



UNIVERSIDAD NACIONAL DE COLOMBIA

Desarrollo de un sistema para la detección e identificación de microplásticos en cuerpos de agua utilizando espectroscopia de impedancia

Juan Daniel Sarmiento Abello

Universidad Nacional de Colombia
Facultad de Ingeniería
Departamento de Ingeniería Eléctrica y Electrónica
Bogotá, Colombia
2024

Desarrollo de un sistema para la detección clasificación de microplásticos en cuerpos de agua utilizando espectroscopia de impedancia

Juan Daniel Sarmiento Abello

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Magister en Ingeniería Electrónica

Director:

Ph.D. Diego Alexander Tibaduiza Burgos

Codirectora:

Ph.D. Maribel Anaya Vejar

Línea de Investigación:

Inteligencia artificial y procesamiento de señales

Grupos de Investigación:

Grupo de investigación en ingeniería electrónica (GMUN)

**Grupo de investigación en electrónica de alta frecuencia y telecomunicaciones
(CMUN)**

Universidad Nacional de Colombia

Departamento de ingeniería eléctrica y electrónica

Bogotá, Colombia

2024

A mi esposa Ana

Agradecimientos

Agradezco a mi esposa Ana, quien ha sido un apoyo incondicional en mi vida. También, quiero dar un especial agradecimiento al ingeniero Diego Alexander Tibaduiza Burgos y a la profesora Maribel Anaya Vejar quienes fueron los directores de esta tesis y me ofrecieron su apoyo, consejo y guía en el desarrollo de los objetivos. Igualmente, agradezco al semillero de Investigación en electrónica, procesamiento de señales e inteligencia artificial, pues allí encontré el apoyo permanente de mis compañeros, así como el espacio y equipos necesario para llevar a cabo los experimentos.

Resumen

Los microplásticos son partículas contaminantes de origen artificial que podrían presentar un riesgo para la salud humana si ingresan a la cadena trófica. Por tal motivo, identificarlos y clasificarlos en diferentes entornos es crucial al momento de ejecutar acciones para prevenir o mitigar su impacto. En consecuencia, en el presente trabajo se desarrolló una metodología que permite evaluar en qué entornos es más probable detectar y clasificar por tamaños - partículas de $500\mu\text{m}$, $1000\mu\text{m}$ y $1400\mu\text{m}$ - dos tipos de microplástico como lo son: el tereftalato de polietileno (PET) y el poliestireno expandido (EPS) en distintos entornos de agua. Debido a la complejidad (alta selectividad) que estos entornos pueden manifestar se escogió una lengua electrónica - en conjunto con una red de sensores -, pues este instrumento es idóneo para detectar elementos en soluciones de alta selectividad. Así las cosas, se evaluaron de forma independiente tres algoritmos de aprendizaje automático, máquinas de soporte vectorial, árboles de decisión y k vecinos más cercanos en ambientes compuestos por: agua potable y tres soluciones conformadas por agua potable y materia orgánica inerte, agua potable, materia orgánica inerte y materia inorgánica y agua potable con materia inorgánica, materia orgánica inerte y materia orgánica viva. Es importante mencionar que los dos tipos de microplástico se midieron de forma independiente en cada uno de los entornos y se obtuvieron rendimientos superiores al 80 % en la detección y clasificación por tamaño.

Palabras clave: (Microplástico, metodología, aprendizaje automático, lengua electrónica, selectividad, contaminantes, inteligencia artificial).

Abstract

Development of a system for the detection and identification of microplastics in bodies of water using impedance spectroscopy

Microplastics are contaminating particles of artificial origin that could present a risk to human health if they enter the food chain, for this reason identifying and classifying them in different environments is crucial when carrying out actions to prevent or mitigate their impact. For this reason in this work, a methodology was developed that allows for evaluating in which environments it is more likely to detect and classify by size, particles of $500\mu\text{m}$, $1000\mu\text{m}$ and $1400\mu\text{m}$, two types of microplastics such as Polyethylene Terephthalate (PET) and Expanded Polystyrene (EPS) in different water environments. Due to the complexity (high selectivity) that these environments can manifest, an electronic language was chosen in conjunction with a network of sensors because this instrument is ideal for detecting elements in highly selectivity solutions. Three machine learning algorithms, support vector machines,

decision trees and k driest neighbors were evaluated independently in environments composed of: drinking water and three solutions consisting of drinking water and inert organic matter. Drinking water, inert organic matter and inorganic matter and drinking water inorganic matter, inert organic matter and living organic matter. The two types of microplastic were measured independently in each of the environments and performances greater than 80 % were obtained in detection and classification by size.

Keywords: (Microplastic, methodology, machine learning, electronic language, selectivity, contaminants, artificial intelligence)

Lista de Figuras

1-1. Representación de un sensor electroquímico, donde R es el receptor y T_R , el transductor. Fuente autor	5
1-2. Arquitectura de una lengua electrónica. Fuente autor	8
1-3. Representación de un arreglo de electrodos utilizado en una lengua impedimétrica. Fuente autor	10
1-4. Representación en el diagrama de Bode (A), diagramas de Nyquist (B) para una lengua electrónica impedimétrica	11
2-1. Arquitectura de la lengua electrónica que se utilizó en el presente trabajo. Fuente autor	16
2-2. Red de sensores[1]	16
2-3. Conexión usada en cada nodo de la red usada en esta tesis. Fuente autor . .	17
2-4. Algunas empresas que fabrican sensores electroquímicos de electrodos serigrafados Screen Printed Electrode (SPE). Fuente, autor	18
2-5. Sensor utilizado en esta tesis	18
2-6. Esquemático interno y dispositivo MUX8-R2 [2]	20
2-7. Esquemático conexión PalmSens4 y MUX8-R2 [3]	21
2-8. Representación de la técnica EIS Diagrama de Nyquist (Derecha) y diagrama de Bode (izquierda), generada por el software PSTrace 5	22
3-1. Resumen de la metodología propuesta para la detección del microplástico. Fuente autor	24
3-2. Filtrado de muestras	25
3-3. Medidas disponibles en el tamiz	25
3-4. Comparativa entre partículas de microplástico y huevos de salmón. Escala 200 μ m. Superior PET, inferior EPS. Fuente autor	26
3-5. Muestra de materia inorgánica	27
3-6. Materia orgánica viva (Izquierda) y materia orgánica inerte (Derecha)	27
3-7. Recipiente utilizado para la preparación de las soluciones.	28
3-8. Arquitectura lengua electrónica usada en el presente proyecto, fuente el autor.	28
3-9. Diagrama de impedancia del agua potable.	29
3-10. Configuración de la lengua electrónica.	30
3-11. Muestras del MP utilizado. Parte superior PET, inferior EPS	31

4-1. Estructura del set de datos obtenido. Fuente autor	33
4-2. Despliegue de los datos.	34
4-3. Valores de Z_x y Z_y	35
4-4. Diferencia entre las técnicas Auto Scaling y Group Scaling.	36
4-5. Componentes principales en los que se distribuye la información.	37
4-6. Distribución del conjunto de datos.	38
4-7. Máquina de soporte vectorial. Fuente autor	39
4-8. Kernel usados por Matlab, tomado de[4].	39
4-9. Iteraciones de KNN. Fuente autor	40
4-10. Algoritmo CART. Fuente autor	41
4-11. Cargue de datos en la CL	42
4-12. Representación de los datos (A). Selección del algoritmo (B).	43
4-13. Representación de los datos (A). Selección del algoritmo (B).	44
5-1. Datos apilados (Superior), datos desplegados (Inferior)	46
5-2. Datos escalados	46
5-3. Distribución de los datos en un espacio bidimensional correspondiente a los dos primeros componentes (A). Porcentaje de la distribución de los datos entre los componentes (B).	47
5-4. Distribución de los datos en un espacio bidimensional correspondiente a los dos primeros componentes en la clasificación del microplástico EPS por tamaño en un entorno de agua potable (A). Porcentaje de la distribución de los datos entre los componentes (B).	48
5-5. Distribución de los datos en un espacio bidimensional correspondiente a los dos primeros componentes en la clasificación del microplástico PET por tamaño en un entorno de agua potable (A). Porcentaje de la distribución de los datos entre los componentes (B).	48
5-6. Distribución de los datos en un espacio tridimensional correspondiente a los tres primeros componentes para la clasificación del microplástico PET y el EPS por tamaño, en un entorno de agua contaminada con materia orgánica inerte(A). Porcentaje de la distribución de los datos entre los componentes (B).	49
5-7. Distribución de los datos en un espacio tridimensional correspondiente a los tres primeros componentes para la clasificación del microplástico PET y el EPS por tamaño, esto en un entorno de agua contaminada con materia orgánica inerte y materia inorgánica(A). Porcentaje de la distribución de los datos entre los componentes (B).	50

5-8. Distribución de los datos en un espacio tridimensional correspondiente a los dos primeros componentes para la clasificación del microplástico PET y el EPS por tamaño, esto en un entorno de agua contaminada con materia orgánica inerte, materia inorgánica y materia orgánica viva(A). Porcentaje de la distribución de los datos entre los componentes (B).	50
5-9. Identificación del microplástico utilizando el algoritmo CART.	52
5-10. Identificación del microplástico utilizando el algoritmo SVM.	52
5-11. Identificación del microplástico utilizando el algoritmo SVM.	53
5-12. Clasificación de microplástico PET en agua potable usando el algoritmo árboles de decisión, con una precisión del 95.8 %.	54
5-13. Clasificación de microplástico PET en agua potable usando el algoritmo máquinas de soporte vectorial, con una precisión del 100 %	54
5-14. Clasificación de microplástico PET en agua potable usando el algoritmo K como vecino más cercano, con una precisión del 100 %	55
5-15. Clasificación de microplástico EPS en agua potable usando el algoritmo árboles de decisión, con una precisión del 95.8 %.	56
5-16. Clasificación de microplástico PET en agua potable usando el algoritmo máquinas de soporte vectorial, con una precisión del 87.5 %	56
5-17. Clasificación de microplástico PET en agua potable usando el algoritmo K como vecino más cercano, con una precisión del 91.7 %	56
5-18. Clasificación de microplástico EPS y PET en agua con materia orgánica inerte utilizando el algoritmo árboles de decisión, con una precisión del 91.7 %. . . .	57
5-19. Clasificación de microplástico EPS y PET en agua con materia orgánica inerte utilizando el algoritmo máquinas de soporte vectorial, con una precisión del 79.2 %	57
5-20. Clasificación de microplástico EPS y PET en agua con materia orgánica inerte utilizando el algoritmo K como vecino más cercano, con una precisión del 87.5 %	58
5-21. Clasificación de microplástico EPS y PET en agua con materia orgánica inerte y materia inorgánica utilizando el algoritmo árboles de decisión, con una precisión del 83.3 %.	59
5-22. Clasificación de microplástico EPS y PET en agua con materia orgánica inerte y materia inorgánica utilizando el algoritmo máquinas de soporte vectorial, con una precisión del 95.8 %	59
5-23. EPS y PET en agua con materia orgánica inerte y materia inorgánica utilizando el algoritmo K como vecino más cercano, con una precisión del 91.7 %	59
5-24. Clasificación de microplástico EPS y PET en agua con materia orgánica inerte, materia inorgánica y materia orgánica viva utilizando el algoritmo árboles de decisión, con una precisión del 100 %.	61

- 5-25.**Clasificación de microplástico EPS y PET en agua con materia orgánica inerte, materia inorgánica y materia orgánica viva utilizando el algoritmo máquinas de soporte vectorial, con una precisión del 100 % 61
- 5-26.**Clasificación de microplástico EPS y PET en agua con materia orgánica inerte, materia inorgánica y materia orgánica viva utilizando el algoritmo K como vecino más cercano, con una precisión del 100 % 61

Lista de Tablas

1-1. Resumen de los diferentes tipos de lengua electrónica ET y sus aplicaciones .	13
2-1. Material electrodos de trabajo en la red de sensores.	19
2-2. Características de los sensores utilizados en el presente trabajo.	20
3-1. Tamaños de microplástico seleccionados para realizar los experimentos. . . .	26
3-2. Escenarios propuestos para la validación de a metodología	29
4-1. Relación de los diferentes modelos que utiliza KNN para calcular la distancia.	40
5-1. Componentes utilizados para la óptima distribución de los datos utilizando PCA	51
5-2. Resultados generales para la detección del microplástico. El parámetro precisión se mide por clase y Accuracy mide el rendimiento general del algoritmo	53
5-3. Resultados generales para la clasificación del microplástico PET en agua potable por tamaño.	55
5-4. Resultados generales para la clasificación del microplástico EPS en agua potable por tamaño.	57
5-5. Resultados generales para la clasificación del microplástico EPS y PER en agua con materia orgánica inerte por tamaño.	58
5-6. Resultados generales para la clasificación del microplástico EPS y PET, por tamaño, en una solución compuesta por agua, materia orgánica inerte y materia inorgánica.	60
5-7. Resultados generales para la clasificación del microplástico EPS y PET, por tamaño, en una solución compuesta por agua, materia orgánica inerte y materia inorgánica.	62
B-1. Abreviaturas	74

1. Introducción

Los microplásticos son partículas antropogénicas cuya presencia en el agua representa uno de los grandes retos medioambientales a los que se enfrenta la humanidad, esto debido a los posibles impactos en la salud de las personas[5] y en los ecosistemas [6]. Estos materiales han sido detectados en cuerpos de agua utilizada para el consumo humano [7] y en algunos alimentos de consumo frecuente como la sal[8]. Su origen se debe a la fragmentación causada por la interacción con su medio circundante como, por ejemplo: la radiación solar y la fricción con el agua, rocas y materia orgánica[9]. También a su implementación en la industria cosmética[10] y textil[11]. Sin duda, la detección y cuantificación de estos micro materiales son los principales obstáculos al momento de crear soluciones para mitigar su impacto. La detección e identificación de microplásticos empleando la espectroscopia de impedancia utiliza un potencióstato y un clúster de sensores conectados a un multiplexor, este sistema puede obtener los valores de impedancia asociados a las micropartículas en un medio específico, como lo es el agua. Cada elemento por detectar tiene un perfil de impedancia que se puede obtener al aplicar una onda senoidal $V_0 \sin(\omega t)$ a diferentes frecuencias, lo que permite discriminar posibles concentraciones de otros materiales. La representación y análisis de la información se realiza utilizando diagramas de Nyquist, diagramas de Bode y aproximaciones con circuitos eléctricos equivalentes. Debido a que cada sensor realiza una representación individual y/o en varias de las mencionadas anteriormente, se requiere de técnicas para el procesamiento de la información que permita obtener la agrupación de los parámetros relevantes con el objetivo de detectar e identificar el microplástico, filtrando así posibles errores causados por elementos que tengan perfiles similares a estas micropartículas. Se sabe que la espectroscopia de impedancia es una técnica fiable para la detección e identificación de microplásticos en el agua [12], por lo cual el objetivo principal de este proyecto es desarrollar una metodología que permita determinar si hay microplásticos en cuerpos de agua dulce, así como poder identificar qué tipo de estas micropartículas están presentes en estos cuerpos. Para esto se evaluarán diferentes tipos de sensores en una variedad de entornos, con miras a identificar cuál presenta la menor resistencia a la transferencia de carga. La idea es evaluar su sensibilidad en diferentes ambientes, y a medida que las partículas son de menor tamaño.

1.1. Objetivos

A continuación se presentan los objetivos planteados en esta tesis:

Objetivo general:

Desarrollar un sistema para la detección y clasificación de microplásticos en cuerpos de agua dulce basado en un sistema de lengua electrónica.

Objetivos específicos:

- Definir una red de sensores basada en el criterio de una elevada selectividad cruzada.
- Evaluar diferentes clasificadores en entornos distintos, con el fin de determinar cuál presenta mejor rendimiento.
- Desarrollar y evaluar una metodología para la detección y clasificación de microplásticos centrada en Polietileno Tereftalato PET y el Poliestireno Expandido (EPS).

1.1.1. Resultados obtenidos

A continuación se presenta un resumen de los resultados obtenidos por cada uno de los objetivos anteriormente relacionados;

Objetivo 1. Definir una red de sensores basada en el criterio una elevada selectividad cruzada.

Se efectuó un análisis comparativo con diferentes tipos de sensores electroquímicos para determinar qué red de sensores se adapta mejor a los requerimientos del presente trabajo. Para lo anterior, se hizo un estudio sobre los diferentes tipos de sensores electroquímicos y sus aplicaciones, igualmente se realizó un pequeño estudio de mercado con el fin de validar su disponibilidad, en caso de ser requerida. Esto se muestra en el capítulo 2.

Objetivo 2. Evaluar diferentes clasificadores en entornos distintos con el fin de determinar cuál presenta mejor rendimiento.

Utilizando una lengua electrónica como sistema para la adquisición de datos y la técnica electroquímica de espectroscopia de impedancia, se evaluaron tres algoritmos de clasificación, a saber: Máquina de soporte vectorial (SVM), Árboles de decisión y K vecinos más cercanos (KNN), con el fin de detectar y clasificar dos tipos de microplástico en diferentes entornos de agua, con una exactitud mayor 80 %. Estos resultados se incluyen en capítulo 5: Resultados.

Objetivo 3. Desarrollar y evaluar una metodología para la detección y clasificación de microplásticos centrada en Polietileno Tereftalato PET y el Poliestireno Expandido (EPS).

En el capítulo 3 se presenta el diseño y desarrollo de una serie de experimentos que simularon diferentes entornos ambientales, empezando por agua potable, agua contaminada con materia orgánica inerte, materia orgánica viva y materia inorgánica. Los materiales utilizados para los expertos se agregaron en concentraciones y tamaños equivalentes a los microplásticos utilizados. Se realizó la adquisición de datos en cada uno de los entornos con diferentes tipos de microplástico (Polietileno Tereftalato PET y el Poliestireno Expandido (EPS)), con el objetivo de identificar qué entorno presenta la mayor dificultad para los algoritmos al momento de detectar y clasificar las partículas de microplástico.

Algunos de los resultados del presente trabajo se expusieron en los siguientes eventos:

- Microplastic detection in water using a sensor network, an electronic tongue and spectroscopy impedance ISBN - 978-989-9121-25-6. Event MILAN 24th International Conference on “Innovations in Science & Technology” (MIST-23) in International center of excellence in engineering & technology (Iceeat) Milan (Italy).
- Clasificación de microplásticos usando lenguas electrónicas <https://doi.org/10.26507/paper.3126>. Evento EIE ACOFI 2023 ingeniería para transformar territorio, realizado en Cartagena de Indias (Colombia).
- Electronic Tongues and Noses: A General Overview. <https://doi.org/10.3390/bios14040190>. Jurnal biosensors, editorial Multidisciplinary Digital Publishing Institute (MDPI)

De igual manera, se encuentran en revisión los siguientes dos artículos de revista.

- Microplastic identification using impedance spectroscopy and machine learning algorithms, International Journal of Distributed Sensor Networks. SAGE.

1.2. Marco teórico

1.2.1. Lenguas electrónicas y arreglos de sensores

Para hablar de lenguas electrónicas primero hay que discutir sobre los sensores electroquímicos, esto debido a que en el contexto del presente trabajo una lengua electrónica está definida como un arreglo de sensores de tipo electroquímico conectados a un sistema de adquisición de datos y a un elemento para el procesamiento de la información. Por lo que primero, se expondrá acerca de su funcionamiento, luego se profundizará en los tipos de sensores y se discutirá con mayor detalle los sensores utilizados en las lenguas electrónicas.

1.2.2. Sensor electroquímico

Un sensor electroquímico es un dispositivo que transforma la información proveniente de una reacción electroquímica, óptica, térmica o másica en una señal eléctrica[13].

Está compuesto fundamentalmente por: (i.) un elemento que es selectivo, el cual puede ser de tipo película de polímero molecular(*molecularly imprinted polymer*, MIP)[14], electrodos de iones selectivos(*ion selective electrodes*, ISE)[15], película con algún tipo de biomaterial o nano material como el grafeno[16], enzima o anticuerpos. (ii.) Un elemento instrumental, que traduce la reacción del fenómeno que esté ocurriendo en una señal eléctrica. A la sustancia que interacciona con el sensor se le conoce como analito el nivel de esta interacción se le conoce como selectividad.

La selectividad hace referencia a la capacidad del sensor para medir determinado analito en una muestra específica compuesta por diferentes elementos. La figura 1-1 representa los componentes de un sensor electroquímico.

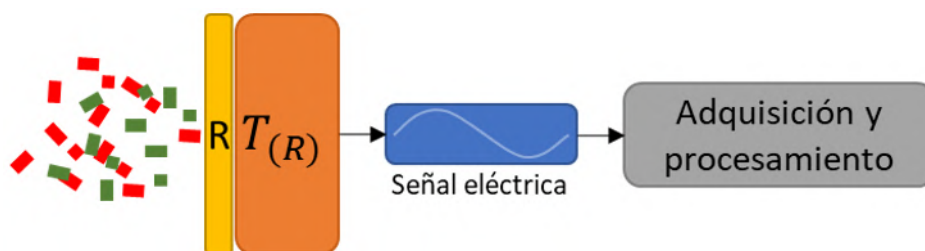


Figura 1-1.: Representación de un sensor electroquímico, donde R es el receptor y T_R , el transductor. Fuente autor

Cuando una partícula (analito) se desplaza por la zona de detección del sensor electroquímico se produce un estímulo en el receptor (Electrodo de trabajo), dicho estímulo es transformado por el transductor en una señal eléctrica. Esta señal es enviada a un sistema de adquisición y procesamiento de datos. La integración de estos dos elementos (Componente selectivo + transductor) son la base de los sensores electroquímicos.

Con el desarrollo de estos sensores se han logrado importantes resultados en la química analítica, principalmente por las ventajas que presentan al realizar medidas *in situ*, su relativo bajo costo de fabricación, su amplio intervalo de linealidad, así como un corto tiempo de respuesta. De igual manera, el avance en la tecnología de materiales ha permitido altos niveles de selectividad, esto les da la capacidad de trabajar en soluciones con diferentes niveles de concentración.

1.3. Tipos de sensores electroquímicos

Para poder realizar una adecuada clasificación de los sensores electroquímicos primero hay que introducir algunas definiciones.

1.3.1. Sensores activos y pasivos

Los sensores activos generan una señal que interactúa con el medio, emitiendo algún tipo de energía y midiendo las propiedades de la respuesta obtenida. Un ejemplo de este tipo de sensor se puede ver en los sistemas de radar[17], o los piezoeléctricos utilizados en monitoreo de salud estructural[18], aunque normalmente estos sensores no cambian las características del medio. Existen sensores activos que sí podrían alterar las propiedades del objeto de estudio (por diseño) con el objetivo de llevarlo a cierto nivel de excitación, en esta categoría están los sensores electroquímicos de los cuales se profundizará más adelante.

Los sensores pasivos solamente adquieren información si el objeto a medir entra en su rango de detección, generalmente se basan en la medición de una variable física como en campo eléctrico, magnético, la corriente y el voltaje. Estas variables también son usadas para medir de forma indirecta otros fenómenos como la masa, presión o velocidad de un objeto.

1.3.2. Sensores absolutos o relativos

En los sensores absolutos o de deflexión[19] la señal de salida está relacionada con el cambio de alguna propiedad física, como frecuencia o deformación, la variación en la magnitud de dicha propiedad es la variable que se mide, un ejemplo son las galgas extensiométricas donde el cambio en la deformación, presión, carga o par ejercido sobre la galga produce un cambio en su resistencia intrínseca. En los sensores relativos la señal de salida depende de un valor de referencia, donde el objetivo es medir el valor diferencial entre la señal medida y el nivel de referencia, un ejemplo de este tipo de sensor son las celdas electroquímicas de tres electrodos. Los sensores electroquímicos en general pertenecen a la categoría de sensores activos y relativos, debido a que envían una señal que interactúa con el analito y normalmente requieren de un voltaje de referencia controlado en uno de sus electrodos.

A continuación se realiza una descripción de los principales sensores electroquímicos.

1.3.3. Sensores potenciométricos

En estos sensores existe una diferencia de potencial entre el electrodo de trabajo y electrodo de referencia debido a un intercambio iónico que ocurre en la superficie del sensor y la solución que se mide. La diferencia de potencial se obtiene a través de la ecuación de Nernst.

$$E = E^0 - \frac{RT}{nF} \ln(Q) \quad (1-1)$$

Donde E es el potencial corregido del electrodo, E^0 es el potencial en condiciones estándar, R es la constante de los gases, T es la temperatura, n es número de moles, F es la constante de Faraday y Q es el cociente de reacción correspondiente.

Dentro de esta clasificación de sensores potenciómetros también existen algunas subclases que vale la pena mencionar, a saber:

Electrodos selectivos de iones

Utilizados en medición de acidez en soluciones, en este tipo de sensores su selectividad depende principalmente de los materiales que conforman la membrana donde ocurre la reacción química, estos pueden ser electrodos de vidrio, líquido o sólido.

Electrodos de alambre recubierto

Este sensor está compuesto por un alambre conductor cubierto por una membrana de algún tipo de polímero como el cloruro de polivinilo. Este tipo de sensor no requiere un electrodo de referencia por, lo que se podría clasificar como un sensor activo y absoluto.

Transistores de efecto de campo

Estos sensores son de electrodos de alambre recubierto, con la diferencia de que la membrana selectiva se encuentra incorporada en la superficie de la compuerta en transistores tipo FET. Debido a su alta impedancia de entrada y baja impedancia de salida, los sensores electroquímicos basados en transistores FET son ideales para monitorear la carga en una membrana de ion selectivo. Es importante mencionar que la fabricación de estos sensores se realiza con los procesos de fabricación de semiconductores y por ello se pueden implementar arreglos con diferentes tipos de membranas selectivas.

1.3.4. Sensores amperométricos

Este tipo de sensor se basa en el intercambio de electrones entre el electrodo de trabajo y el analito o un potencial dado. Si dicho potencial permanece constante, el modo de detección se denomina detección potencioestática y si fluctúa, típicamente entre dos valores, se denomina detección potenciodinámica [20]. Hay que agregar que al aplicar algún tipo de material biológico de reconocimiento molecular sobre la superficie electroactiva del electrodo se crea

la clasificación de los biosensores. En los biosensores el elemento transductor se le conoce como biotransductor y se encarga de transformar la reacción química o biológica en una señal eléctrica u óptica interpretable.

1.3.5. Sensores conductimetricos

Estos sensores miden los cambios en la conductividad del analito en una solución específica o en la membrana selectiva, debido a que la conductividad depende de la estructura atómica del material los sensores conductimetricos son una buena opción para la detección de metales, aunque también se pueden encontrar en biotransductores [21].

1.4. Lenguas electrónicas

Una lengua electrónica ET es un dispositivo formado por una red de sensores conectados a un equipo de medición electroquímica, generalmente un potenciostato, integrado con un sistema de adquisición y procesamiento de datos, figura 1-2. Su objetivo es simular el comportamiento de la lengua humana. Sin embargo, su uso se ha extendido por diferentes áreas de la ciencia y la ingeniería y aunque aún cuenta con una presencia importante en el sector de alimentos[22][23][24], el uso del sensor de sabor tiene un enfoque más óptimo en la detección de sabores[25].

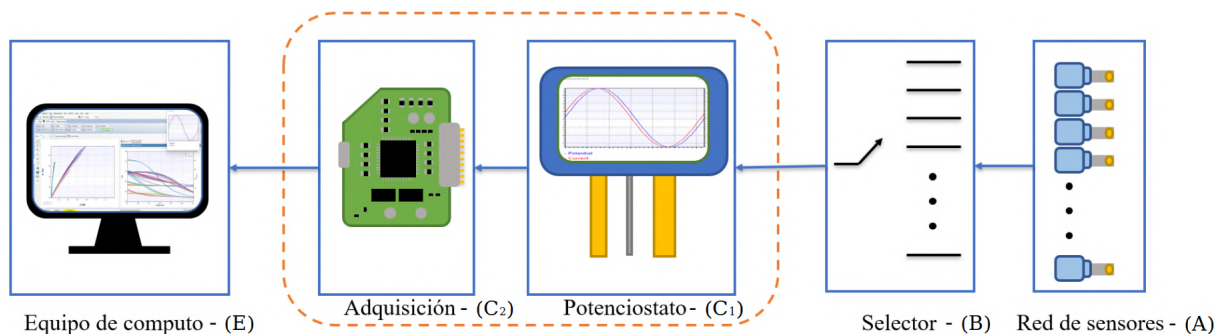


Figura 1-2.: Arquitectura de una lengua electrónica. Fuente autor

Los aspectos fundamentales que caracterizan a las lenguas electrónicas son su selectividad global, la cual está asociada a un bajo nivel de selectividad en los electrodos y un alto rango de selectividad cruzada. La selectividad cruzada (también conocida como sensibilidad cruzada) es un término que expresa la capacidad de un electrodo para detectar una sustancia específica, si el electrodo detecta parcialmente una sustancia o presenta un valor de selectividad cercano para analitos diferentes, se puede decir que tiene una alta sensibilidad cruzada.

1.4.1. Principio de funcionamiento

A continuación se describe el principio de funcionamiento del sistema ET utilizado en el presente trabajo.

El funcionamiento de una ET empieza con la detección del analito realizado por la red de sensores figura 1-2-A, para lograr una baja selectividad se utilizaron electrodos de diferente material. Debido a que las señales enviadas al potencióstato pueden interferirse mutuamente la ET cuenta con un Mux 1-2-B, el cual selecciona el sensor que realiza la medición según un orden preestablecido, la señal que es recibida por el potencióstato 1-2-C₁ dependerá de la técnica electroquímica utilizada, esta puede ser una señal de voltaje, corriente o una relación impedimétrica o conductimétrica. Por último, la información es obtenida por un sistema de adquisición y procesamiento de señales 1-2-C₂, donde se discretiza y envía a un equipo de cómputo 1-2-E para su análisis. Por último se implementan técnicas de machine learning para detección y/o la clasificación del analito objetivo.

En la siguiente sección se relacionan las diferentes arquitecturas que puede presentar un sistema ET.

1.4.2. Lengua electrónica volumétrica (VTI)

Esta lengua se basa en el principio de voltamperometría para realizar el análisis electroquímico, entre los diferentes arreglos de sensores que se pueden utilizar con esta (VTI), están los electrodos de trabajo de tipo metálico como el oro, iridio, platino y rodio[26], en esta configuración se puede utilizar un único electrodo de referencia y es una de las más comunes, a esta configuración también se le conoce como electrodos desnudos[27]. Esta configuración resultó siendo bastante robusta por lo que ha logrado mantenerse a lo largo del tiempo, una de sus aplicaciones está en el estudio sobre la detección y el crecimiento de especies microbianas[28], o la evaluación de algoritmos para la clasificación de té[29]. Desarrollos más recientes utilizan diferentes tipos de electrodos modificados con nanocompuestos en la identificación de alimentos o bebidas fermentadas[30], de igual manera el uso de biosensores[31] ha permitido avances importantes en la selectividad cruzada de las lenguas.

1.4.3. Lengua electrónica potenciométrica (PET)

También existen lenguas electrónicas basadas en arreglos de electrodos potenciométricos, algunas cuentan con electrodos de membrana policloruro de vinilo[32], empleadas en el seguimiento de parámetros fisicoquímicos del agua, o los usados en el análisis del aceite virgen de oliva[33], utilizando electrodos de plata (Ag-AgCl), así como las lenguas basadas en sensores de estado sólido con un contacto sólido hecho de un compuesto epoxi conductor[34] para la clasificación de cerveza. La PET también tiene aplicaciones en la detección de metales

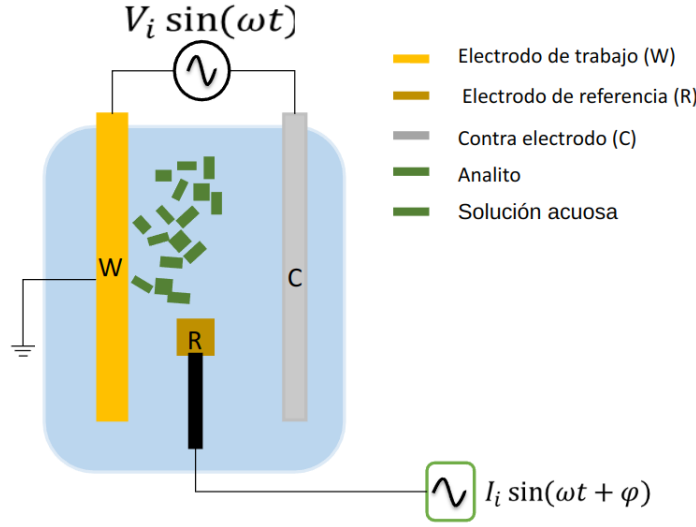


Figura 1-3.: Representación de un arreglo de electrodos utilizado en una lengua impedimétrica. Fuente autor

pesados utilizando electrodos selectivos a iones (ISE) de vidrio de calcogenuro[35] así como en la clasificación de mieles[36].

1.4.4. Lengua electrónica impedimétrica (IET)

Este tipo de IET tiene su uso en la identificación de analitos y la caracterización de sus propiedades fisicoquímicas utilizando parámetros eléctricos en el dominio de los números complejos, adicionalmente es la configuración de IET utilizada en el presente trabajo. Al aplicar una señal de voltaje de baja amplitud (del orden de los milivoltios) en un rango de frecuencias, a un arreglo de electrodos sumergidos en una solución acuosa que contenga uno o varios analitos, como se muestra en la figura 1-3, se puede obtener una respuesta respecto a un punto referencia como una corriente con amplitud y fase distintas a la señal de entrada[37]. La relación entre la entrada y la salida se define en la ecuación (1-2), donde Z es un número complejo que representa una propiedad del analito a una frecuencia y voltaje específico.

$$Z = \frac{V}{I} = \frac{V_i \sin(\omega t)}{I_i \sin(\omega t + \varphi)} = Z_0 \frac{\sin(\omega t)}{\sin(\omega t + \varphi)} \quad (1-2)$$

La manera en que la IET representa la información es variada, con diagramas de Nyquist, diagramas de Bode o presentación de circuitos eléctricos equivalente, esto dota al sistema de diferentes formas de representar la información, lo cual puede ser útil a la hora de realizar el

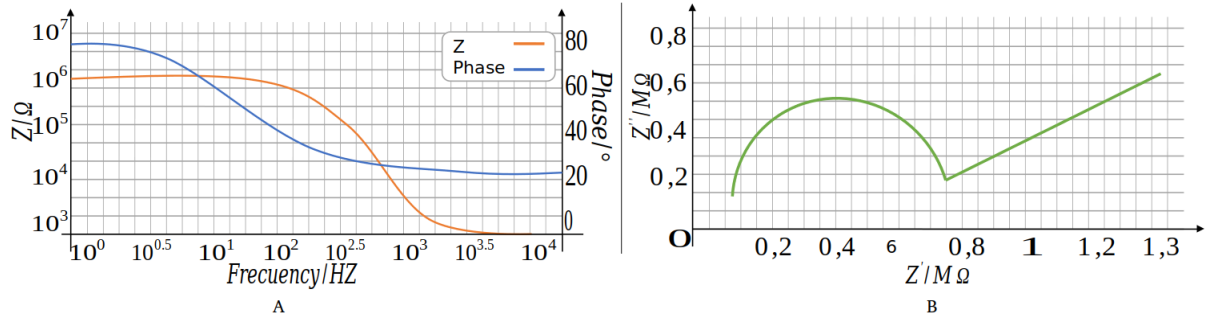


Figura 1-4.: Representación en el diagrama de Bode (A), diagramas de Nyquist (B) para una lengua electrónica impedimétrica

análisis de los datos, como lo es el estudio en el dominio de la frecuencia[38], dándole acceso a ingenieros y científicos de diferentes disciplinas a las IET en general. En la figura 1-4 se pueden ver algunas de las formas de representar la información en este tipo de lengua.

Al igual que la PET, la IET también se emplea en la identificación del origen botánico de la miel[39], utilizando una configuración de electrodos compuesta por oro, carbono, óxido de indio, estaño y silicio dopado con el objetivo es obtener el origen florar de la miel. De igual forma, las IET se han usado en la identificación de sabores[40] usando arreglos de electrodos compuestos por una matriz a nano escala conocida como electrodo interdigitado desnudo, esta matriz de electrodos también puede estar recubierta con ácido esteárico, un aligonero de polianilina y polipirrol. Otra aplicación donde las IET han obtenido interesantes resultados es en la industria láctea[41] donde se utilizan electrodos modificados con nanopartículas de plata y diferentes tipos de enzimas, igualmente se han implementado en el monitoreo de variables ambientales[21] presentando importantes ventajas debido al uso de microtransductores conductimétricos en biosensores, esto les permite eliminar el electrodo de referencia. De igual forma, las mediciones en modo diferencial son óptimas al momento de requerir un alto nivel de atenuación.

1.4.5. Lengua electrónica híbrida (HET)

Esta clase de lengua hace uso de diferentes tipos de sensores en un solo sistema, como sensores potenciométricos, conductimétricos y/o amperométricos, adicionalmente entre las distintas clases de sensores también pueden existir electrodos de diferentes compuestos. El uso de estos dispositivos puede incidir mejoras para los algoritmos de clasificación. Las (HET) han sido probadas con éxito en la evaluación de parámetros organolépticos en frutas[42] utilizando un conjunto de sensores potenciométricos y voltamétricos y las técnicas de amperometría y conductimetría. Una arquitectura similar es usada para realizar el monitoreo de bioprocesos[43] como el 2-feniletanol, así como el seguimiento en el proceso de fermentación para la cerveza[44], donde se utiliza una red de 10 electrodos miniaturizados de iones selec-

tivos para el sistema potenciométrico, también transductores voltamétricos de 3 electrodos basados en silicio para el sistema voltamétrico. Un tipo de (HET) es la basada en sistemas optofluídico calorimétrico en conjunto con una red de sensores potenciométricos con tecnología ISFET (sensor potenciométrico basado en el transistor de efecto de campo de iones selectivos), modificados con membranas poliméricas sensibles a los iones y una pequeña red de electrodos amperométricos basados en microelectrodos de oro (Au)[45]. Este sistema es utilizado en la caracterización y cuantificación de diferentes vino tintos de uva. Una configuración que debería ser común es la (HET) basada en potencimetría, voltamperometría y conductividad, utilizada en la clasificación de diferentes tipos de leche fermentada[46], compuesta por una red de sensores de oro, iridio, paladio, platino, renio y rodio en conjunto con un electrodo de referencia Ag/AgCl.

En la tabla 1-1 se muestra un resumen de los diferentes tipos de ET según su aplicación y el tipo de sensor utilizado.

En el siguiente capítulo se abordan las técnicas utilizadas en el presente trabajo para el preprocesado de datos provenientes de la lengua electrónica, así como los algoritmos que han mostrado los resultados en la clasificación del microplástico.

1.5. Preprocesado de datos y algoritmos de clasificación

En la siguiente sección se describen los principales aspectos a tener en cuenta al momento de realizar un análisis de datos multivariados y/o multidimensionales (en quimiometría) debido a que es de vital importancia que la información ingresada a los algoritmos de clasificación cuente con niveles bajos o nulos de ruido, así como un bajo nivel de dimensionalidad con el fin de evitar resultados sesgados o buenos rendimientos solo a nivel de entrenamiento.

1.5.1. Reducción de ruido

Los datos obtenidos en el mundo real generalmente no tienen la forma requerida por los algoritmos de clasificación, por lo que se debe realizar un preprocesado con el objetivo de mejorar su calidad.

En análisis de datos quimiométricos el ruido hace referencia a fluctuaciones pequeñas, rápidas y aleatorias a la salida del sistema, en el presente trabajo estas fluctuaciones están dadas por las respuestas en el arreglo de sensores. Existen varias técnicas que pueden ayudar con el tratamiento del ruido como despliegue de datos multidimensional[47]; a saber, la eliminación de líneas base, la alineación de picos, la selección de picos y el escalado[48]. Sin embargo, la utilización de estas técnicas generalmente consume la mayor parte del tiempo y recursos computacionales en el procesamiento total de los datos.

Tipo de ET	Arreglo de sensores	Aplicación
Volumétrica (VET)	Electrodos de oro, iridio, platino y rodio Nanocompuestos Biosensores	Crecimiento de especies microbianas Clasificación de té Identificación de alimentos y bebidas Identificación de alimentos y bebidas
Potenciométrica PET	membrana policloruro de vinilo Plata (Ag-AgCl) FET de epoxi conductor Vidrio de calcogenuro	Seguimiento de parámetros físicoquímicos del agua Análisis aceite virgen de oliva Clasificación de cerveza Detección de metales pesados
impedimétrica IET	oro, carbono, óxido de indio, estaño y silicio Electrodos interdigitados recubiertos con ácido esteárico con aligonero de polianilina y polipirrol Nanopartículas de plata Microtransductores conductimétricos en biosensores	Identificación de mieles Identificación de sabores Parametrización de leche Monitoreo de variables ambientales
híbrida (HET)	Potenciométricos y voltamétricos Potenciométricos y voltamétricos Electrodos miniaturizados de iones selectivos ISFET (sensor potenciométrico basado en el transistor de efecto de campo de iones selectivos) mas microelectrodos de oro	Identificación de parámetros organolépticos en frutas Control de bioprocesos en levaduras Fermentación para la cerveza Caracterización y cuantificación diferentes vinos tintos de uva

Tabla 1-1.: Resumen de los diferentes tipos de lengua electrónica ET y sus aplicaciones

Las técnicas puntuales utilizadas en el presente trabajo se presentarán en secciones posteriores.

1.5.2. Reducción de dimensionalidad

Comprende un conjunto de técnicas que se utilizan para reducir la alta dimensionalidad en los datos en sistemas quimiométricos, como los generados por lenguas electrónicas[49]. Entre los métodos más comúnmente utilizados están: análisis de componentes principales (PCA), transformada rápida de Fourier, transformada wavelet discreta y la integral de corte en ventana de voltamograma[50]. Sin embargo, para el caso en el que los datos tienen un alto nivel de no linealidad son útiles técnicas como fusionar la información global y local con el objetivo de desplegar datos altamente no lineales[51], o el análisis de mínimos cuadrados trilineales[52]. Para el caso de las lenguas electrónicas, debido a que tienen un alto valor de selectividad cruzada, la información resultante está representada por muchas variables, por lo que la técnica de PCA es una buena opción para identificar qué variables representan mejor la información de los datos[53].

1.5.3. Algoritmos de clasificación

Luego del preprocesado inicial, los datos deben contar con un nivel de ruido suficientemente atenuado y tener una forma geométrica que los algoritmos de clasificación puedan aceptar como entrada. En el capítulo 4 se describe en profundidad los algoritmos utilizados en el presente trabajo.

2. Selección del potencióstato y la red de sensores

En este capítulo se describe de forma general la arquitectura de la lengua electrónica utilizada en esta tesis, también se realiza una descripción general del equipo, el cual es un potencióstato de la marca PalmSens referencia PalmSens4, igualmente se definen las características de los electrodos empleados en la red de sensores.

2.1. Arquitectura de la lengua electrónica

La arquitectura de lengua electrónica utilizada en la presente tesis está conformada por cuatro grandes bloques: primero, arquitectura de lengua electrónica; segundo, red de sensores; tercero, multiplexor; cuarto, analizador electroquímico PalmSens4. Es importante comenzar diciendo que una red de sensores está compuesta por ocho electrodos, de tres terminales, conectados a un nodo común, figura **2-1-A**. La señal que envía el potencióstato hacia la red de sensores solo puede ser recibida por un sensor a la vez, esto implica que en la respuesta de la red también se da un sensor a la vez, por lo que se debe agregar un elemento que seleccione el sensor que va a recibir la señal del potencióstato. Adicionalmente, esta selección se debe realizar en un orden específico. De esto se encarga el bloque de selección o Mux, figura **2-1-B**. El Mux funciona como puente entre la red de sensores y el potencióstato, este último es el instrumento que se encarga de enviar la señal de estímulo y recibir su respuesta **2-1-C**. Según la técnica electroquímica con la que esté configurado envía un tipo de respuesta al siguiente bloque, el cual es un sistema de cómputo que procesa las señales **2-1-D**.

A continuación se realiza una descripción detallada de cada uno de los bloques.

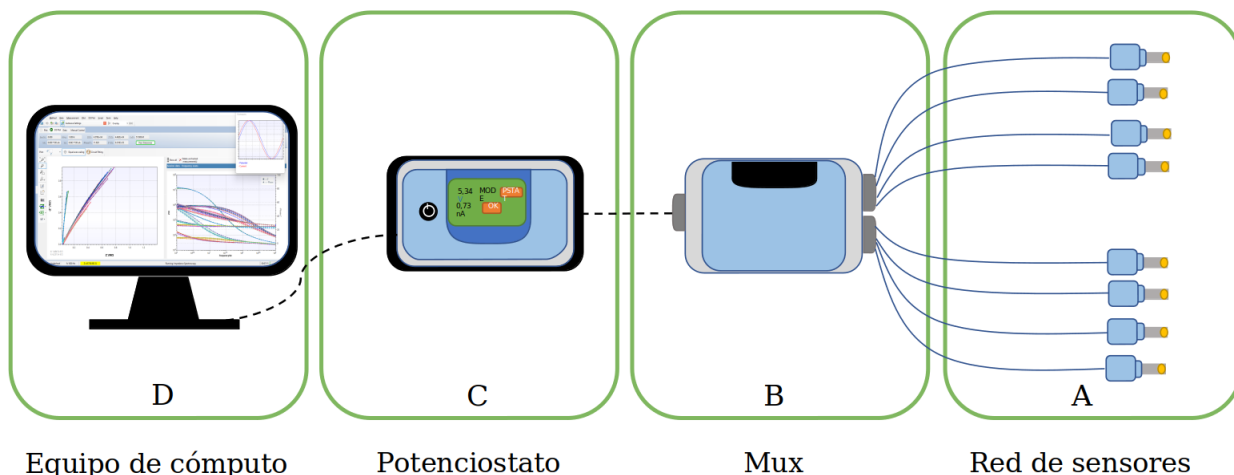


Figura 2-1.: Arquitectura de la lengua electrónica que se utilizó en el presente trabajo.
Fuente autor

2.2. Red de sensores

La red de sensores está compuesta por dos conjuntos de cuatro nodos unidos a un conector tipo D-Sub central, como se muestra en la figura 2-2-A. Cada nodo está compuesto por seis terminales, figura 2-2-B, la conexión final varía en función del experimento a realizar. Para esta tesis el acoplamiento se muestra en la figura 2-3.

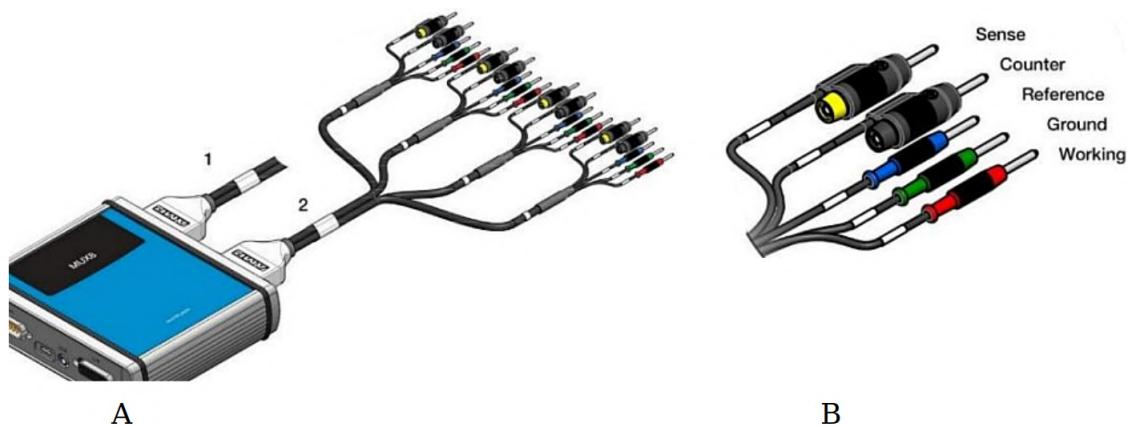


Figura 2-2.: Red de sensores[1]



Figura 2-3.: Conexión usada en cada nodo de la red usada en esta tesis. Fuente autor

Uno de los elementos más importantes de una lengua electrónica, sin duda, son los electrodos (Sensores electroquímicos), por lo que su elección requiere un análisis previo, con el objetivo de identificar el que mejor se adapte a la aplicación. A continuación se muestra un resumen de los diferentes tipos de electrodos utilizados comercialmente y la elección escogida para el presente trabajo.

2.2.1. Sensores electroquímicos basados en electrodos serigrafiados (SPE)

Como resultado del avance en las técnicas de miniaturización y la escalabilidad de la tecnología, se desarrolló una arquitectura donde los diferentes tipos de electrodos se encuentran impresos sobre un sustrato que no interactúa con el analito. Su aplicación está centrada en plataformas de detección para el análisis clínico, medioambientales e industriales [54].

En el presente trabajo se utilizaron los electrodos SPE para la identificación del microplástico en diferentes entornos. La marca seleccionada es BVT technologies. Sin embargo, en la industria existen varias marcas, como se muestra en la figura 2-4.

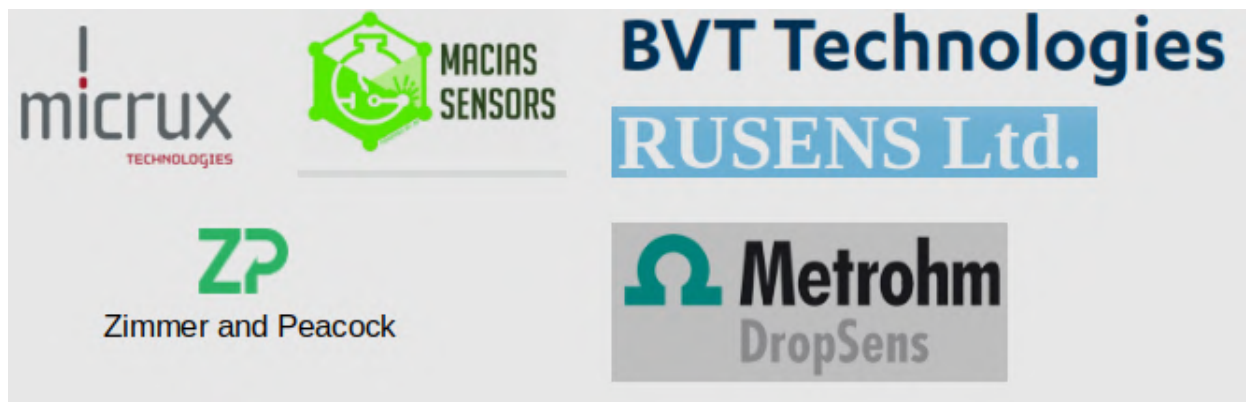


Figura 2-4.: Algunas empresas que fabrican sensores electroquímicos de electrodos serigrafados Screen Printed Electrode (SPE). Fuente, autor

La razón por la que se escogió este tipo de electrodos es que tienen un factor de forma pequeño, lo cual permite un transporte y uso general óptimo. Respecto a la marca elegida, su elección se debe a su relativa buena disponibilidad en el mercado, así como a la variedad de materiales que ofrece para sus electrodos de trabajo. La figura 2-5 muestra el tipo de sensor utilizado.



Figura 2-5.: Sensor utilizado en esta tesis

2.2.2. Selección de los electrodos

Para lograr un alto rango de selectividad cruzada es relevante que los electrodos de trabajo sean de diferentes materiales, por lo que se eligió una red de sensores de diferente material en

el electrodo de trabajo en configuración con tres terminales (electrodo de referencia, contra electrodo y electrodo de trabajo) impresas sobre un sustrato de corindón [55], este es un material cerámico que no interactúa químicamente con el analito o la solución. En la tabla se muestra la configuración de la red.

Posición del sensor	Electrodo de trabajo	Electrodo referencia
1	Oro	AC1.W1.R1 DW=2
2	Platino	AC1.W2.R2 DW=2
3	Grafito	AC1.W4.R2 DW=2
4	Plata	AC1.W3.R2 DW=2
5	Aleación de oro y platino	AC1.WS.R2 DW=2
6	Oro	AC1.W1.R1 DW=2
7	Oro	AC1.W1.R1 DW=2
8	Platino	AC1.W2.R2 DW=2

Tabla 2-1.: Material electrodos de trabajo en la red de sensores.

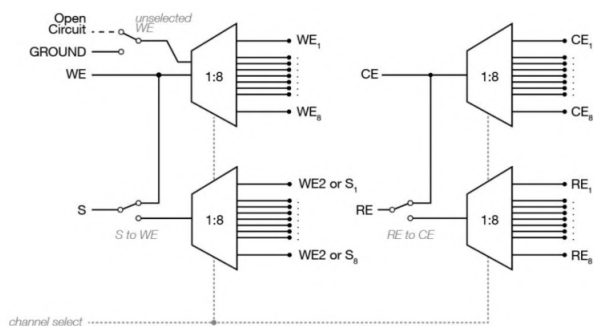
La tabla **2-1** está compuesta por la posición del sensor en la red, columna número uno, el electrodo de trabajo y la referencia del sensor, suministrada por el fabricante, en las columnas dos y tres. Para entender mejor las características de cada elemento se va a realizar una descripción de su referencia. Esta se puede dividir en cuatro secciones como se muestra en la tabla **2-2**. Donde ACn, con n = grupo de sensores, hace referencia al tipo sensor (A) amperímetro y (C) base de corindón. La siguiente sección con la sigla Wn, con n= 1,2,3..., describe el material del electrodo de trabajo, por ejemplo W1 es un electrodo de trabajo compuesto por oro mientras que W2 es Platino. Para el caso del n = S el electrodo de trabajo está integrado por una aleación de oro y platino. El valor Rn con R = 1,2,3..., describe el material en el electrodo de referencia, un ejemplo se puede ver con R1 = Plata / Cloruro de plata o R2 = Plata cubierta por Cloruro de plata. La denominación DW = n describe el diámetro del electrodo de trabajo, con n = 2 el electrodo tiene un diámetro de 2mm.

Definición de ACn	Electrodo trabajo (Wn)	Electrodo referencia (Rn)	Diámetro electrodo de trabajo (DW=n)
A=Amperímetro	n=1, Oro	n=1, Plata/Cloruro de Plata	n=2/ 2mm
C=Base de corindón	n=2, Platino	n=2, Plata cubierta por Cloruro de plata	
n=1 grupo al que pertenece	n=3, Plata		
	n=S, Aleación oro/-platino		

Tabla 2-2.: Características de los sensores utilizados en el presente trabajo.

2.3. Multiplexor MUX8-R2

El MUX8-R2 es un demultiplexor de cuatro a ocho (cuatro entradas ocho salidas) compuesto por cuatro demultiplexores más pequeños, de 1 a 8 conectados en serie paralela, como se muestra en la figura 2-6-A. Este es configurable por software permitiéndole trabajar en diversas maneras. En función de la técnica electroquímica que se vaya a utilizar se puede trabajar en modo secuencial o en alternancia. En el modo alternancia es posible medir varios canales al mismo tiempo, pero solo está disponible para las técnicas Cronoamperometría, Amperometría de Resistencia Cero, Cronopotenciometría y Potenciometría de circuito abierto.



A



B

Figura 2-6.: Esquemático interno y dispositivo MUX8-R2 [2]

2.4. Analizador electroquímico PalmSens4

El PalmSens4 es un potenciostato, galvanostato y analizador de impedancia que se utiliza en análisis electroquímico, este equipo fue usado en esta tesis para realizar el estudio de los cambios en la impedancia del agua provocados por el microplástico. Puede trabajar en un rango de frecuencias que va desde los $10\mu\text{Hz}$ hasta 1MHz (para espectroscopía de impedancia electroquímica EIS), con una amplitud máxima de 250 mV y una resolución en corriente que varía de 100 pA a 10 mA (9 rangos) y cuenta con una entrada analógica de 18 bits. Este equipo se puede configurar para ejecutar diversos tipos de técnicas electroquímicas, entre las que se encuentra la espectroscopia de impedancia electroquímica (EIS) y se puede trabajar con voltaje, frecuencia o voltaje fijo. Si bien es de un solo canal analógico, cuenta con un puerto digital de una entrada y cuatro salidas. La entrada analógica se puede conectar al MUX8-R2, permitiéndole ampliar sus capacidades con una red ocho de sensores, en la figura 2-7 se muestra un esquemático de la conexión entre el PalmSens4 y el MUX8-R2. Si se desea profundizar en el analizador PalmSens4 se puede consultar <http://surl.li/oppas>.

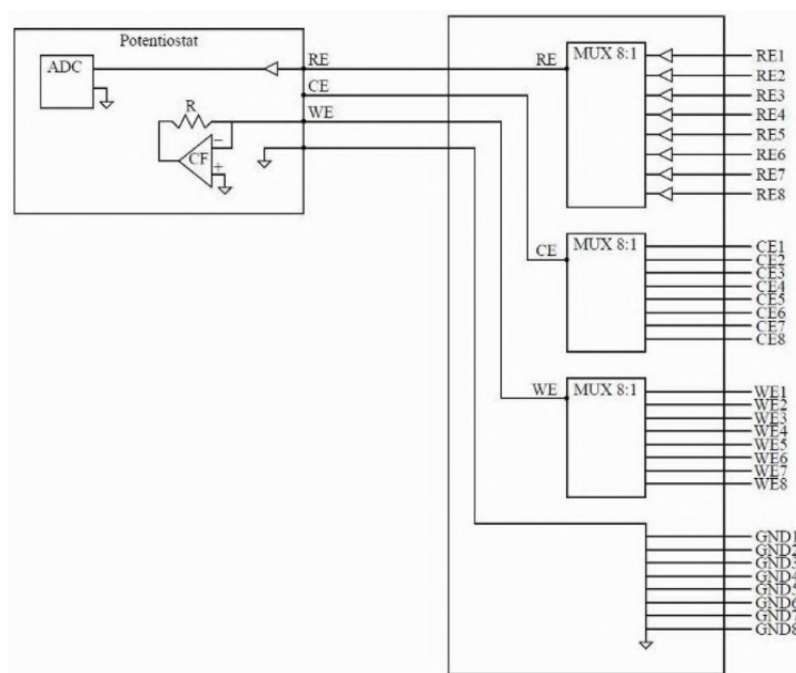


Figura 2-7.: Esquemático conexión PalmSens4 y MUX8-R2 [3]

2.5. Software PSTrace 5

El software PSTrace 5 es utilizado para el cómputo de señales electroquímicas, tiene tres modos de trabajo: científico, en el cual se pueden analizar todas las técnicas electroquímicas ofrecidas en el PalmSens4. Este modo se configura para realizar las mediciones de (EIS),

y dos modos de aplicación específica, análisis de corrosión y analítico. El modo analítico está diseñado para usarse con (bio)sensores. En este software se forman las bases de datos correspondientes a los experimentos llevados a cabo. Aunque solo es posible ejecutar análisis cualitativos, se pueden exportar los datos en diferentes formatos como MATLAB (.m) o Excel (.xls) para su procesamiento masivo. Con la técnica (EIS) este software puede realizar aproximaciones de circuitos eléctricos equivalentes que describan el fenómeno electroquímico que se esté midiendo. En la figura 2-8 se muestran algunos parámetros de configuración y los resultados se expresan en diagramas de Bode o Nyquist.

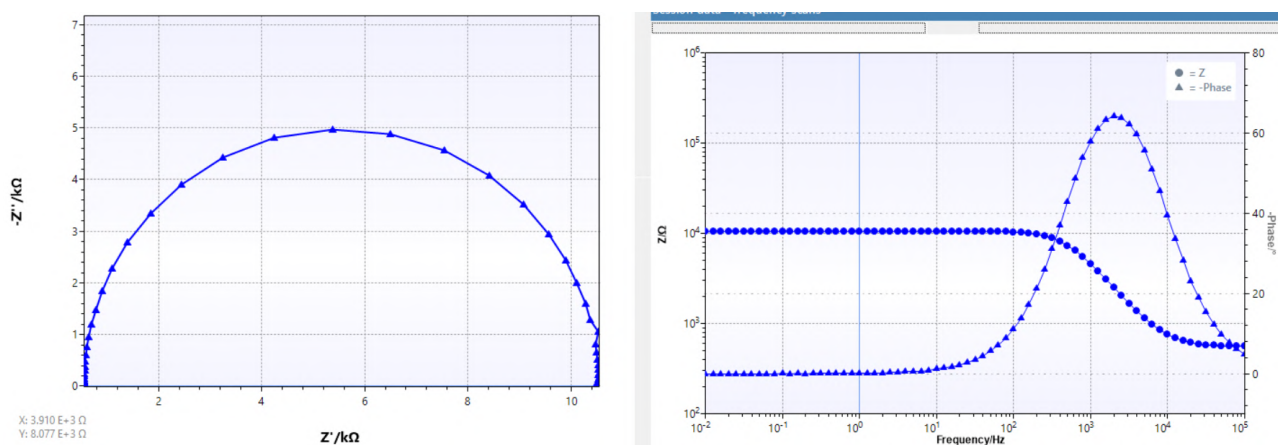


Figura 2-8.: Representación de la técnica EIS Diagrama de Nyquist (Derecha) y diagrama de Bode (izquierda), generada por el software PStTrace 5

Por último, los elementos mencionados se integran para formar la lengua electrónica impedimétrica utilizada en esta tesis figura 2-1.

En la siguiente sección se abordará la metodología desarrollada para la identificación de microplástico usando la espectroscopia de impedancia (EIS) y la arquitectura de lengua descrita en este capítulo.

3. Metodología para la identificación de microplástico

En este capítulo se detalla el desarrollo de la metodología propuesta para la detección y clasificación del microplástico PET y EPS en diversos entornos usando una lengua impedimétrica.

3.1. Descripción general de la metodología

La metodología propuesta en esta tesis pretende desarrollar una serie de experimentos que permitan cuantificar la probabilidad de identificación y clasificación por tamaño de dos tipos de microplástico en diferentes entornos acuáticos con el uso de una lengua electrónica, la cual utiliza la técnica espectroscopia de impedancia(EIS) como método de detección. La (EIS) ya ha sido probada con éxito en la identificación de esferas microplásticas[56]. Sin embargo, es importante validar experimentalmente qué condiciones prestan los mayores desafíos al momento de identificar y clasificar estos materiales. Igualmente, es necesario validar qué algoritmos computacionales entregan una mejor exactitud y en qué ambientes, al momento de identificarlos y clasificarlos por tamaño.

Para esto se elaboró una serie de experimentos que consisten en el desarrollo de seis escenarios, dos de patrón y cuatro con microplástico. Estos últimos simulan las condiciones ambientales en las que se podría encontrar el microplástico, las cuales consisten en: **I.** Agua potable (AP) como patrón 1, **II.** Una solución compuesta por (AP) Materia orgánica inerte (MOI), materia inorgánica (MIN) y materia orgánica viva (MOV) como patrón 2. **III.** Agua potable (AP) con partículas de microplástico (MP) **IV.** Agua potable (AP), partículas de microplástico (MP) y materia inorgánica (MIN). **V.** Agua potable (AP), partículas de microplástico (MP), materia inorgánica (MIN) y materia orgánica inerte (MOI). Por último, **VI.** Agua potable (AP), partículas de microplástico (MP), materia inorgánica (MIN), materia orgánica inerte (MOI) y materia orgánica viva. La medición de estos experimentos se llevó a cabo con cada tipo de microplástico (PET y EPS) por separado y en concentraciones iguales. Los tamaños de microplástico utilizados son los siguientes: 500μ , 1000μ m y 1400μ m, para el PET y el EPS. En la tabla **3-2** se muestra el contenido de los escenarios que agrupa cada uno de los experimentos realizados. El paso siguiente consiste en la medición y adquisición de los datos. Este trabajo lo realizó la lengua electrónica impedimétrica. Los

datos de impedancia no están en función del tiempo sino, como una representación gráfica en el plano complejo. El siguiente paso es realizar un escalamiento de los datos, para esto se agruparon los datos por sensor y debido al alto grado de selectividad cruzada causada por los diferentes materiales que componen los electrodos de trabajo, los datos se encontraron en diferentes escalas por lo que se probaron dos algoritmos de escalamiento para evaluar su resultado. Posteriormente, se utilizó el algoritmo PCA con el objetivo de reducir la alta dimensionalidad presentada por el conjunto de datos. Por último, se dividió el conjunto de datos en tres subconjuntos, entrenamiento validación y test; posteriormente se evaluaron diversos algoritmos de machine learning. En la figura 3-1 se muestra un resumen de la metodología propuesta.

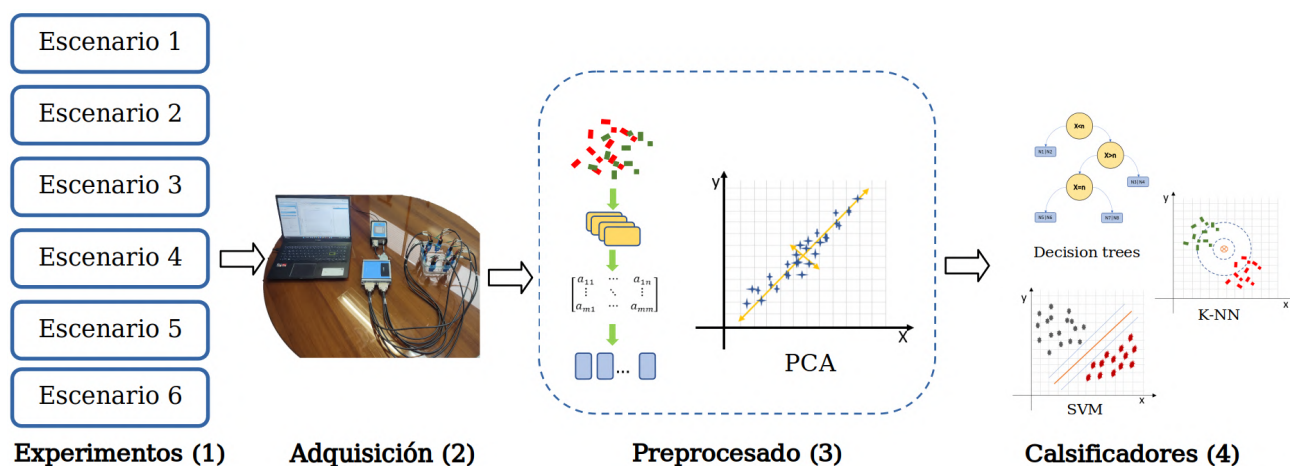


Figura 3-1.: Resumen de la metodología propuesta para la detección del microplástico.
Fuente autor

Los pasos de la metodología presentada en la figura 3-1 se repitieron por cada tipo de microplástico.

3.2. Preparación de las muestras

A continuación se describe el proceso de preparación de las muestras.

Muestras de microplástico

Para el microplástico se utilizaron muestras de Tereftalato de polietileno (PET) y Poliestireno expandido (EPS), debido a su generalizada producción para elementos de consumo masivo, de un solo uso como empaques para alimentos y bebidas, en ejemplo de ello es el PET. De igual modo, el EPS se emplea en la fabricación de recipientes de un solo uso, e igualmente

se utiliza en soportes para bases en construcción de viviendas, cavas y confitería, por lo que su exposición al medio ambiente es relativamente alta.

Para lograr los tamaños requeridos los materiales seleccionados pasaron por un proceso de molienda para posteriormente ser filtradas en un equipo de tamizado, figura 3-2.

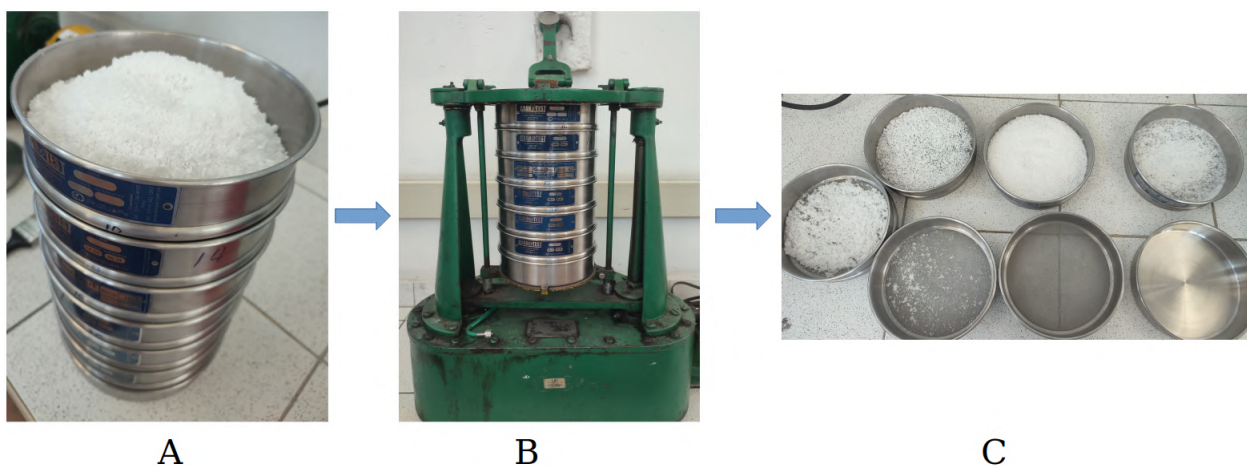


Figura 3-2.: Filtrado de muestras

Las mallas disponibles en el equipo de tamizado se muestran en la figura 3-3, los tamaños seleccionados para los experimentos son: 1.4mm, 1mm y 0.5mm.

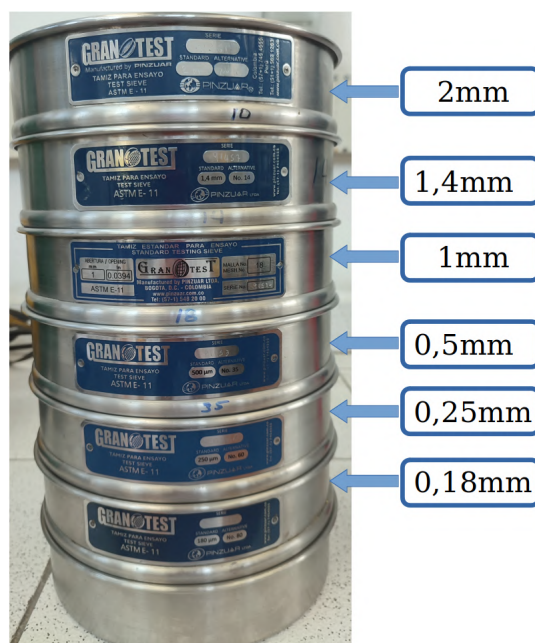


Figura 3-3.: Medidas disponibles en el tamiz

En la figura 3-4 se muestran los resultados obtenidos con los tamaños aproximados para cada tipo de microplástico y en la tabla 3-1 se resumen las muestras de microplástico obtenidas.

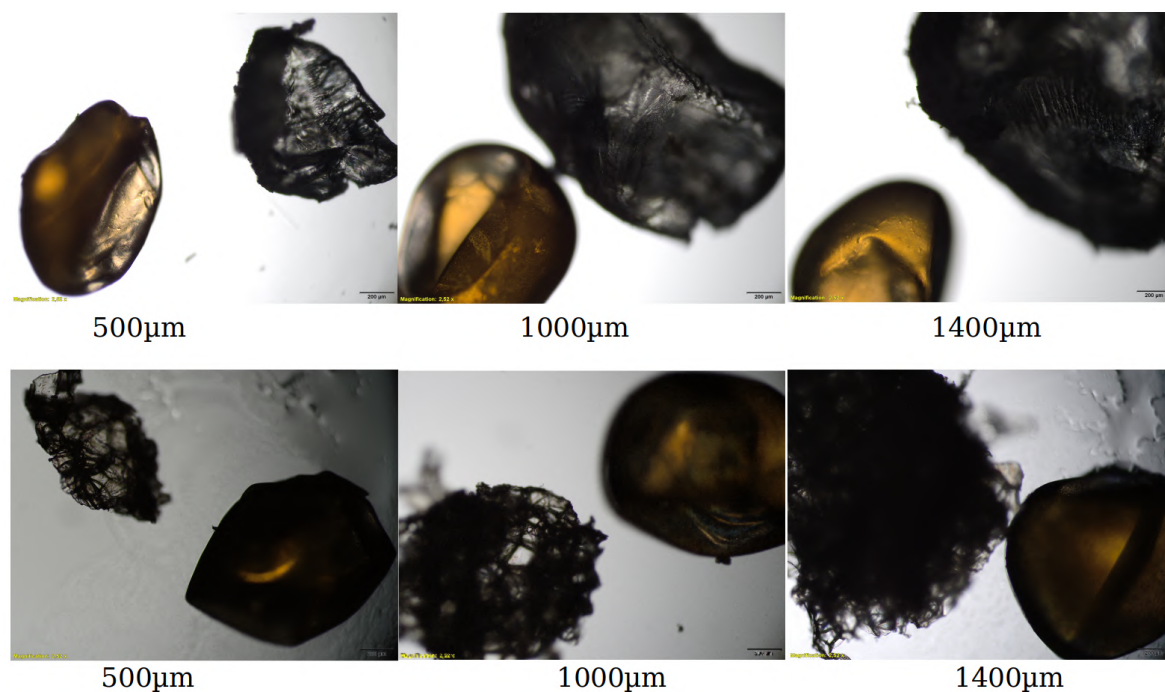


Figura 3-4.: Comparativa entre partículas de microplástico y huevos de salmón. Escala $200\mu\text{m}$. Superior PET, inferior EPS. Fuente autor

Material	Tamaño1	Tamaño2	Tamaño3
PET	$500\mu\text{m}$	$1000\mu\text{m}$	$1400\mu\text{m}$
EPS	$500\mu\text{m}$	$1000\mu\text{m}$	$1400\mu\text{m}$

Tabla 3-1.: Tamaños de microplástico seleccionados para realizar los experimentos.

Muestras de materia inorgánica

La materia inorgánica está compuesta por una mezcla de áridos granulados de origen natural, entre los que se encuentra la arena de río, arena de miga y perlita expandida, como se muestra en la figura 3-5.



Figura 3-5.: Muestra de materia inorgánica

Muestras materia orgánica

Para la materia orgánica se utilizaron dos tipos de muestras diferentes (i) **Orgánica inerte:** compuesta por un agregado de tierra, cascara de semilla y madera molida, figura 3-6-A. (ii) **Orgánica viva:** para esta muestra se utilizaron huevos de salmón, debido a que su tamaño se encuentra entre el rango de microplásticos utilizados en la presente tesis, figura 3-6-B. Una comparación entre el tamaño de los huevos de salmón y las partículas de microplástico se puede ver en la figura 3-4.



Figura 3-6.: Materia orgánica viva (Izquierda) y materia orgánica inerte (Derecha)

En la figura 3-4 se puede apreciar a una escala de $200\ \mu\text{m}$ el tamaño de la materia orgánica viva y las partículas de microplástico. En la figura 3-4-(Superior), el material es PET y la figura 3-4-(Inferior), el microplástico es EPS.

Para la implementación de cada escenario se eligió de forma arbitraria, pero en iguales proporciones, cada uno de los elementos que conforman las distintas soluciones. Se escogió un recipiente de vidrio con un volumen de 1250 mililitros (ml) de sección trapezoidal y simétrica (prisma), como se muestra en la figura 3-7. Los sensores se instalaron en las

paredes del recipiente en grupos de dos, en el orden de la tabla 2-1. La integración de la lengua electrónica se muestra en la figura 3-8.

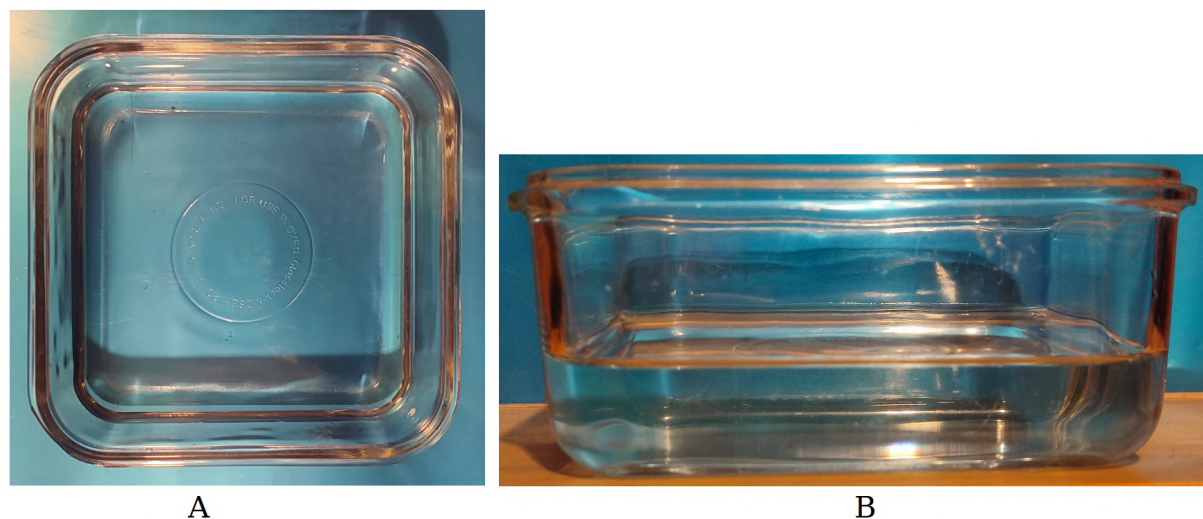


Figura 3-7.: Recipiente utilizado para la preparación de las soluciones.

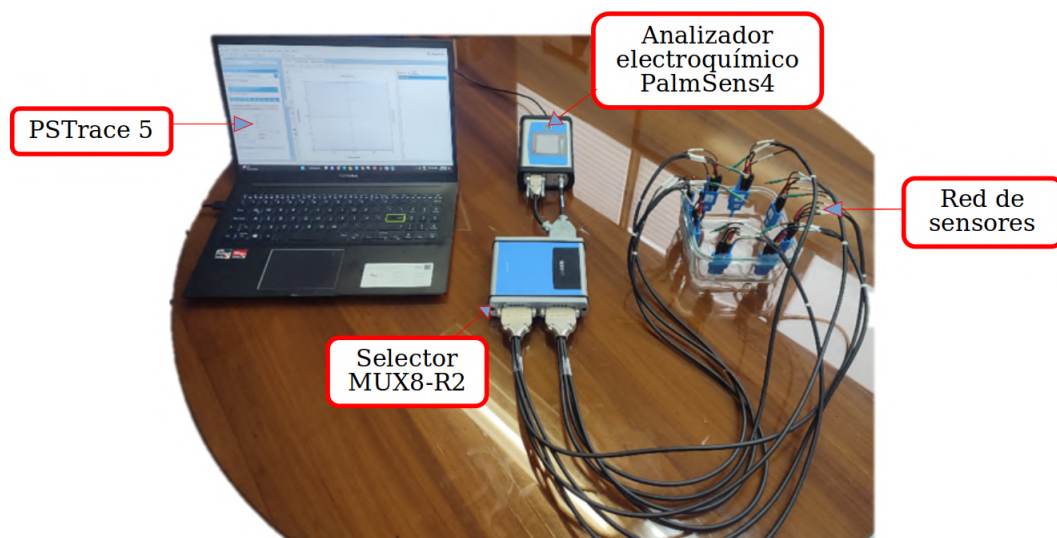


Figura 3-8.: Arquitectura lengua electrónica usada en el presente proyecto, fuente el autor.

Experimentos	
Escenario 1	AP
Escenario 2	AP + MOI + MIN + MOV
Escenario 3	AP + MP 500 μm 1000 μm , 1400 μm
Escenario 4	AP + MIN + MP 500 μm 1000 μm , 1400 μm
Escenario 5	AP + MOI + MIN + MP 500 μm 1000 μm , 1400 μm
Escenario 6	AP + MOI + MIN + MOV + MP 500 μm 1000 μm , 1400 μm

Tabla 3-2.: Escenarios propuestos para la validación de a metodología

A continuación se realiza una descripción detallada de los escenarios presentados en la tabla 3-2.

Escenario 1 agua potable

En este escenario la medición se llevó a cabo utilizando una muestra de agua potable, procedente del suministro de la red local. Se agregaron 420ml del líquido, al recipiente donde se realizan las mediciones, figura 3-7-B. Debido a que el agua puede dirigirse a un sistema de almacenamiento - antes de ser consumida - se tomaron muestras de dos puntos diferentes y se compararon sus medidas de impedancia, los resultados no presentaron cambios aparentes, las gráficas de impedancia correspondiente al agua potable se muestran en la figura 3-9. En la figura 3-10, se muestra la configuración utilizada en la lengua electrónica.

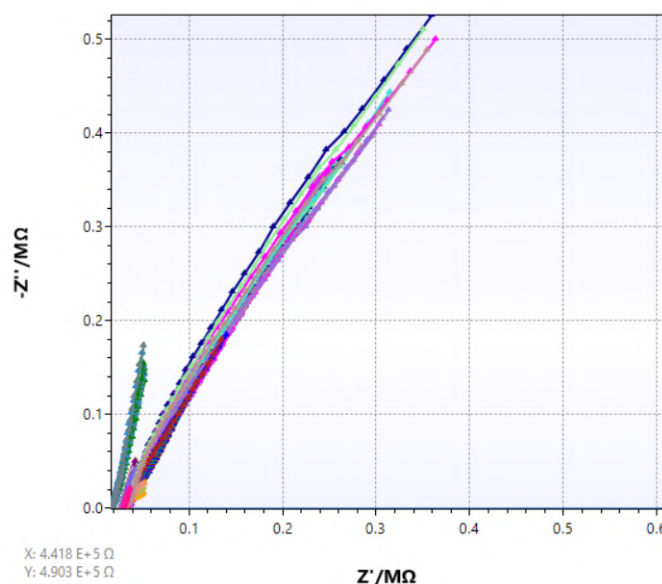


Figura 3-9.: Diagrama de impedancia del agua potable.

Para la configuración de la lengua electrónica se utilizó todo el rango de corrientes que el equipo permitía, y se aplicó una señal de voltaje de 250 mV AC sin componente continuo, el

rango de frecuencias con el que se logró la menor atenuación de ruido fue de 1 ZH a 1 kHz. Se utilizó un tiempo de equilibrio de 3 segundos para estabilizar la señal antes de iniciar cada medición, figura 3-10. Esta configuración fue la utilizada en la medición de todos los escenarios.

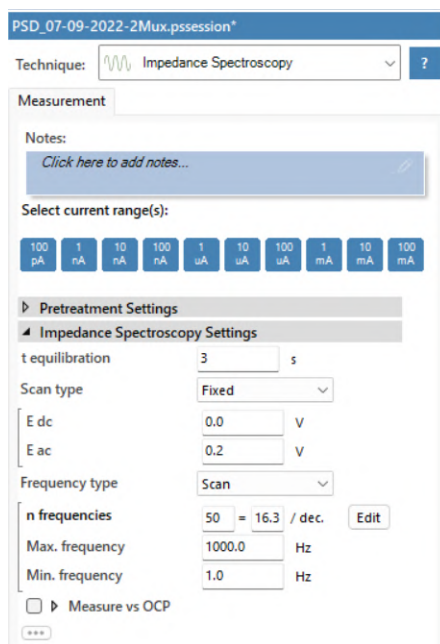


Figura 3-10.: Configuración de la lengua electrónica.

Escenario 2

En este escenario, la solución está compuesta por 420 ml de agua potable, la cual contiene materia orgánica inerte, materia inorgánica y materia orgánica viva en proporciones de 1.4 gramos sin partículas de microplástico. Esta es una solución patrón y se utilizó para compararla con otra con iguales características, pero con partículas de microplástico. Su objetivo es validar si en este entorno es posible detectar el microplástico.

Escenario 3

En este escenario se agregaron 1.4 gramos de microplástico por cada tamaño a la solución de 420 ml de agua. En primer lugar, se realizó la medición con el MP de 500 μm . Posteriormente, se agregaron los tamaños de 1000 y 1400 μm . La medición se repitió para cada muestra agregada. Las mediciones se realizaron de manera independiente para el PET y el EPS, en la figura 3-11 se pueden ver las muestra de MP utilizadas.

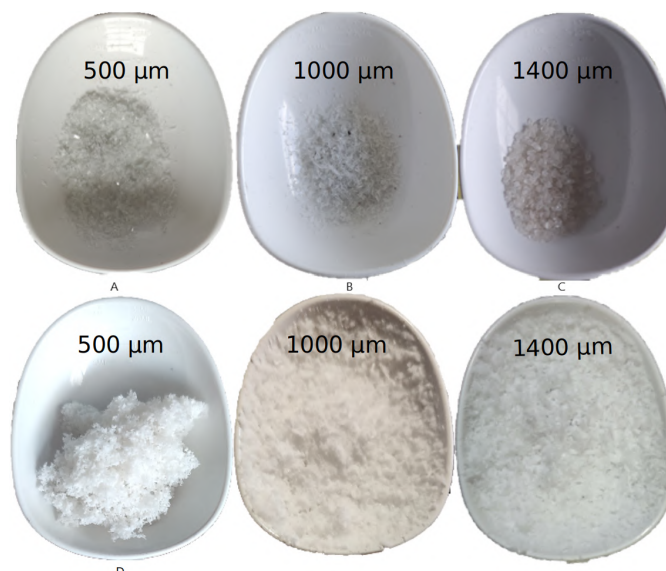


Figura 3-11.: Muestras del MP utilizado. Parte superior PET, inferior EPS

Escenarios 4, 5 y 6

Para estos escenarios se repitieron los pasos del escenario 3. Los materiales se fueron agregando en proporciones iguales. Por cada material se realizó una nueva ronda de mediciones, a su vez este proceso se repitió por cada tamaño de microplástico.

4. Algoritmos utilizados

En esta sección se discute sobre el proceso de adquisición de los datos, así como de los diferentes conjuntos de datos obtenidos. También, se profundiza en los pasos 3 y 4 de la metodología, exponiendo los diferentes algoritmos utilizados en el preprocesado, los clasificadores y la configuración del software que se utilizó, el cual fue Matlab versión R2022b.

4.1. Adquisición de datos

Para la adquisición de los datos se configuró el multiplexor, figura **2-1-B**, en modo secuencial, esto implica que solo un sensor realiza un experimento mientras los demás permanecen inactivos. En la configuración general del equipo figura **3-10** se mostró que el rango de frecuencias está entre 1 Hz y 1000Hz, con un número de 50 datos obtenidos por cada sensor, en esta configuración y teniendo en cuenta el tiempo de equilibrio, una medición completa de la red de sensores está entre 10 y 15 minutos y por cada experimento, tabla **3-2**, se realizaron 20 mediciones. Los valores de cada medición representan componentes reales e imaginarios de los valores de impedancia, por simplicidad representados como (Z_x, Z_y) .

Como cada escenario está planteado en función de una cantidad fija de experimentos que definen los entornos donde se quiere detectar el microplástico, el conjunto de datos completo se puede expresar como un arreglo de conjuntos más pequeños según las mediciones apiladas como se muestra en la figura **4-1**.

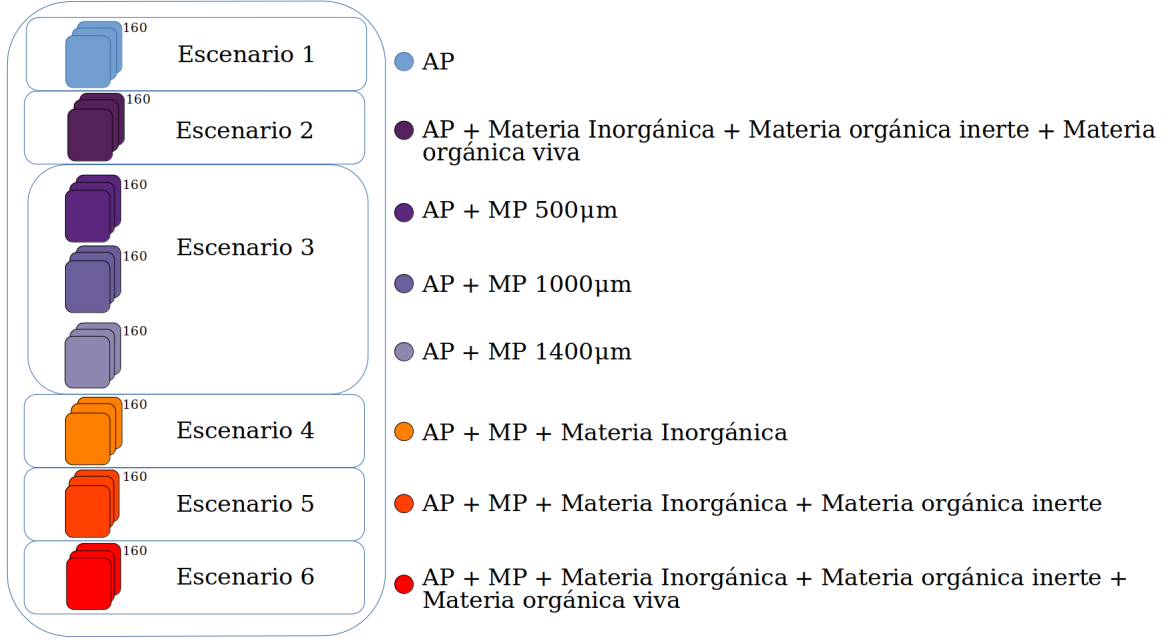


Figura 4-1.: Estructura del set de datos obtenido. Fuente autor

El conjunto de datos - obtenido por experimento - con los parámetros anteriormente nombrados se puede representar como una matriz de tres dimensiones $[M_s \ C \ R_s]$ como se muestra a continuación. Donde M_s representa la cantidad de mediciones por sensor, C las columnas donde se almacenan los valores que entrega la medición y R_s el tamaño de la red de sensores.

$$\begin{pmatrix} [20 \ (Z_x, Z_y) \ Sensor8] & & \\ & \ddots & \\ & & [20 \ (Z_x, Z_y) \ Sensor2] & \\ & & & [20 \ (Z_x, Z_y) \ Sensor1] \end{pmatrix} \quad (4-1)$$

La matriz 4-1 representa la forma en que la lengua electrónica entrega los datos, debido a que cada sensor realizó 20 mediciones, el total de mediciones por experimento es de 160. Una forma simplificada de la matriz 4-1 se puede ver en el arreglo 4-2.

$$\begin{pmatrix} [160 \ (Z_x, Z_y) \ Sensor8] & & \\ & \ddots & \\ & & [40 \ (Z_x, Z_y) \ Sensor2] & \\ & & & [20 \ (Z_x, Z_y) \ Sensor1] \end{pmatrix} \quad (4-2)$$

En esta representación se puede ver cómo se agrupan los datos a medida que aumenta el tamaño de la red de sensores.

Las representaciones 4-1 y 4-2 pueden no tener la forma óptima al momento de ingresar los datos a algún tipo de clasificador, esto debido a que los datos presentan un alto nivel de dimensionalidad, a causa del elevado grado de selectividad cruzada identificada por los sensores. Adicionalmente, los datos se encuentran en escalas diferentes. Si esto no se ha corregido antes de procesar los datos con algún clasificador, los resultados podrían estar sesgados o reflejar sobre ajuste; por tal motivo, en la siguiente sección se va a presentar el preprocesado que se efectuó a los datos para simplificar al máximo su representación.

4.2. Preprocesado

4.2.1. Despliegue

Gracias a que los datos se encontraban apilados en una matriz tridimensional, se realizó un proceso de despliegue que consiste en una reorganización, como se muestra en la figura 4-2. Para lograr esta estructura se cargaron los datos en Matlab por sensor, con cada uno de los experimentos luego se obtuvo su transpuesta.

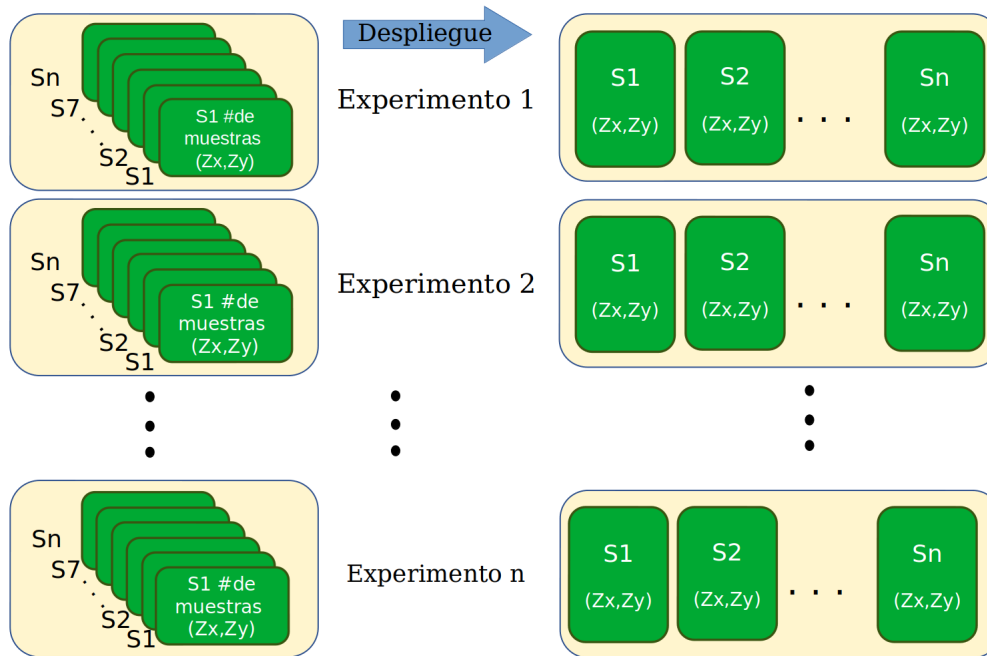


Figura 4-2.: Despliegue de los datos.

Esta forma de desplegar los datos permite que la información procedente de los experimentos se estructure como una matriz bidimensional, donde las filas representan la cantidad de mediciones realizadas por experimento y las columnas los resultados de dicha medición por sensor. La matriz que representa la nueva estructura de datos se pueden describir como:

$$\begin{pmatrix} [40 \ (Z_x, Z_y)^T \ Sensor1] & [40 \ (Z_x, Z_y)^T \ Sensor2] & \cdots & [40 \ (Z_x, Z_y)^T \ Sensor8] \\ [40 \ (Z_x, Z_y)^T \ Sensor1] & [40 \ (Z_x, Z_y)^T \ Sensor2] & \cdots & [40 \ (Z_x, Z_y)^T \ Sensor8] \\ \vdots & \vdots & \ddots & \vdots \\ [40 \ (Z_x, Z_y)^T \ Sensor1] & [40 \ (Z_x, Z_y)^T \ Sensor2] & \cdots & [M_s \ C^T \ R_s] \end{pmatrix} \quad (4-3)$$

La matriz 4-3 representa el conjunto de datos completo después de realizado el despliegue de datos, como los datos obtenidos por la lengua electrónica representan la información de la impedancia en los ejes coordenados, la medición pasa a tener 40 valores, 20 de Z_x y 20 de Z_y . En la gráfica **4-3** se muestra los valores de impedancia de Z_x y Z_y en función de la cantidad de mediciones que realizaba cada sensor, como se explicó anteriormente.

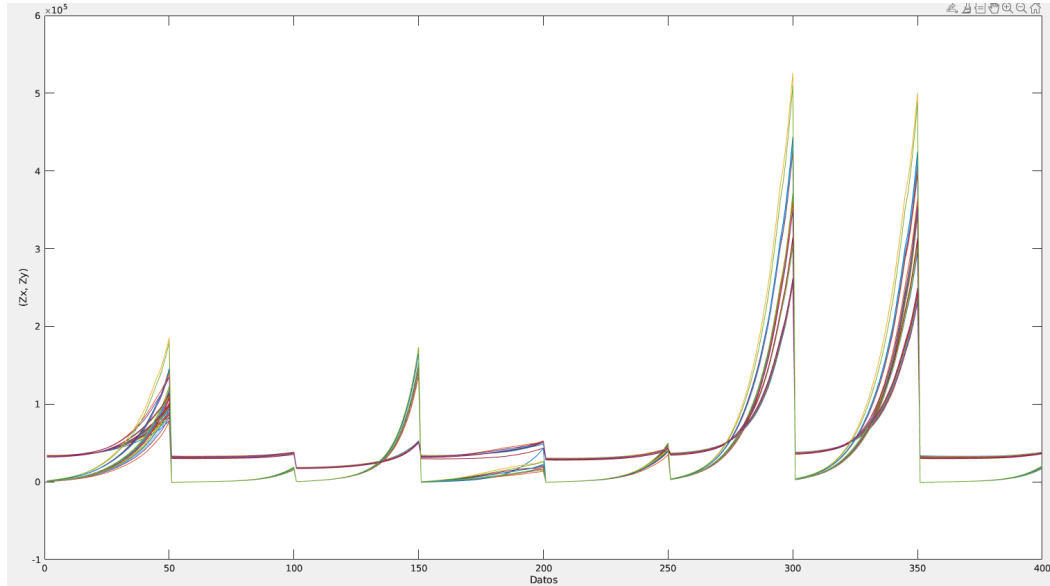


Figura 4-3.: Valores de Z_x y Z_y .

Como se puede ver los datos mostrados en la gráfica anterior está en diferentes escalas, por lo que el paso siguiente al despliegue es normalizar.

4.2.2. Normalización de los datos

Para la normalización se utilizaron dos técnicas group scaling y auto scaling, a cada una de las matrices $[Z_x, Z_y]^T$.

Group scaling

Esta técnica consiste en obtener la desviación estándar (4-4) en cada uno de los puntos obtenidos al momento de realizar la medición y al mismo tiempo que se calcula la media aritmética, ecuación (4-5), en todo el conjunto de datos.

$$\sigma = \frac{\sum_i \sum_k^K (X_{ijk} - \mu_j)}{IK} \quad (4-4)$$

$$\mu_{jk} = \frac{\sum_i X_{ijk}}{I} \quad (4-5)$$

Auto scaling

La técnica de auto scaling toma un enfoque diferente, ya que en esta se aplica la desviación estándar (4-4) y la media aritmética por cada uno de los datos en la medición. Para lograr una mejor claridad entre las dos técnicas descritas anteriormente se muestra la figura 4-4, donde se tomó la gráfica referente a la medición de un solo sensor en un formato que permita visualizar cada dato obtenido por el sensor. Como ejemplo se tomó la medición 1 del primer sensor, correspondiente al experimento de agua potable.

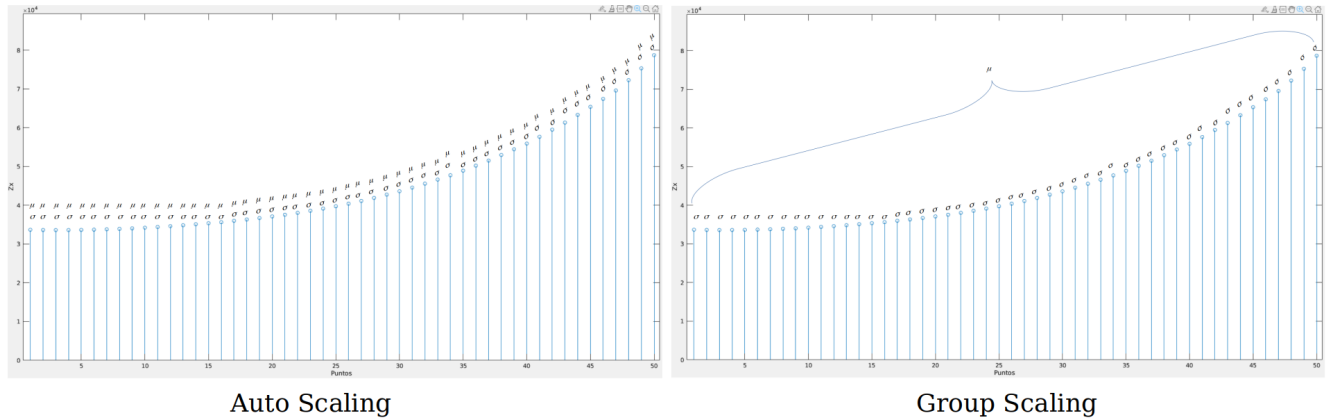


Figura 4-4.: Diferencia entre las técnicas Auto Scaling y Group Scaling.

Análisis de componentes principales PCA

Para obtener una representación donde se exprese la máxima variabilidad de los datos se utilizó el algoritmo PCA, como estos se encuentran distribuidos en un espacio bidimensional 4-3, al aplicar PCA se produce una proyección donde se encuentra la mayor cantidad de datos no correlacionados. Se puede definir como:

$$X \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ T \\ \vdots \end{bmatrix} \begin{bmatrix} E \end{bmatrix} \quad (4-6)$$

Este conjunto de matrices se puede representar como una ecuación, a saber:

$$X = TP^T + E \quad (4-7)$$

Donde P^T corresponde a los coeficientes de los componentes principales, también conocidos como cargas, para la matriz de datos X, T es la matriz de puntuaciones de los componentes principales, representada como un puntaje (score), así como las varianzas de los componentes principales. E es una matriz Residual.

Para evaluar los resultados de PCA primero se deben obtener los valores P^T , con el objetivo de identificar en cuántos componentes (dimensiones) se distribuye la mayor cantidad de información, para esto se mide la varianza por componente, el resultado se muestra en la gráfica 4-5.

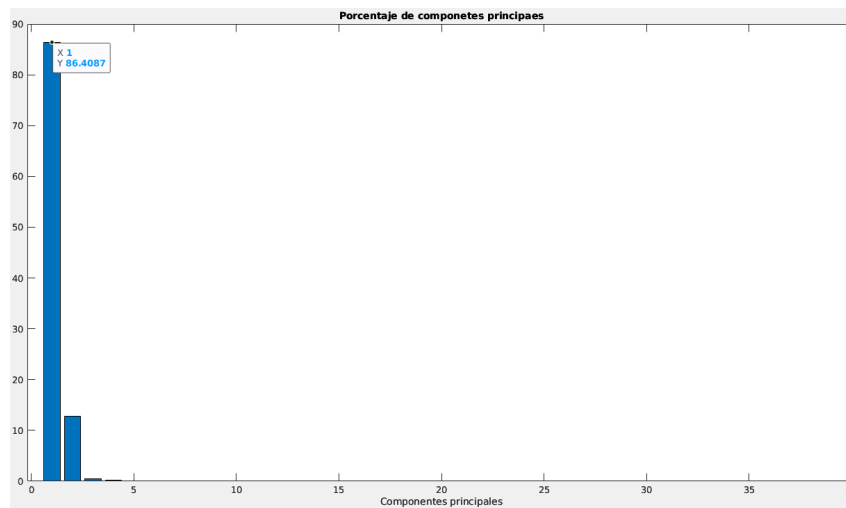


Figura 4-5.: Componentes principales en los que se distribuye la información.

Como se puede ver en la figura 4-5, el 86.40 % de la varianza de los datos se concentra en el primer componente y aunque se grafican los primeros 40 componentes se puede evidenciar

que los primeros 4 componentes contienen más del 90 % de la variabilidad de los datos. El siguiente aspecto a validar es la distribución de los datos, este aspecto es de vital importancia debido a que puede influir en el rendimiento de los algoritmos de clasificación.

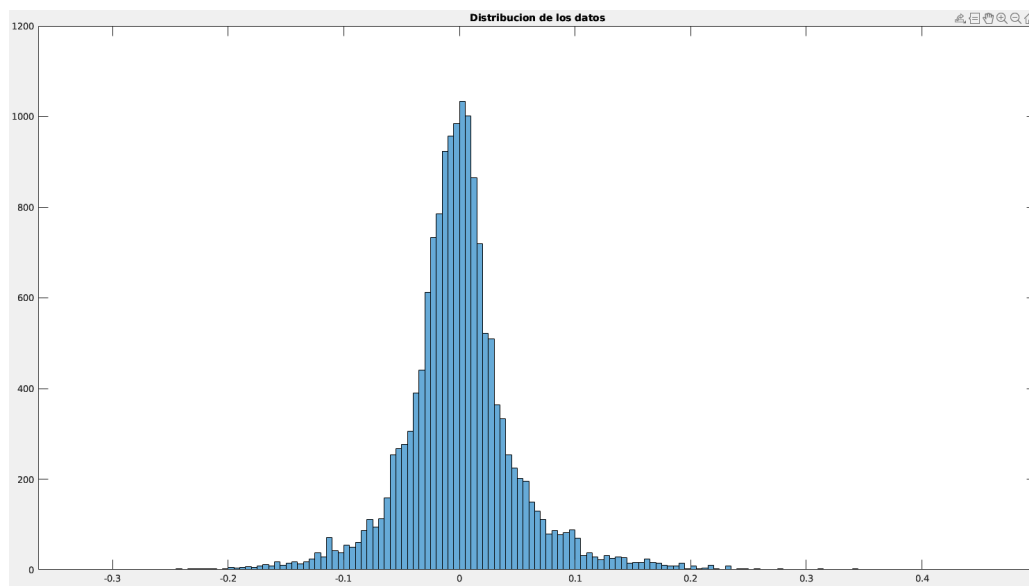


Figura 4-6.: Distribución del conjunto de datos.

Como se puede ver, la media es aproximadamente cero ($\mu \approx 0$) y la desviación estándar se encuentra entre -0.2 y 0.2 ($-0,2 < \sigma < 0,2$).

A continuación se discuten los algoritmos utilizados para la clasificación del microplástico y cómo se implementan en Matlab utilizando la herramienta **Classification Learner**.

4.3. Clasificadores

Para la detección y clasificación por tamaño del microplástico se evaluaron tres algoritmos en cada uno de los experimentos, Máquinas de soporte vectorial (SVM), árboles de decisión y k vecinos más cercanos (KNN), por tal motivo, a continuación se discuten los mecanismos sobre los que trabajan estos algoritmos.

4.3.1. Máquinas de soporte vectorial SVM

Las máquinas de soporte vectorial son un algoritmo supervisado enfocado a la clasificación binaria, aunque también se puede usar en la regresión lineal. Este algoritmo se puede integrar con redes neuronales para evaluar su rendimiento o como etapa de salida. El objetivo de las SVM es encontrar el mejor hiperplano que funcione como frontera para clasificar datos, los vectores de soporte corresponden a los segmentos de datos más cercanos entre sí, correspondientes a las clases que se deseen clasificar, este proceso se ilustra en la figura 4-7.

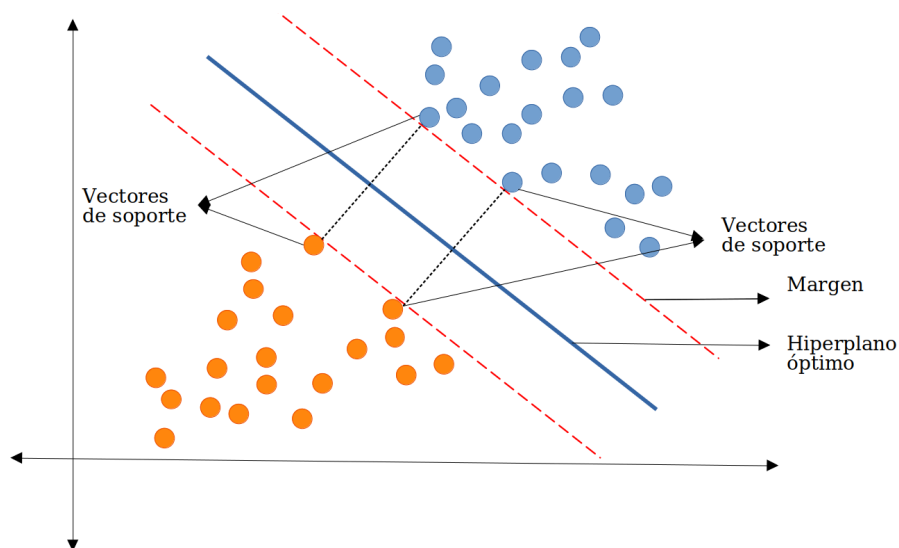


Figura 4-7.: Máquina de soporte vectorial. Fuente autor

Las márgenes son líneas paralelas al hiperplano que pasan por los vectores de soporte y cuya distancia al hiperplano debe ser equidistante para evitar sesgos en el algoritmo.

De manera simplificada, si se quiere clasificar un conjunto de datos en dos clases P_n y Q_n en un espacio bidimensional se podría definir un hiperplano como $W_1X_1 + W_2X_2 + b$, para el caso de la figura 4-7, donde W_1 , W_2 y b son los coeficientes que describen el hiperplano óptimo. Sin embargo, existen muchos conjuntos de datos que no pueden ser separados por un hiperplano, para solucionar este problema las SVM utilizan una función no lineal comúnmente denominada como Kernel, que eleva el nivel de dimensionalidad de los datos obteniendo un nuevo espacio vectorial que permita agregar un hiperplano para clasificar los conjuntos de datos. Los Kernel usados por Matlab se muestran a continuación.

Tipo de SVM	Kernel de Mercer	Descripción
Función de base radial (RBF) o gaussiana	$K(x_1, x_2) = \exp\left(-\frac{\ x_1 - x_2\ ^2}{2\sigma^2}\right)$	Aprendizaje de una clase. σ representa la anchura del kernel.
Lineal	$K(x_1, x_2) = x_1^T x_2$	Aprendizaje de dos clases.
Polinómica	$K(x_1, x_2) = (x_1^T x_2 + 1)^\rho$	ρ representa el orden del polinomio.
Sigmoide	$K(x_1, x_2) = \tanh(\beta_0 x_1^T x_2 + \beta_1)$	Representa un kernel de Mercer solo para determinados valores β_0 y β_1 .

Figura 4-8.: Kernel usados por Matlab, tomado de[4].

Luego de obtener el hiperplano en el nuevo espacio vectorial, las SVM retornan al espacio original para terminar la clasificación.

4.3.2. K-vecinos más cercanos KNN

El KNN es un algoritmo del dominio no supervisado, no perimétrico y basado en instancias. Esto implica que no requiere una distribución específica de los datos, adicionalmente KNN no aprende explícitamente del modelo, en cambio, memoriza las instancias de formación que posteriormente se utilizan para la fase de predicción. Cuando se ingresa un dato nuevo al modelo, el algoritmo calcula la distancia entre el dato de entrada y los datos de entrenamiento, posteriormente el dato es clasificado en el conjunto de datos numéricamente más cercano. Algunas de las distancias paramétrica que utiliza KNN se muestran en la siguiente tabla.

Distancia paramétrica	Nombre
$d_{(x_i, x_j)} = \sqrt{\sum_{r=1}^p (x_{ir} - x_{jr})^2}$	Distancia euclidiana
$D_{Chebyshev}(p, q) := \max_i (p_i - q_i)$	Distancia Chebyshev
$D(X, Y) = (\sum_{i=1}^n x_i - y_i ^p)^{\frac{1}{p}}$	Distancia de Minkowski
$sim_j(A, B) = \frac{ A \cap B }{ A \cup B }$	Índice de Jaccard

Tabla 4-1.: Relación de los diferentes modelos que utiliza KNN para calcular la distancia.

Para lograr cluster de datos que permitan agrupar las distintas clases, el algoritmo debe iterar varias veces acarreando un costo computacional relativamente elevado.

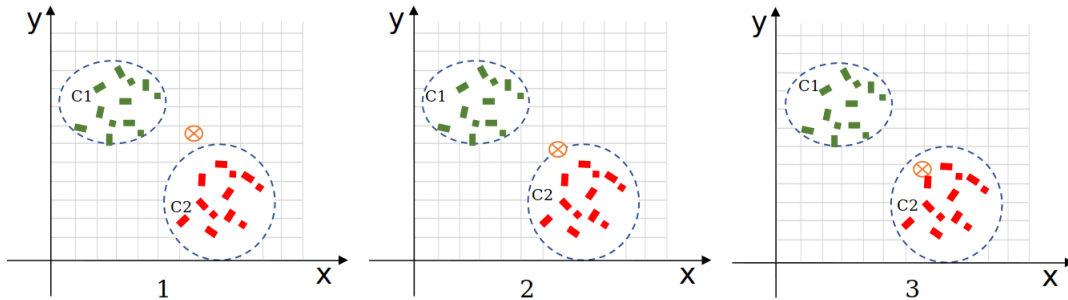


Figura 4-9.: Iteraciones de KNN. Fuente autor

Los principales hiperparámetros que se deben validar al momento de implementar KNN son el número de vecinos, generalmente entre más pequeño este dato, menos riesgo de sesgo se presenta y la métrica de distancia la cual se define en términos de una desigualdad, como lo son ($=, <, >$), que sirva como criterio para aproximar un dato a un conjunto o vecindad.

4.3.3. Árboles de clasificación y regresión CART

El CART es un algoritmo sencillo de implementar, permite manipular datos relativamente complejos, con resultados simples de interpretar. La manera en que funciona el algoritmo CART consiste en realizar una serie de particiones sobre el conjunto de datos de un límite dado que permitan clasificar los datos en conjuntos homogéneos, con el fin de medir la homogeneidad de los datos se utiliza el índice Gini, el cual está definido por la ecuación de Brown.

$$G_i = \left| 1 - \sum_{k=1}^{n-1} (X_{k+1} - X_k)(Y_{k+1} + Y_k) \right| \quad (4-8)$$

Entre más cercano a cero se encuentre el coeficiente G_i más homogéneos serán los conjuntos de datos que se estén clasificando. La manera en que el algoritmo CART evalúa los resultados de las particiones es ejecutando una función de costo que asigna un puntaje al dato de la partición inicial (Nodo padre), o anterior utilizando una función de costo que calcula el valor promedio ponderado de los datos mal clasificados como se describe en la ecuación 4-9.

$$V_p = \frac{G_i(D_i + D_j)}{N} \quad (4-9)$$

Donde D_i corresponde al conjunto de datos mal clasificados, D_j son los datos clasificados correctamente y N son los datos de la partición anterior. Esta función de costo se calcula por cada nodo y hace que el árbol se despliegue a medida que se reduce su valor mejorando la clasificación de los datos, mientras que los nodos con una alta función de costo se van eliminando. En la figura 4-10 se muestra cómo conforme que V_p y G_i disminuyen, los datos mejoran su clasificación.

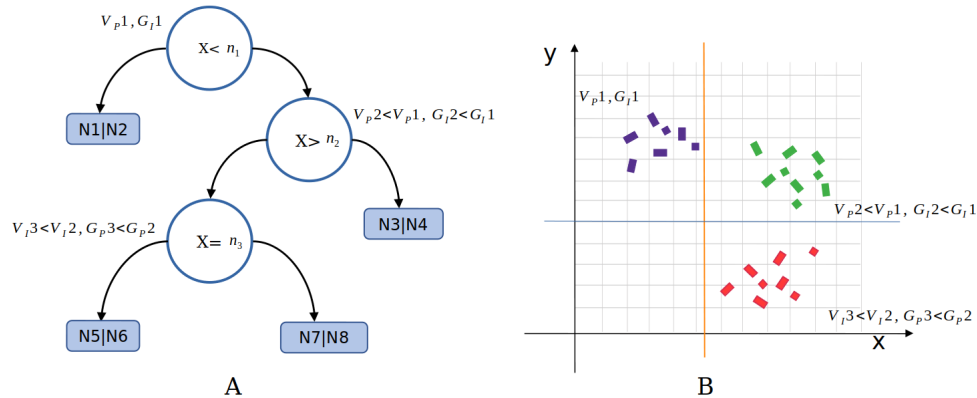


Figura 4-10.: Algoritmo CART. Fuente autor

A continuación se muestra una guía básica sobre el uso de la herramienta que se utilizó para realizar el proceso de clasificación; el Classification Learner (CL) de Matlab.

4.3.4. Classification Learner CL

El CL es una aplicación que implementa algoritmos de inteligencia artificial, se utiliza principalmente para el entrenamiento y evaluación de distintos algoritmos como las redes neuronales, análisis discriminantes, clasificadores Niave Bayes, máquinas de soporte vectorial, árboles de decisión o K-vecinos más cercanos. El CL permite ejecutar distintos modelos de forma concurrente, así como comparar su rendimiento.

A continuación se muestra una guía básica correspondiente al uso del CL de Matlab usando como ejemplos el algoritmo KNN presentando en la sección anterior.

Análisis exploratorio de datos con el CL

Este es el primer paso antes de implementar cualquier modelo. El CL permite realizar un estudio de los datos sin ningún tipo de preprocesado, aunque para el presente trabajo sí se realizó un preprocesado el cual se mostró en la sección 4.2.

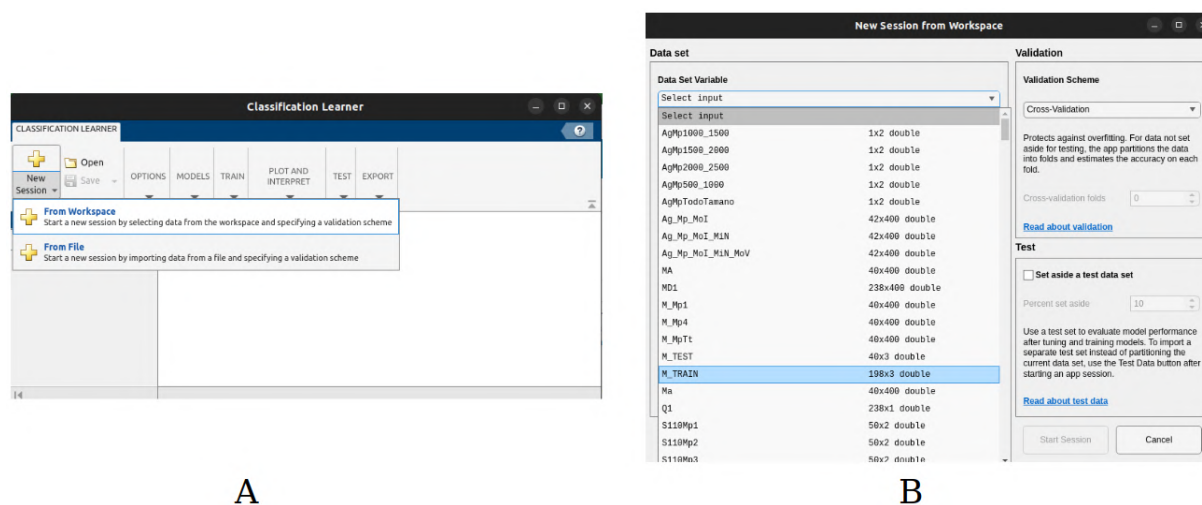


Figura 4-11.: Cargue de datos en la CL

Selección de algoritmos y configuración

Se pueden cargar datos desde el espacio de trabajo en Matlab o desde un archivo en otra ubicación. Si se desea realizar un análisis exploratorio inicial o directamente evaluar algún algoritmo es importante que los datos estén etiquetados, ya que la herramienta detecta las etiquetas de forma automática. En la parte derecha de la figura 4-11-B, se pueden configurar varios esquemas de validación como la (i). Cross-Validation, para los datos que no están reservados para pruebas, (ii). Holdout Validation utilizada en grandes conjuntos de datos

donde se puede escoger un porcentaje de los mismos para realizar el proceso de validación, y por último, la (iii). Resubstitution Validation, en ese modo la herramienta utiliza los mismos datos de entrenamiento para realizar la validación, por lo que no se protege contra el sobre ajuste. En la sección Test se puede extraer un conjunto de datos de prueba del set de datos original.

Cuando se han cargado los datos, la toolbox permite realizar una validación utilizando el algoritmo de CART y despliega los datos como se muestra en la figura 4-12-A.

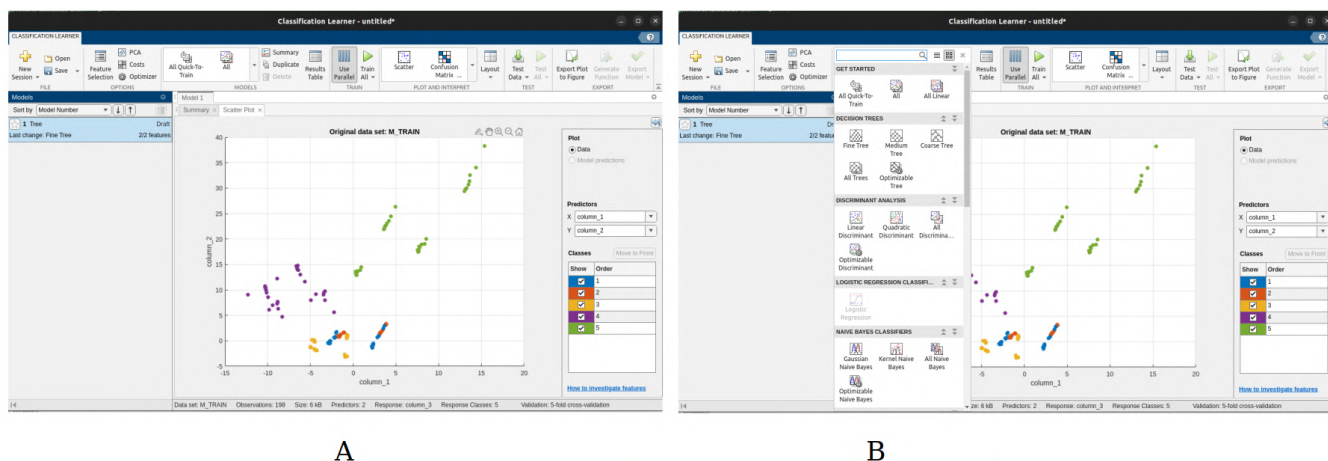


Figura 4-12.: Representación de lo datos (A). Selección del algoritmo (B).

En la sección MODELS se pueden escoger los algoritmos que se desea evaluar, figura 4-12-B. Si selecciona cualquier algoritmo con la opción de optimizable, se pueden inicializar los hiperparámetros para el CL, este los ajustará según los resultados que vaya obteniendo. Al momento de evaluar varios algoritmos al tiempo, el CL tiene la opción de concurrencia en la sección train-use parallel.

Test y resultados

Otra forma de validar los resultados del entrenamiento es cargando una matriz de test con datos previamente etiquetados, figura 4-13-A. Los resultados del test se pueden obtener de la sección plot and interpret, figura 4-13-B.

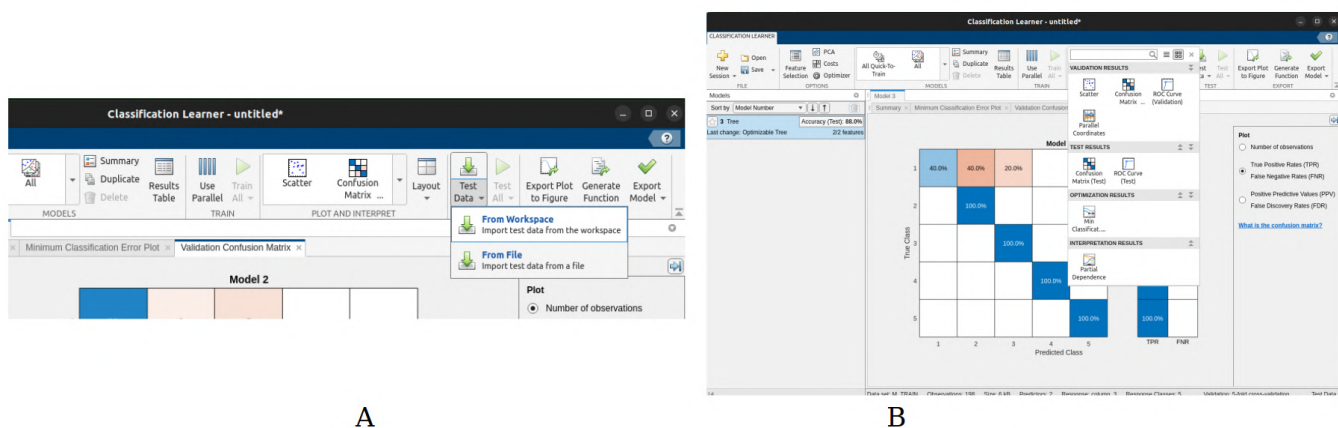


Figura 4-13.: Representación de los datos (A). Selección del algoritmo (B).

Las gráficas de resultados, así como los algoritmos utilizados que se deseen exportar, se muestran en la sección EXPORT. los resultados de los clasificadores de pueden obtener en la sección PLOTS AND RESULTS.

En el siguiente capítulo se van a discutir los resultados que se obtuvieron con los algoritmos para la clasificación de los diferentes microplásticos.

5. Resultados

En este capítulo se discuten los diferentes resultados obtenidos con la metodología expuesta. De igual forma, se presentan los resultados obtenidos en las fases de escalado, reducción de dimensionalidad, entrenamiento y test. En este último capítulo se comparan las respuestas alcanzadas por los diferentes algoritmos en cada uno de los experimentos. Primero, se muestran los resultados de la detección del microplástico, PET y EPS, en dos entornos: *(i)*. Agua potable y *(ii)*. Agua con una mezcla de materia orgánica inerte, materia orgánica viva y materia inorgánica, denominada en este trabajo como agua contaminada. Posteriormente, se muestran los resultados de la clasificación por tipo de microplástico en cada uno de los entornos propuestos en el capítulo 3.

5.1. Escalado de los datos

Una vez obtenida la información de los sensores se debe realizar el proceso de despliegue, esto gracias a la manera en que los datos son entregados, en forma de pila. En efecto, esta estructura de datos no es la adecuada para ingresarlo a los clasificadores. La técnica que se utilizó para este proceso se describió en la sección 4.2.1.

Ahora bien, los resultados del despliegue de datos se muestran en la siguiente figura.

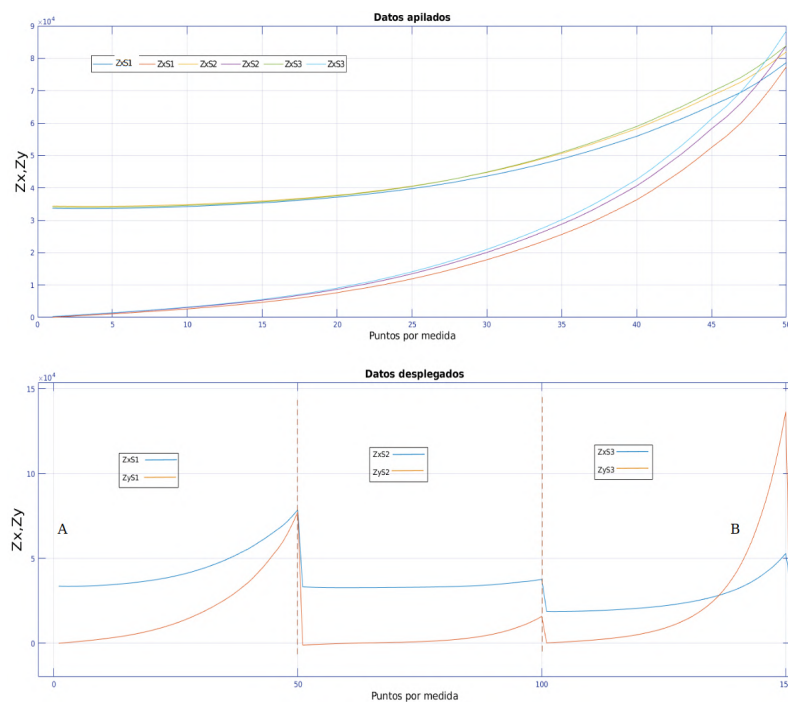


Figura 5-1.: Datos apilados (Superior), datos desplegados (Inferior)

Los datos de la figura 5-1 no se encuentran escalados. A continuación se muestran los resultados correspondientes al de escalado en los datos.

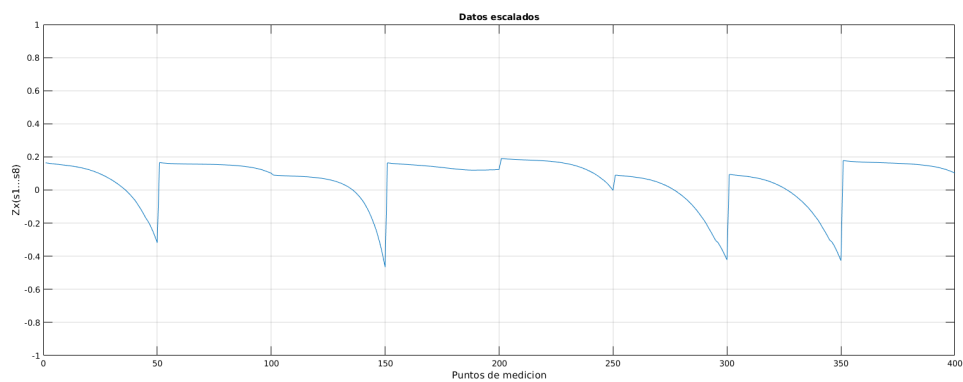


Figura 5-2.: Datos escalados

Como se muestra en la figura 5-2, los datos quedaron escalados entre 1 y -1.

5.2. Reducción de dimensionalidad utilizando PCA

A continuación se muestran los resultados obtenidos utilizando el análisis de componentes principales para reducir la dimensionalidad de los diferentes conjuntos de datos. En general, se utilizó un número de componentes $k = 2$; sin embargo, este parámetro tuvo variaciones en función de la mezcla que caracterizaba la solución.

5.2.1. PCA en la detección de microplástico

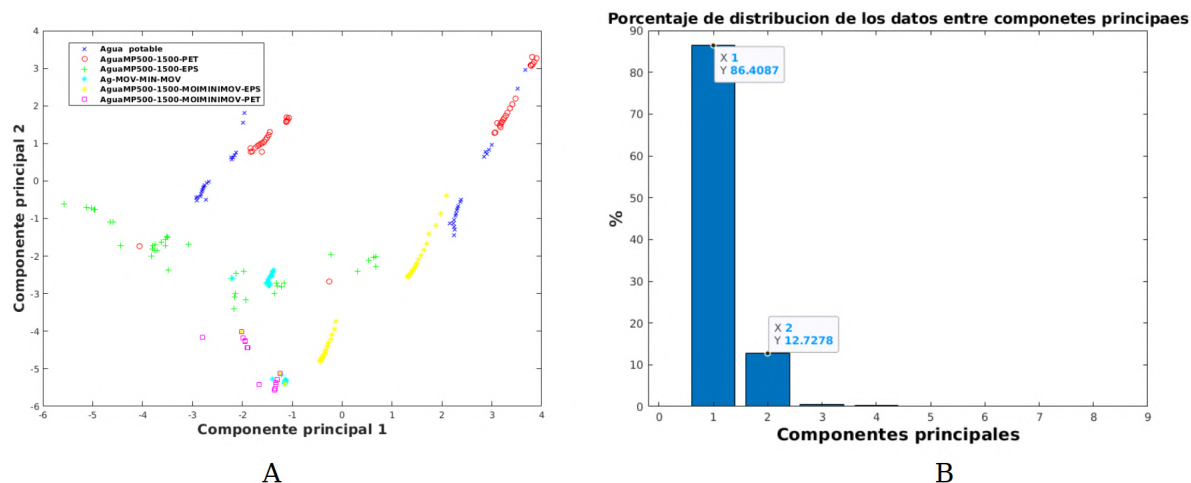


Figura 5-3.: Distribución de los datos en un espacio bidimensional correspondiente a los dos primeros componentes (A). Porcentaje de la distribución de los datos entre los componentes (B).

Al aplicar PCA en la detección del microplástico, la máxima variabilidad de los datos se distribuye en el primer componente con un 86.4 %; en el segundo componente este porcentaje se reduce a 12.7 % por lo que estos dos componentes representan el 99.1 % de los datos.

5.2.2. PCA en la clasificación por tamaño del PET y EPS en agua potable

Resultados de PCA para el EPS

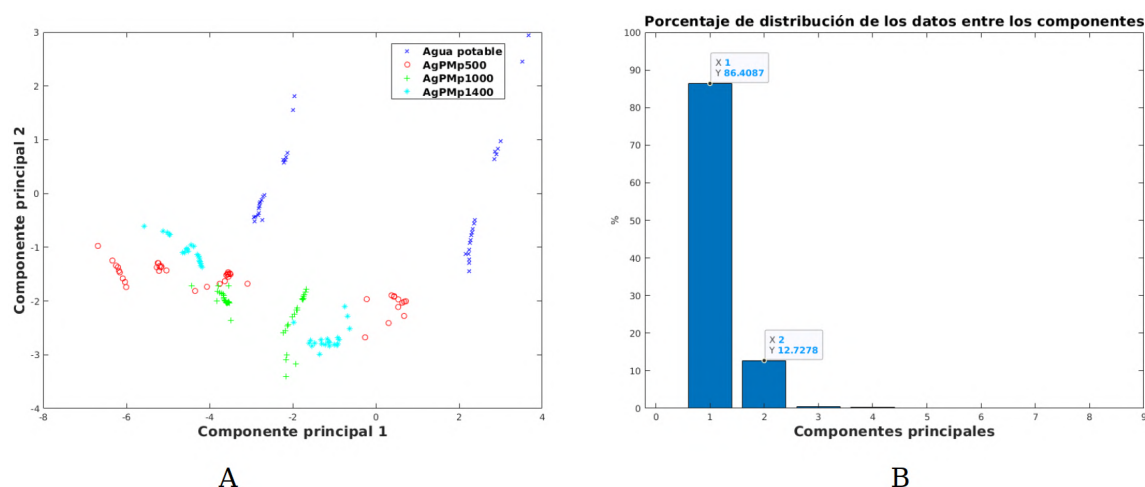


Figura 5-4.: Distribución de los datos en un espacio bidimensional correspondiente a los dos primeros componentes en la clasificación del microplástico EPS por tamaño en un entorno de agua potable (A). Porcentaje de la distribución de los datos entre los componentes (B).

Resultados PCA para el PET

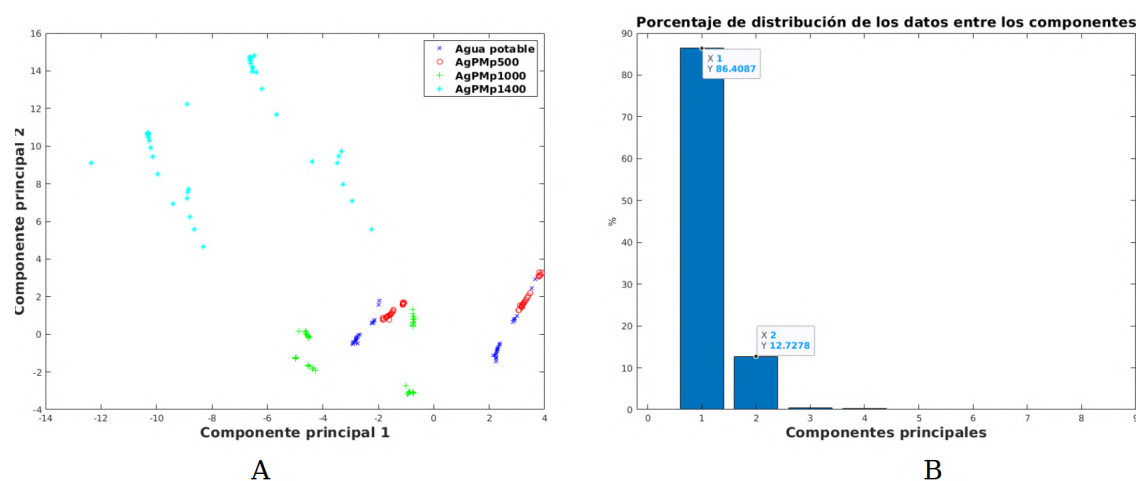


Figura 5-5.: Distribución de los datos en un espacio bidimensional correspondiente a los dos primeros componentes en la clasificación del microplástico PET por tamaño en un entorno de agua potable (A). Porcentaje de la distribución de los datos entre los componentes (B).

5.2.3. PCA en la clasificación por tamaño del PET y EPS diferentes entornos

Resultados de PCA para la reducción de dimensionalidad de los datos correspondientes a una solución compuesta por agua mezclada con materia orgánica inerte, microplástico EPS en dimensiones de $500\mu\text{m}$, $1000\mu\text{m}$ y $1400\mu\text{m}$ y microplástico PET en dimensiones de $1000\mu\text{m}$. La composición de la materia orgánica inerte se describe en el capítulo 3.2.

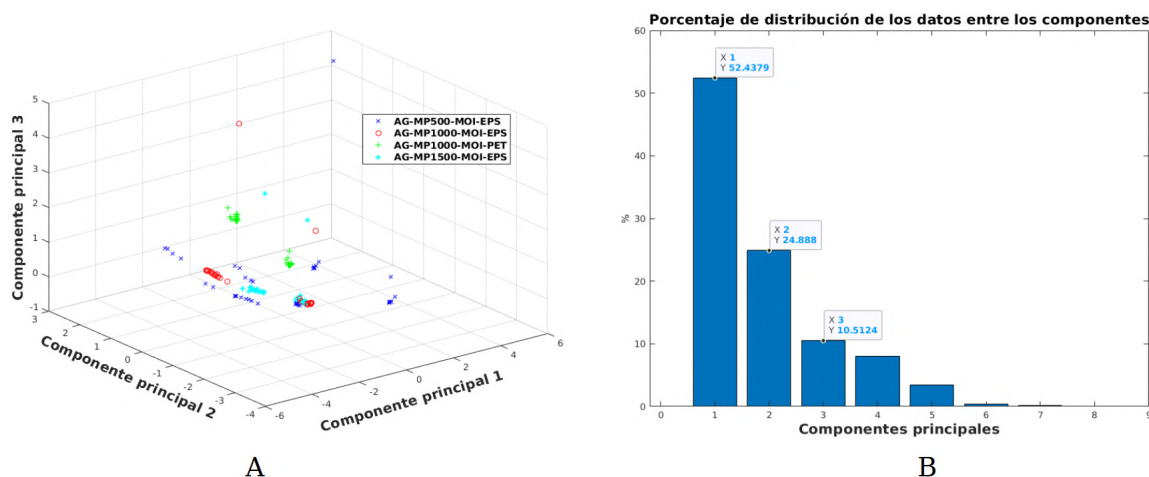


Figura 5-6.: Distribución de los datos en un espacio tridimensional correspondiente a los tres primeros componentes para la clasificación del microplástico PET y el EPS por tamaño, en un entorno de agua contaminada con materia orgánica inerte(A). Porcentaje de la distribución de los datos entre los componentes (B).

En este caso, se seleccionó un número de componentes $k = 3$ debido a que se obtuvo una mejor representación de la información, con un 87.7 % de los datos.

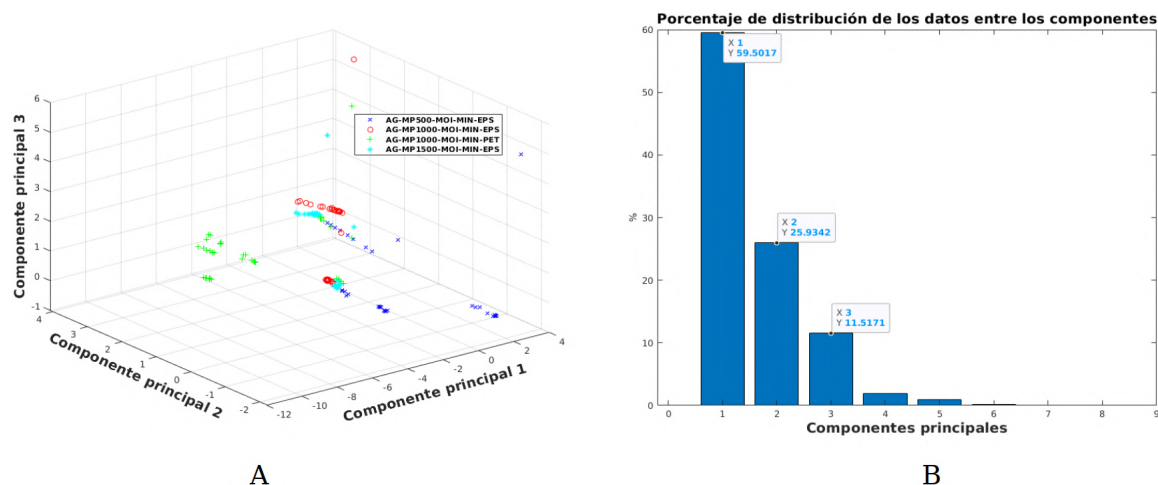


Figura 5-7.: Distribución de los datos en un espacio tridimensional correspondiente a los tres primeros componentes para la clasificación del microplástico PET y el EPS por tamaño, esto en un entorno de agua contaminada con materia orgánica inerte y materia inorgánica(A). Porcentaje de la distribución de los datos entre los componentes (B).

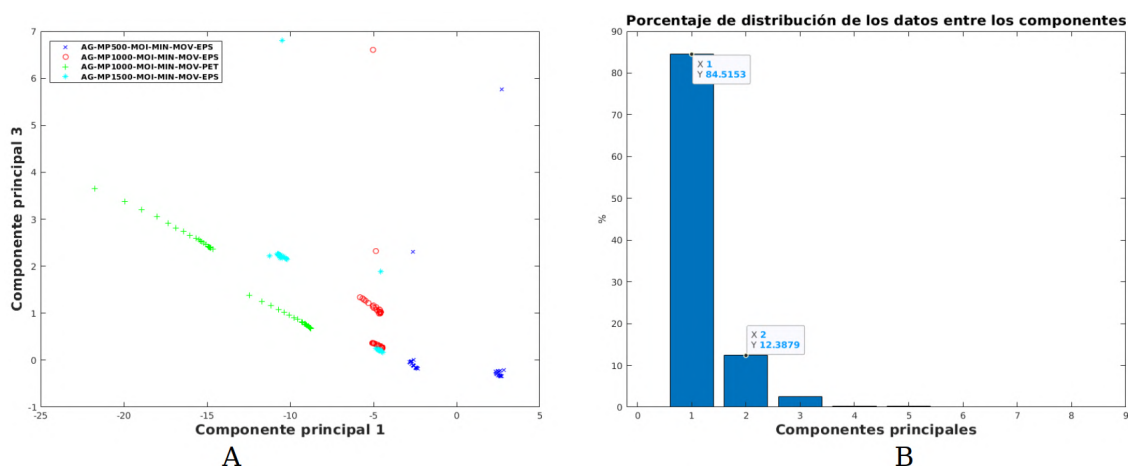


Figura 5-8.: Distribución de los datos en un espacio tridimensional correspondiente a los dos primeros componentes para la clasificación del microplástico PET y el EPS por tamaño, esto en un entorno de agua contaminada con materia orgánica inerte, materia inorgánica y materia orgánica viva(A). Porcentaje de la distribución de los datos entre los componentes (B).

En la siguiente tabla se muestra un resumen con los resultados de PCA en los diferentes experimentos.

Entorno	# de componentes	% de datos
Detección de MP en agua potable y agua contaminada	2	99.1 %
Clasificación por tamaño de MP EPS en agua potable	2	99.1 %
Clasificación por tamaño de MP PET en agua potable	2	99.1 %
Clasificación por tamaño de MP EPS y PET en agua potable con materia orgánica inerte	3	87.71 %
Clasificación por tamaño de MP EPS y PET en agua potable, materia orgánica inerte y materia inorgánica	3	96.9 %
Clasificación por tamaño de MP EPS y PET en agua potable, materia orgánica inerte, materia inorgánica y materia orgánica viva	2	97.88 %

Tabla 5-1.: Componentes utilizados para la óptima distribución de los datos utilizando PCA

En la siguiente sección, se muestran resultados alcanzados por los distintos algoritmos en la detección y clasificación de microplástico EPS y PET.

5.3. Detección

La matriz de confusión de la figura 5-9 está compuesta por seis clases, en las primeras tres se evalúa el algoritmo CART para identificar los microplásticos PET y EPS, en diferentes tamaños, en agua potable. De la clase cuatro a la seis se evaluó la precisión en un entorno de agua contaminada.

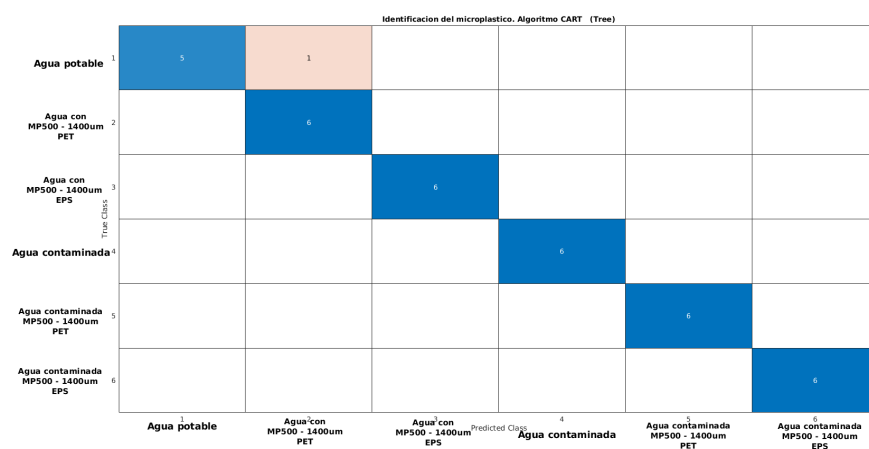


Figura 5-9.: Identificación del microplástico utilizando el algoritmo CART.

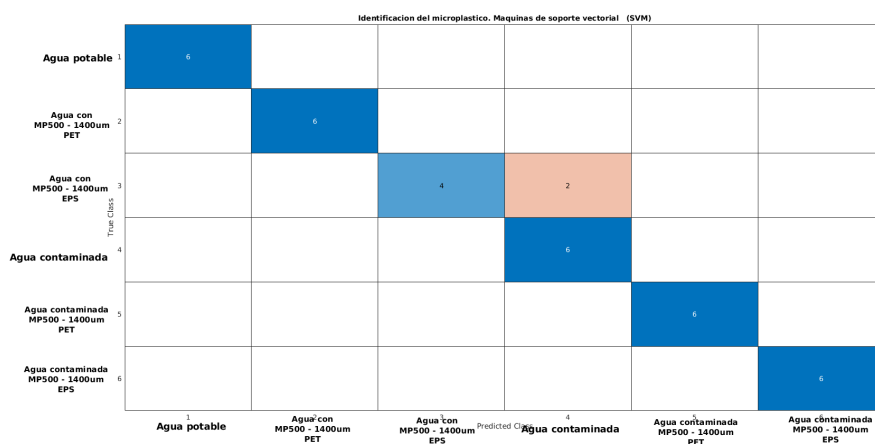


Figura 5-10.: Identificación del microplástico utilizando el algoritmo SVM.

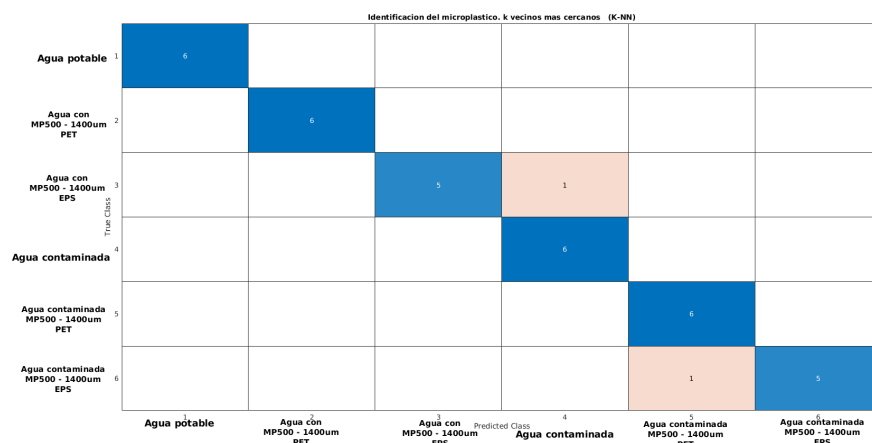


Figura 5-11.: Identificación del microplástico utilizando el algoritmo SVM.

A continuación se presenta una tabla resumen con las métricas de evaluación para la detección de los dos materiales.

Resultados detección PET y EPS		
CART	Precision	85.7 %
	Recall	83.3 %
	F1-score	84.48 %
	Accuracy	97.2 %
SVM	Precisión	75 %
	Recall	66.7 %
	F1-score	70.60 %
	Accuracy	94.4 %
K-NN	Precisión	85.7 % / 85.7 %
	Recall	83.3 % / 83.3 %
	F1-score	85.34/85.34 %
	Accuracy	94.4 %

Tabla 5-2.: Resultados generales para la detección del microplástico. El parámetro precisión se mide por clase y Accuracy mide el rendimiento general del algoritmo

Como el objetivo inicial es la detección del microplástico en cualquiera de los dos entornos, inicialmente se le da más importancia a la métrica de recall debido a que esta medida se asocia con la sensibilidad que tienen los algoritmos para detectar cambios. La métrica de precisión evalúa las capacidades de los algoritmos para realizar una efectiva clasificación y se mide por clase. Por último, la métrica de F1-score se conoce como la resonancia de sistema y mide la efectividad de los algoritmos para la clasificación, así como las variaciones del sistema. Las métricas recall, precisión y F1-score se miden por clase, mientras que la

medida de accuracy representa el comportamiento general del sistema. Como se muestra en la tabla 5-7 el algoritmo que mejor detectó los dos tipos de microplástico en cada entorno es CART. De igual manera, se logró un alto rendimiento al momento de la clasificación por material, razón por la que se presentó nivel F1-score. Si bien el valor de F1-score no presenta mucha relevancia, es este apartado y para los diferentes escenarios va a hacerlo. Es importante mencionar que aunque los resultados generales (accuracy) de SVM y K-NN son iguales, SVM presentó un rendimiento relativamente menor. Esto se explica si se evalúan los resultados por clase. Como se puede ver en la tabla 5-7 las máquinas de soporte vectorial solo obtuvieron 66.7 % en el parámetro Recall, debido a que no lograron detectar todas las partículas de EPS en ninguno de los dos entornos propuestos, figuras 5-9, 5-10 y 5-11 mientras que los algoritmos CART y K-NN superaron el 80 %.

A continuación se presentan los resultados correspondientes a clasificación del microplástico (PET y EPS).

5.4. Clasificación del microplástico por tamaño en diferentes entornos

5.4.1. Clasificación de PET en agua potable

Las figuras muestran las matrices de confusión muestran los resultados de la clasificación de partículas de microplástico compuestas por PET y EPS por tamaño ($500\mu\text{m}$, $1000\mu\text{m}$ y $1400\mu\text{m}$) en agua potable.

Matriz de confusion algoritmo CART

1	5		1	
2		6		
3			6	
4				6
	1	2	3	4
	Agua potable	Agua potable MP500 μm	Agua potable MP1000 μm	Agua potable MP1400 μm

Figura 5-12.: Clasificación de microplástico PET en agua potable usando el algoritmo árboles de decisión, con una precisión del 95.8 %.

Matriz de confusion algoritmo SVM

1	6			
2		6		
3			6	
4				6
	1	2	3	4
	Agua potable	Agua potable MP500 μm	Agua potable MP1000 μm	Agua potable MP1400 μm

Figura 5-13.: Clasificación de microplástico PET en agua potable usando el algoritmo máquinas de soporte vectorial, con una precisión del 100 %

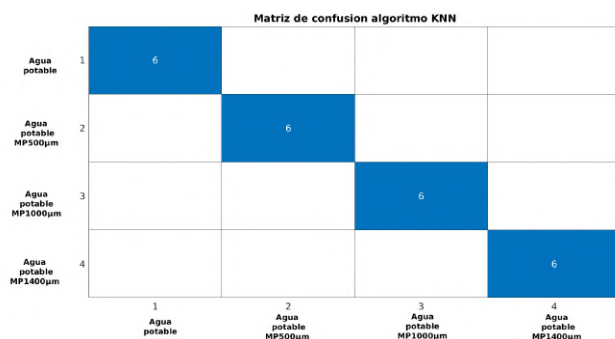


Figura 5-14.: Clasificación de microplástico PET en agua potable usando el algoritmo K como vecino más cercano, con una precisión del 100 %

Clasificación de PET en agua potable		
CART	Precisión	85.7 %
	Recall	83.3 %
	F1-score	84.48 %
	Accuracy	95.8 %
SVM	Precisión	100 %
	Recall	100 %
	F1-score	100 %
	Accuracy	100 %
K-NN	Precisión	100 %
	Recall	100 %
	F1-score	100 %
	Accuracy	100 %

Tabla 5-3.: Resultados generales para la clasificación del microplástico PET en agua potable por tamaño.

5.4.2. Clasificación de EPS en agua potable

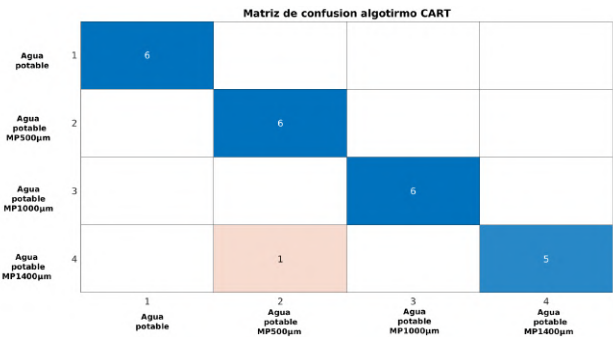


Figura 5-15.: Clasificación de microplástico EPS en agua potable usando el algoritmo árboles de decisión, con una precisión del 95.8 %.

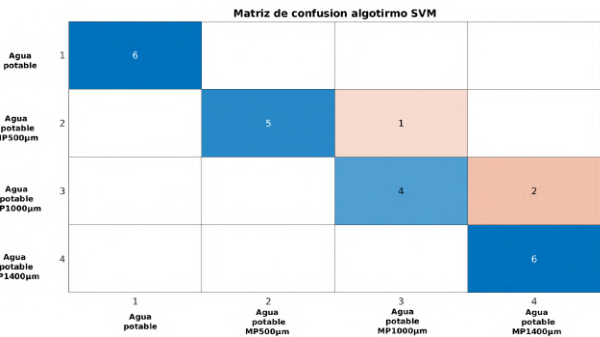


Figura 5-16.: Clasificación de microplástico PET en agua potable usando el algoritmo máquinas de soporte vectorial, con una precisión del 87.5 %

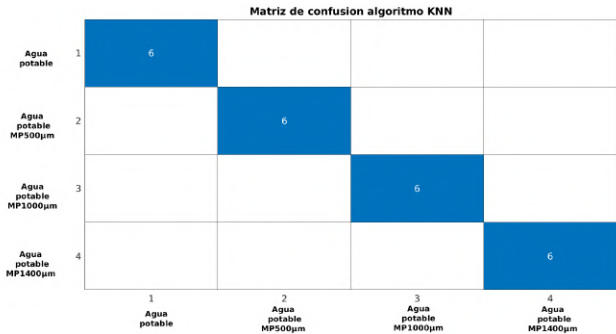


Figura 5-17.: Clasificación de microplástico PET en agua potable usando el algoritmo K como vecino más cercano, con una precisión del 91.7 %

Clasificación de EPS en agua potable		
CART	Precisión	85.7 %
	Recall	83.3 %
	F1-score	84.48 %
	Accuracy	95.8 %
SVM	Precesión	80.0 %/75 %
	Recall	83.3 %/66.7 %
	F1-score	81.61 %/70.6 %
	Accuracy	87.5 %
K-NN	Precisión	83.3 %/85.7 %
	Recall	83.3 %/83.3 %
	F1-score	83.3 %/84.6 %
	Accuracy	91.7 %

Tabla 5-4.: Resultados generales para la clasificación del microplástico EPS en agua potable por tamaño.

5.4.3. Clasificación de EPS y PET en una mezcla de agua y materia orgánica inerte

Para este experimento se agregaron partículas de microplástico PET con un tamaño de 1000 μ .

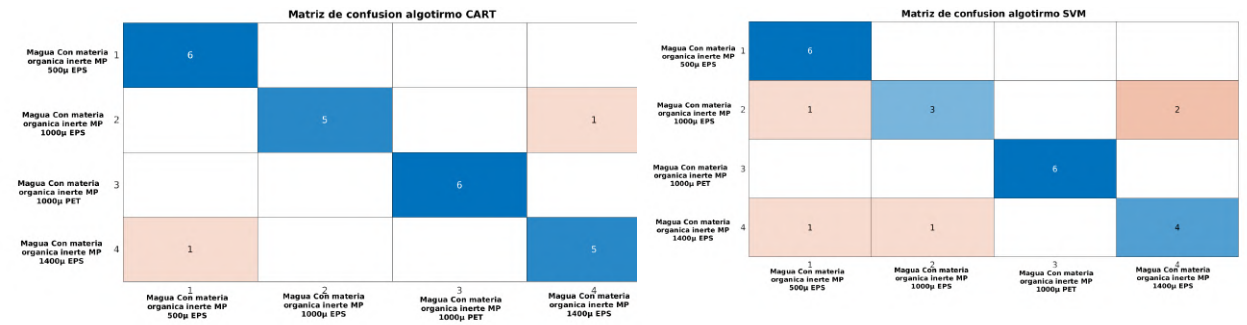


Figura 5-18.: Clasificación de microplástico EPS y PET en agua con materia orgánica inerte utilizando el algoritmo árboles de decisión, con una precisión del 91.7 %.

Figura 5-19.: Clasificación de microplástico EPS y PET en agua con materia orgánica inerte utilizando el algoritmo máquinas de soporte vectorial, con una precisión del 79.2 %

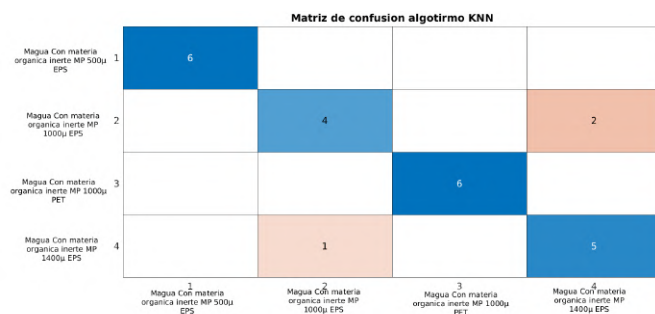


Figura 5-20.: Clasificación de microplástico EPS y PET en agua con materia orgánica inerte utilizando el algoritmo K como vecino más cercano, con una precisión del 87.5 %

Clasificación por tamaño del EPS y PET en agua con materia orgánica inerte		
CART	Precisión	85.7 %
	Recall	83.3 %
	F1-score	84.48 %
	Accuracy	91.7 %
SVM	Precisión	75 %/75 %/76.7 %
	Recall	50 %/66.7 %
	F1-score	60 %/70.6 %
	Accuracy	79.2 %
K-NN	Precisión	80 %/71.4 %
	Recall	66.7 %/83.3 %
	F1-score	72.7 %/76.8 %
	Accuracy	87.5 %

Tabla 5-5.: Resultados generales para la clasificación del microplástico EPS y PER en agua con materia orgánica inerte por tamaño.

5.4.4. Clasificación de EPS y PET en una mezcla de agua, materia orgánica inerte y materia inorgánica

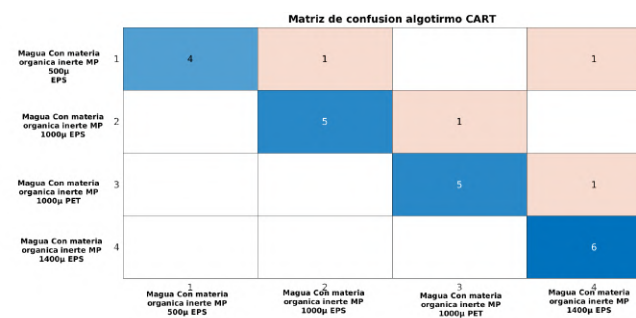


Figura 5-21.: Clasificación de microplástico EPS y PET en agua con materia orgánica inerte y materia inorgánica utilizando el algoritmo árboles de decisión, con una precisión del 83.3 %.

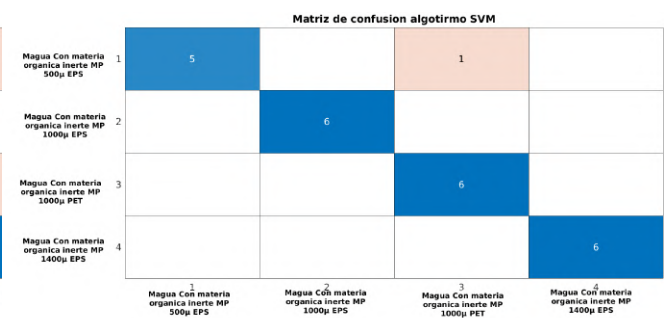


Figura 5-22.: Clasificación de microplástico EPS y PET en agua con materia orgánica inerte y materia inorgánica utilizando el algoritmo máquinas de soporte vectorial, con una precisión del 95.8 %

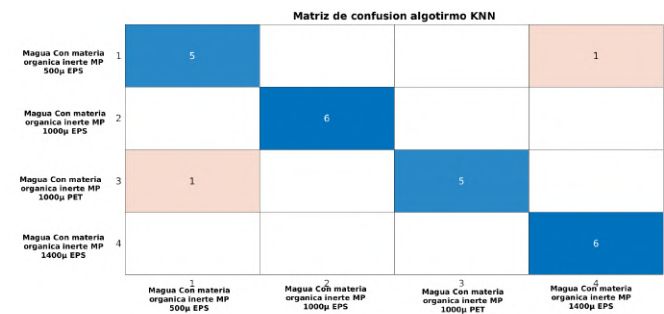


Figura 5-23.: EPS y PET en agua con materia orgánica inerte y materia inorgánica utilizando el algoritmo K como vecino más cercano, con una precisión del 91.7 %

Clasificación por tamaño del EPS y PET en agua con materia orgánica inerte y materia inorgánica		
CART	Precisión	83 %/83 %/75 %
	Recall	66.7 %/83.3 %/83.3 %
	F1-score	73.9 %/83.1 %/78.9 %
	Accuracy	83.3 %
SVM	Precisión	85.7 %
	Recall	83.3 %
	F1-score	84.48 %
	Accuracy	95.8 %
K-NN	Precisión	85.7 % / 83.3 %
	Recall	83.3 %/83.3 %
	F1-score	84.48 % /83.3 %
	Accuracy	91.7 %

Tabla 5-6.: Resultados generales para la clasificación del microplástico EPS y PET, por tamaño, en una solución compuesta por agua, materia orgánica inerte y materia inorgánica.

5.4.5. Clasificación de EPS y PET en una mezcla de agua, materia orgánica inerte, materia inorgánica y materia orgánica viva

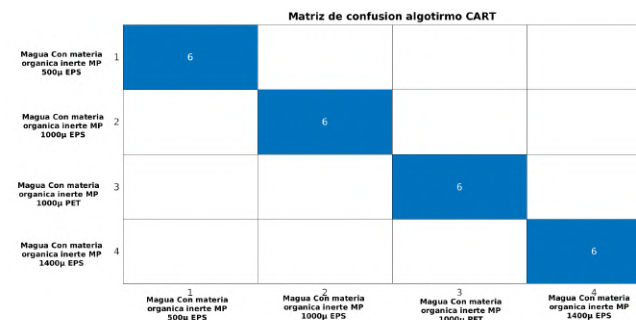


Figura 5-24.: Clasificación de microplástico EPS y PET en agua con materia orgánica inerte, materia inorgánica y materia orgánica viva utilizando el algoritmo árboles de decisión, con una precisión del 100 %.

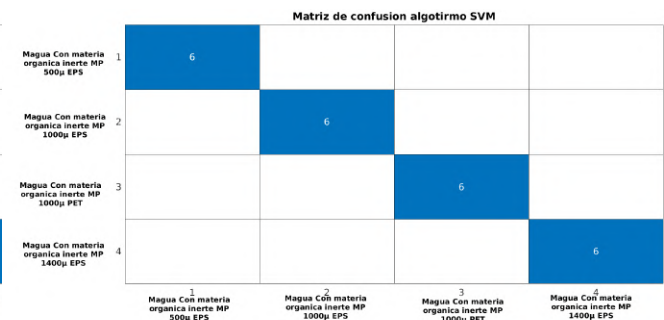


Figura 5-25.: Clasificación de microplástico EPS y PET en agua con materia orgánica inerte, materia inorgánica y materia orgánica viva utilizando el algoritmo máquinas de soporte vectorial, con una precisión del 100 %

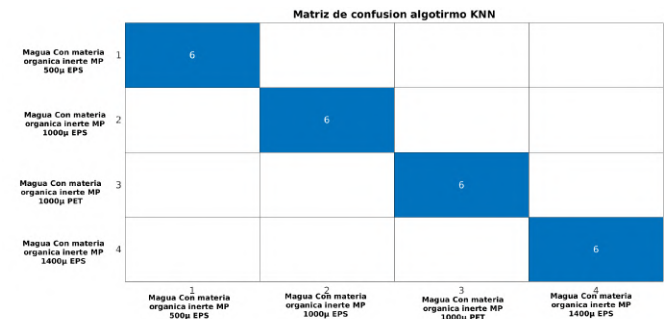


Figura 5-26.: Clasificación de microplástico EPS y PET en agua con materia orgánica inerte, materia inorgánica y materia orgánica viva utilizando el algoritmo K como vecino más cercano, con una precisión del 100 %

Clasificación por tamaño del EPS y PET en agua con materia orgánica inerte, materia inorgánica y materia orgánica viva		
CART	Precisión	100 %
	Recall	100 %
	F1-score	100 %
	Accuracy	100 %
SVM	Precisión	100 %
	Recall	100 %
	F1-score	100 %
	Accuracy	100 %
K-NN	Precisión	100 %
	Recall	100 %
	F1-score	100 %
	Accuracy	100 %

Tabla 5-7.: Resultados generales para la clasificación del microplástico EPS y PET, por tamaño, en una solución compuesta por agua, materia orgánica inerte y materia inorgánica.

6. Conclusiones y recomendaciones

6.1. Conclusiones

- En esta tesis se desarrolló un sistema compuesto por una lengua electrónica trabajando en conjunto con diversos algoritmos de inteligencia artificial que permitió diferenciar el agua, del agua con microplástico, para dos tipos de material diferente (PET y ESP), en entornos diferentes como lo son el agua potable y el agua con una variedad de contaminantes.
- En el desarrollo de esta tesis se logró definir una red de sensores con el nivel de selectividad requerido para medir de manera óptima diferentes entornos de agua con microplástico.
- Se evaluaron tres diferentes clasificadores, Máquinas de soporte vectorial (SVM), K vecinos mas cercanos (KNN) y árboles de decisión (CART). En la detección y clasificación de microplástico (PET y EPS), con resultados mayores al 95 % de precisión en detección y entre el 79 % y el 100 % de precisión en la clasificación por tamaño.
- En esta tesis se desarrolló una metodología basada en inteligencia artificial que permitió la detección y clasificación por tamaño de microplástico PET y EPS en cuatro entornos diferentes como lo son: agua potable, agua potable con materia orgánica inerte, agua potable con materia orgánica inerte y materia inorgánica y agua potable con materia orgánica inerte, materia inorgánica y materia orgánica viva.

6.2. Trabajos futuros

- Si bien se alcanzaron los objetivos propuestos en la presente tesis, es importante resaltar algunos aspectos y limitaciones que podrían aplicarse en investigaciones futuras.
- En un futuro, se debe considerar un set de datos de mayor tamaño al usado en esta tesis, esto con el objetivo de poder evaluar otro software y contar con más datos para el test.
- Otro elemento que se identificó como una limitante son los sensores, debido a que su área de detección efectiva es de tan solo 2mm, por lo que el diseño de estos elementos

es de vital importancia para entornos donde la cantidad de polímero por volumen de agua sea escaso.

A. Código

En este anexo se muestran los códigos utilizados para cargar la información correspondiente al agua potable y hacer el preprocesado de los datos para la detección del microplástico. Solo se incluyen estos dos ejemplos debido a que el proceso se es igual para cada set de datos.

Cargue datos agua

```
1 %% Agua potable.
2 clear all % Eliminar elementos del espacio de trabajo, liberando memoria
   del sistema
3 close all % Cerrar una o mas figuras
4 clc % Borrar la ventana de comandos
5 format long % Establecer formato de visualizacion de salida en larga
6
7 % Con el comado "readmatrix" se lee una matriz (o tabla de datos numericos
   )
8 % desde un archivo. Dobde 'PSD_08-09-2022-2MuxAgua' es el nombre del
9 % archivo, 'sheet' el nombre de la hoja "160" el numero de la hoja y
10 % 'Range','G3:H52' es el fragmento de informacion de se va a cargar,
11
12 Sa11=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',160,'Range','G3:H52');
13 Sa12=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',152,'Range','G3:H52');
14 Sa13=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',144,'Range','G3:H52');
15 Sa14=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',136,'Range','G3:H52');
16 Sa15=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',128,'Range','G3:H52');
17 Sa16=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',120,'Range','G3:H52');
18 Sa17=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',112,'Range','G3:H52');
19 Sa18=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',104,'Range','G3:H52');
20 Sa19=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',96,'Range','G3:H52');
21 Sa110=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',88,'Range','G3:H52');
22 Sa111=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',80,'Range','G3:H52');
23 Sa112=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',72,'Range','G3:H52');
24 Sa113=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',64,'Range','G3:H52');
25 Sa114=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',56,'Range','G3:H52');
26 Sa115=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',48,'Range','G3:H52');
27 Sa116=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',40,'Range','G3:H52');
28 Sa117=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',32,'Range','G3:H52');
29 Sa118=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',24,'Range','G3:H52');
30 Sa119=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',16,'Range','G3:H52');
31 Sa120=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',8,'Range','G3:H52');
```

```

32
33 Sa1=[Sa11 Sa12 Sa13 Sa14 Sa15 Sa16 Sa17 Sa18 Sa19 Sa110 ...
34     Sa111 Sa112 Sa113 Sa114 Sa115 Sa116 Sa117 Sa118 Sa119 Sa120]';
35
36 Sa21=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',159,'Range','G3:H52');
37 Sa22=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',151,'Range','G3:H52');
38 Sa23=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',143,'Range','G3:H52');
39 Sa24=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',135,'Range','G3:H52');
40 Sa25=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',127,'Range','G3:H52');
41 Sa26=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',119,'Range','G3:H52');
42 Sa27=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',111,'Range','G3:H52');
43 Sa28=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',103,'Range','G3:H52');
44 Sa29=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',95,'Range','G3:H52');
45 Sa210=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',87,'Range','G3:H52');
46 Sa211=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',79,'Range','G3:H52');
47 Sa212=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',71,'Range','G3:H52');
48 Sa213=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',63,'Range','G3:H52');
49 Sa214=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',55,'Range','G3:H52');
50 Sa215=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',47,'Range','G3:H52');
51 Sa216=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',39,'Range','G3:H52');
52 Sa217=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',31,'Range','G3:H52');
53 Sa218=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',23,'Range','G3:H52');
54 Sa219=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',15,'Range','G3:H52');
55 Sa220=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',7,'Range','G3:H52');
56
57 Sa2=[Sa21 Sa22 Sa23 Sa24 Sa25 Sa26 Sa27 Sa28 Sa29 Sa210 ...
58     Sa211 Sa212 Sa213 Sa214 Sa215 Sa216 Sa217 Sa218 Sa219 Sa220 ]';
59
60 Sa31=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',158,'Range','G3:H52');
61 Sa32=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',150,'Range','G3:H52');
62 Sa33=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',142,'Range','G3:H52');
63 Sa34=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',134,'Range','G3:H52');
64 Sa35=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',126,'Range','G3:H52');
65 Sa36=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',118,'Range','G3:H52');
66 Sa37=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',110,'Range','G3:H52');
67 Sa38=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',102,'Range','G3:H52');
68 Sa39=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',94,'Range','G3:H52');
69 Sa310=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',86,'Range','G3:H52');
70 Sa311=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',78,'Range','G3:H52');
71 Sa312=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',70,'Range','G3:H52');
72 Sa313=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',62,'Range','G3:H52');
73 Sa314=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',54,'Range','G3:H52');
74 Sa315=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',46,'Range','G3:H52');
75 Sa316=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',38,'Range','G3:H52');
76 Sa317=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',30,'Range','G3:H52');
77 Sa318=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',22,'Range','G3:H52');
78 Sa319=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',14,'Range','G3:H52');
79 Sa320=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',6,'Range','G3:H52');

```

```

80
81 Sa3=[Sa31 Sa32 Sa33 Sa34 Sa35 Sa36 Sa37 Sa38 Sa39 Sa310 ...
82     Sa311 Sa312 Sa313 Sa314 Sa315 Sa316 Sa317 Sa318 Sa319 Sa320 ]';
83
84 Sa41=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',157,'Range','G3:H52');
85 Sa42=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',149,'Range','G3:H52');
86 Sa43=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',141,'Range','G3:H52');
87 Sa44=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',132,'Range','G3:H52');
88 Sa45=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',125,'Range','G3:H52');
89 Sa46=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',117,'Range','G3:H52');
90 Sa47=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',109,'Range','G3:H52');
91 Sa48=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',101,'Range','G3:H52');
92 Sa49=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',93,'Range','G3:H52');
93 Sa410=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',85,'Range','G3:H52');
94 Sa411=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',77,'Range','G3:H52');
95 Sa412=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',69,'Range','G3:H52');
96 Sa413=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',61,'Range','G3:H52');
97 Sa414=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',53,'Range','G3:H52');
98 Sa415=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',45,'Range','G3:H52');
99 Sa416=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',37,'Range','G3:H52');
100 Sa417=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',29,'Range','G3:H52');
101 Sa418=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',21,'Range','G3:H52');
102 Sa419=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',13,'Range','G3:H52');
103 Sa420=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',5,'Range','G3:H52');
104
105 Sa4=[Sa41 Sa42 Sa43 Sa44 Sa45 Sa46 Sa47 Sa48 Sa49 Sa410 ...
106     Sa411 Sa412 Sa413 Sa414 Sa415 Sa416 Sa417 Sa418 Sa419 Sa420 ]';
107
108 Sa51=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',156,'Range','G3:H52');
109 Sa52=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',148,'Range','G3:H52');
110 Sa53=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',140,'Range','G3:H52');
111 Sa54=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',132,'Range','G3:H52');
112 Sa55=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',124,'Range','G3:H52');
113 Sa56=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',116,'Range','G3:H52');
114 Sa57=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',108,'Range','G3:H52');
115 Sa58=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',100,'Range','G3:H52');
116 Sa59=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',92,'Range','G3:H52');
117 Sa510=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',84,'Range','G3:H52');
118 Sa511=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',76,'Range','G3:H52');
119 Sa512=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',68,'Range','G3:H52');
120 Sa513=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',60,'Range','G3:H52');
121 Sa514=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',52,'Range','G3:H52');
122 Sa515=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',44,'Range','G3:H52');
123 Sa516=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',36,'Range','G3:H52');
124 Sa517=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',28,'Range','G3:H52');
125 Sa518=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',20,'Range','G3:H52');
126 Sa519=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',12,'Range','G3:H52');
127 Sa520=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',4,'Range','G3:H52');

```

```

128
129 Sa5=[Sa51 Sa52 Sa53 Sa54 Sa55 Sa56 Sa57 Sa58 Sa59 Sa510 ...
130      Sa511 Sa512 Sa513 Sa514 Sa515 Sa516 Sa517 Sa518 Sa519 Sa520 ]';
131
132 Sa61=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',155,'Range','G3:H52');
133 Sa62=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',147,'Range','G3:H52');
134 Sa63=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',139,'Range','G3:H52');
135 Sa64=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',131,'Range','G3:H52');
136 Sa65=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',123,'Range','G3:H52');
137 Sa66=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',115,'Range','G3:H52');
138 Sa67=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',107,'Range','G3:H52');
139 Sa68=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',99,'Range','G3:H52');
140 Sa69=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',91,'Range','G3:H52');
141 Sa610=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',83,'Range','G3:H52');
142 Sa611=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',75,'Range','G3:H52');
143 Sa612=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',67,'Range','G3:H52');
144 Sa613=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',59,'Range','G3:H52');
145 Sa614=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',51,'Range','G3:H52');
146 Sa615=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',43,'Range','G3:H52');
147 Sa616=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',35,'Range','G3:H52');
148 Sa617=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',27,'Range','G3:H52');
149 Sa618=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',19,'Range','G3:H52');
150 Sa619=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',11,'Range','G3:H52');
151 Sa620=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',3,'Range','G3:H52');
152
153 Sa6=[Sa61 Sa62 Sa63 Sa64 Sa65 Sa66 Sa67 Sa68 Sa69 Sa610 ...
154      Sa611 Sa612 Sa613 Sa614 Sa615 Sa616 Sa617 Sa618 Sa619 Sa620 ]';
155
156 Sa71=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',154,'Range','G3:H52');
157 Sa72=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',146,'Range','G3:H52');
158 Sa73=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',138,'Range','G3:H52');
159 Sa74=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',130,'Range','G3:H52');
160 Sa75=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',122,'Range','G3:H52');
161 Sa76=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',114,'Range','G3:H52');
162 Sa77=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',106,'Range','G3:H52');
163 Sa78=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',98,'Range','G3:H52');
164 Sa79=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',90,'Range','G3:H52');
165 Sa710=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',82,'Range','G3:H52');
166 Sa711=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',74,'Range','G3:H52');
167 Sa712=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',66,'Range','G3:H52');
168 Sa713=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',58,'Range','G3:H52');
169 Sa714=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',50,'Range','G3:H52');
170 Sa715=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',42,'Range','G3:H52');
171 Sa716=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',34,'Range','G3:H52');
172 Sa717=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',26,'Range','G3:H52');
173 Sa718=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',18,'Range','G3:H52');
174 Sa719=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',10,'Range','G3:H52');
175 Sa720=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',2,'Range','G3:H52');

```

```

176
177 Sa7=[Sa71 Sa72 Sa73 Sa74 Sa75 Sa76 Sa77 Sa78 Sa79 Sa710 ...
178     Sa711 Sa712 Sa713 Sa714 Sa715 Sa716 Sa717 Sa718 Sa719 Sa720 ]';
179
180 Sa81=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',153,'Range','G3:H52');
181 Sa82=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',145,'Range','G3:H52');
182 Sa83=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',137,'Range','G3:H52');
183 Sa84=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',129,'Range','G3:H52');
184 Sa85=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',121,'Range','G3:H52');
185 Sa86=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',113,'Range','G3:H52');
186 Sa87=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',105,'Range','G3:H52');
187 Sa88=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',97,'Range','G3:H52');
188 Sa89=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',89,'Range','G3:H52');
189 Sa810=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',81,'Range','G3:H52');
190 Sa811=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',73,'Range','G3:H52');
191 Sa812=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',65,'Range','G3:H52');
192 Sa813=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',57,'Range','G3:H52');
193 Sa814=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',49,'Range','G3:H52');
194 Sa815=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',41,'Range','G3:H52');
195 Sa816=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',33,'Range','G3:H52');
196 Sa817=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',25,'Range','G3:H52');
197 Sa818=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',17,'Range','G3:H52');
198 Sa819=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',9,'Range','G3:H52');
199 Sa820=readmatrix('PSD_08-09-2022-2MuxAgua','sheet',1,'Range','G3:H52');
200
201 Sa8=[Sa81 Sa82 Sa83 Sa84 Sa85 Sa86 Sa87 Sa88 Sa89 Sa810 ...
202     Sa811 Sa812 Sa813 Sa814 Sa815 Sa816 Sa817 Sa818 Sa819 Sa820 ]';
203
204 MA=[Sa1 Sa2 Sa3 Sa4 Sa5 Sa6 Sa7 Sa8];
205
206 %Guardar la medion.
207 save('/home/juan/Documentos/Tesis/P_TEIS/EPS2_0/Medicion1_Mp500uM/Agua/MA.
    mat')
208
209 % Borrar el espacio de trabajo
210
211 clear Sa1
212 clear Sa2
213 clear Sa3
214 clear Sa4
215 clear Sa5
216 clear Sa6
217 clear Sa7
218 clear Sa8
219
220 clear Sa11 Sa12 Sa13 Sa14 Sa15 Sa16 Sa17 Sa18 Sa19 Sa110...
221     Sa111 Sa112 Sa113 Sa114 Sa115 Sa116 Sa117 Sa118 Sa119 Sa120
222

```

```

223 clear Sa21 Sa22 Sa23 Sa24 Sa25 Sa26 Sa27 Sa28 Sa29 Sa210...
224     Sa211 Sa212 Sa213 Sa214 Sa215 Sa216 Sa217 Sa218 Sa219 Sa220
225
226
227 clear Sa3 Sa31 Sa32 Sa33 Sa34 Sa35 Sa36 Sa37 Sa38 Sa39 Sa310...
228     Sa311 Sa312 Sa313 Sa314 Sa315 Sa316 Sa317 Sa318 Sa319 Sa320
229
230 clear Sa41 Sa42 Sa43 Sa44 Sa45 Sa46 Sa47 Sa48 Sa49 Sa410...
231     Sa411 Sa412 Sa413 Sa414 Sa415 Sa416 Sa417 Sa418 Sa419 Sa420
232
233
234 clear Sa51 Sa52 Sa53 Sa54 Sa55 Sa56 Sa57 Sa58 Sa59 Sa510...
235     Sa511 Sa512 Sa513 Sa514 Sa515 Sa516 Sa517 Sa518 Sa519 Sa520
236
237
238 clear Sa61 Sa62 Sa63 Sa64 Sa65 Sa66 Sa67 Sa68 Sa69 Sa610...
239     Sa611 Sa612 Sa613 Sa614 Sa615 Sa616 Sa617 Sa618 Sa619 Sa620
240
241 clear Sa71 Sa72 Sa73 Sa74 Sa75 Sa76 Sa77 Sa78 Sa79 Sa710...
242     Sa711 Sa712 Sa713 Sa714 Sa715 Sa716 Sa717 Sa718 Sa719 Sa720
243
244 clear Sa81 Sa82 Sa83 Sa84 Sa85 Sa86 Sa87 Sa88 Sa89 Sa810...
245     Sa811 Sa812 Sa813 Sa814 Sa815 Sa816 Sa817 Sa818 Sa819 Sa820

```

Preprocesado datos en la detección del microplástico

```

1 % Medicion de agua potable Vs agua potable com microplasticos de tama os
2 % entre 400, 1000 y 1400um para el PET y el EPS y medicion correspondiente
3 % a la solucion de agua potable contaminada con materia organica inerte (
4   MOI),
5 % materia inorganica (MIN) y materia organica viva(MoV) Vs la misma
6   solucion
7 % con microplasticos a adidos de tama os 400,1000 y 1400um.
8
9 clear all % Eliminar elementos del espacio de trabajo, liberando memoria
10 del sistema
11 close all % Cerrar una o mas figuras
12 clc % Borrar la ventana de comandos
13
14 % Cargue de las matrices
15 load MA.mat;
16 load AguaMP400_1400_MOIMINIMOV_EPS.mat;
17 load AguaMP400_1400_MOIMINIMOV_PET.mat;
18 load AguaMP400_1400_EPS.mat;
19 load AguaMP400_1400_PET.mat;
20 load Ag_MOV_MIN_MOV.mat;
21
22 MD1 = [MA ; AguaMP400_1400_PET;AguaMP400_1400_EPS;Ag_MOV_MIN_MOV;
23   AguaMP400_1400_MOIMINIMOV_PET;AguaMP400_1400_MOIMINIMOV_EPS];%;

```

```

20
21
22 Ma=size(MA)
23 amP400_1400_EPS=size(AguaMP400_1400_EPS)
24 amP400_1400_PET=size(AguaMP400_1400_PET)
25 aG_MOV_MIN_MOV=size(Ag_MOV_MIN_MOV)
26 amP400_1400_IMINMOV_EPS=size(AguaMP400_1400_MOIMINIMOV_EPS)
27 amP400_1400_IMINMOV_PET=size(AguaMP400_1400_MOIMINIMOV_PET)
28
29
30 train = MA;
31 tes1 = MD1;
32
33 Ttrain = size(MA)
34 Ttes1 = size(MD1)
35
36
37 %%
38 %carga los datos del agua
39 TRAIN=train;
40 %carga los datos del sistema DE PRUEBA
41 TEST1=tes1;
42
43 clear tes1
44 clear train
45 o=8;
46 %%
47 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%preprocesado
48 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
49 ncomp=o;
50 [train1,test1]=preprocessing(TRAIN,TEST1,'grps',o);
51 [load_model1, scores_model1,latent_model1] = pca(train1);
52
53
54 [m1,n1] = size(latent_model1);
55
56 k1=sum(latent_model1);
57
58 %visualizacion de la distribucion de la informacion entre los compoentes.
59 for i1=1:m1
60     per_comp1(i1,1)=(100*(latent_model1(i1,1)))/k1;
61 end
62
63 Test1 = [train1;test1];
64
65 figure
66

```

```

67 bar(per_comp1);
68 ylabel('%')
69 xlabel('Componentes principales')
70 title('Porcentaje de componetes principaes')
71
72 [scores1, T1, Q1]=tqs1(test1,load_model1(:,1:ncomp),latent_model1(1:ncomp)
    );
73 %Visualizcion de los datos
74 figure
75 subplot(1,2,1);
76 plot(scores1(1:50,1),scores1(1:50,2),'bx');
77 hold on ;
78 plot(scores1(51:122,1),scores1(51:122,2),'kd');
79 hold on ;
80 legend( 'Test','Train');
81 title ('Agua potable com MP Vs Agua potable sin MP');
82 subplot(1,2,2);
83 plot(scores1(123:162,1),scores1(123:162,2),'bx');
84 hold on
85 plot(scores1(163:246,1),scores1(163:246,2),'kd');
86 hold on ;
87 title ('Agua Contminada com MP Vs Agua Contminada sin MP');
88
89 figure
90
91 plot(scores1(1:50,1),scores1(1:50,2),'bx')
92 hold on
93 plot(scores1(51:82,1),scores1(51:82,2),'ro')
94 hold on
95 plot(scores1(83:122,1),scores1(83:122,2),'g+')
96 hold on
97 plot(scores1(123:165,1),scores1(123:165,2),'c*')
98 hold on
99 plot(scores1(165:206,1),scores1(165:206,2),'y*')
100 hold on
101 plot(scores1(207:246,1),scores1(207:246,2),'ms')
102 hold on
103 legend('MA','AguaMP500-1500-PET', 'AguaMP500-1500-EPS', 'Ag-MOV-MIN-MOV',
    'AguaMP500-1500-MOIMINIMOV-EPS', 'AguaMP500-1500-MOIMINIMOV-PET')%, '
    AgMpTodoTamano')
104 title ('scores')
105
106 figure
107 plot(T1(1:50,:),Q1(1:50:),'bx')
108 hold on
109 plot(T1(51:82,:),Q1(51:82:),'ro')
110 hold on
111 plot(T1(83:122,:),Q1(83:122:),'g+')

```

```
112 hold on
113 plot(T1(123:165,:),Q1(123:165,:), 'c+')
114 hold on
115 plot(T1(165:206,:),Q1(165:206,:), 'y*')
116 hold on
117 plot(T1(207:246,:),Q1(207:246,:), 'ms')
118 hold on
119 legend('MA', 'AguaMP500-1500-PET', 'AguaMP500-1500-EPS', 'Ag-MOV-MIN-MOV',
        'AguaMP500-1500-MOIMINIMOV-EPS', 'AguaMP500-1500-MOIMINIMOV-PET')%, '
        AgMpTodoTamano')
120 title('T Vs Q')
121 %Guardar los datos
122 save('/home/juan/Documentos/Tesis/P_TEIS/
        Deteccion_MP_500_1400_diferentesEntornos__PET_EPS/
        M_TRAIN_Deteccion_MP_500_1400_DifEntornos__PET_EPS_Scores1.mat')
```

B. Abreviaturas

A continuación se muestra la lista de abreviaturas

MOI	Materia orgánica inerte
MIN	Materia inorgánica
MOV	Materia orgánica viva
MP	Microplástico
K-NN	K vecinos mas cercanos
SVM	Máquinas de soporte vectorial
CART	Árboles de clasificación y regresión
PET	Polietileno tereftalato
EPS	Poliestireno expandido
CL	Classification Learner

Tabla B-1.: Abreviaturas

C. Publicaciones

- Microplastic detection in water using a sensor network, an electronic tongue and spectroscopy impedance ISBN - 978-989-9121-25-6. Event MILAN 24th International Conference on “Innovations in Science & Technology” (MIST-23) in International center of excellence in engineering & technology (Iceeat) Milan (Italy).
- Clasificación de microplásticos usando lenguas electrónicas <https://doi.org/10.26507/paper.3126>. Evento EIE ACOFI 2023 ingeniería para transformar territorio, realizado en Cartagena de Indias (Colombia).

Bibliografía

- [1] PalmSens, “Sensor cable with banana connectors for use with mux8-r2 - palmsens.”
- [2] PalmSens, “Mux8-r2 multiplexer - palmsens.”
- [3] L. Stratmann, “Open circuit potential measurements with a mux8-r2 optimal usage of the mux8-r2’s input buffers open circuit potential measurements with a mux8-r2,” 2019.
- [4] MathWorks, “Conceptos clave de support vector machine (svm) - matlab simulink.”
- [5] J. López-Vázquez, R. Rodil, M. J. Trujillo-Rodríguez, J. B. Quintana, R. Cela, and M. Miró, “Mimicking human ingestion of microplastics: Oral bioaccessibility tests of bisphenol a and phthalate esters under fed and fasted states,” *Science of The Total Environment*, vol. 826, p. 154027, 6 2022.
- [6] C. J. Silva, A. L. Silva, C. Gravato, and J. L. Pestana, “Ingestion of small-sized and irregularly shaped polyethylene microplastics affect chironomus riparius life-history traits,” *Science of The Total Environment*, vol. 672, pp. 862–868, 7 2019.
- [7] M. Kosuth, S. A. Mason, and E. V. Wattenberg, “Anthropogenic contamination of tap water, beer, and sea salt,” *PLOS ONE*, vol. 13, p. e0194970, 4 2018.
- [8] D. Yang, H. Shi, L. Li, J. Li, K. Jabeen, and P. Kolandhasamy, “Microplastic pollution in table salts from china,” *Environmental science technology*, vol. 49, pp. 13622–13627, 10 2015.
- [9] R. C. Thompson, Y. Olson, R. P. Mitchell, A. Davis, S. J. Rowland, A. W. John, D. McGonigle, and A. E. Russell, “Lost at sea: Where is all the plastic?,” *Science*, vol. 304, p. 838, 5 2004.
- [10] S. M. Praveena, S. N. M. Shaifuddin, and S. Akizuki, “Exploration of microplastics from personal care and cosmetic products and its estimated emissions to marine environment: An evidence from malaysia,” *Marine Pollution Bulletin*, vol. 136, pp. 135–140, 11 2018.
- [11] I. E. Napper and R. C. Thompson, “Release of synthetic microplastic plastic fibres from domestic washing machines: Effects of fabric type and washing conditions,” *Marine Pollution Bulletin*, vol. 112, pp. 39–45, 11 2016.

- [12] B. C. Colson and A. P. Michel, "Flow-through quantification of microplastics using impedance spectroscopy," *ACS Sensors*, vol. 6, pp. 238–244, 1 2021.
- [13] A. M. S. A. Manel del Valle, "Sensores electroquímicos: introducción a los quimiosensores y biosensores ... - salvador alegret, manel del valle, arben merkoçi - google libros."
- [14] B. Mostafiz, S. A. Bigdeli, K. Banan, H. Afsharara, D. Hatamabadi, P. Mousavi, C. M. Hussain, R. Keçili, and F. Ghorbani-Bidkorbeh, "Molecularly imprinted polymer-carbon paste electrode (mip-cpe)-based sensors for the sensitive detection of organic and inorganic environmental pollutants: A review," *Trends in Environmental Analytical Chemistry*, vol. 32, p. e00144, 12 2021.
- [15] X. V. Chen and P. Bühlmann, "Ion-selective potentiometric sensors with silicone sensing membranes: A review," *Current Opinion in Electrochemistry*, vol. 32, p. 100896, 4 2022.
- [16] A. T. Lawal, "Recent developments in electrochemical sensors based on graphene for bioanalytical applications," *Sensing and Bio-Sensing Research*, vol. 41, p. 100571, 8 2023.
- [17] A. Bole, A. Wall, and A. Norris, "Basic radar principles," *Radar and ARPA Manual*, pp. 1–28, 1 2014.
- [18] D. A. Tibaduiza, L. E. Mujica, M. Anaya, and J. Rodellar, *Combined and I Indices Based on Principal Component Analysis for Damage Detection and Localization*. 2011.
- [19] R. P. Areny, "Sensores y acondicionadores de señal," p. 480, 2004.
- [20] M. Scampicchio, D. Ballabio, A. Arecchi, S. M. Cosio, and S. Mannino, "Amperometric electronic tongue for food analysis," 9 2008.
- [21] N. Jaffrezic-Renault and S. V. Dzyadevych, "Conductometric microbiosensors for environmental monitoring," *Sensors 2008, Vol. 8, Pages 2569-2588*, vol. 8, pp. 2569–2588, 4 2008.
- [22] L. Escuder-Gilabert and M. Peris, "Review: Highlights in recent applications of electronic tongues in food analysis," *Analytica Chimica Acta*, vol. 665, pp. 15–25, 4 2010.
- [23] G. Orlandi, R. Calvini, G. Foca, L. Pigani, G. V. Simone, and A. Ulrici, "Data fusion of electronic eye and electronic tongue signals to monitor grape ripening," *Talanta*, vol. 195, pp. 181–189, 4 2019.
- [24] C. D. Natale, R. Paolesse, A. Macagnano, A. Mantini, A. D'Amico, A. Legin, L. Lvova, A. Rudnitskaya, and Y. Vlasov, "Electronic nose and electronic tongue integration for improved classification of clinical and food samples," *Sensors and Actuators B: Chemical*, vol. 64, pp. 15–21, 6 2000.

- [25] Y. Tahara and K. Toko, "Electronic tongues-a review," *IEEE SENSORS JOURNAL*, vol. 13, p. 3001, 2013.
- [26] P. Ivarsson, Y. Kikkawa, F. Winqvist, C. Krantz-Rülcker, N.-E. Höjer, K. Hayashi, K. Toko, and I. Lundström, "Comparison of a voltammetric electronic tongue and a lipid membrane taste sensor," 2001.
- [27] F. Winqvist, P. Wide, and I. Lundström, "An electronic tongue based on voltammetry," *Analytica Chimica Acta*, vol. 357, pp. 21–31, 12 1997.
- [28] C. Söderström, F. Winqvist, and C. Krantz-Rülcker, "Recognition of six microbial species with an electronic tongue," *Sensors and Actuators B: Chemical*, vol. 89, pp. 248–255, 4 2003.
- [29] M. Palit, B. Tudu, N. Bhattacharyya, A. Dutta, P. K. Dutta, A. Jana, R. Bandyopadhyay, and A. Chatterjee, "Comparison of multivariate preprocessing techniques as applied to electronic tongue based pattern classification for black tea," *Analytica Chimica Acta*, vol. 675, pp. 8–15, 8 2010.
- [30] J. Wang, L. Zhu, W. Zhang, and Z. Wei, "Application of the voltammetric electronic tongue based on nanocomposite modified electrodes for identifying rice wines of different geographical origins," *Analytica Chimica Acta*, vol. 1050, pp. 60–70, 3 2019.
- [31] X. Cetó, S. Pérez, and B. Prieto-Simón, "Fundamentals and application of voltammetric electronic tongues in quantitative analysis," *TrAC Trends in Analytical Chemistry*, vol. 157, p. 116765, 12 2022.
- [32] L. Nuñez, X. Cetó, M. I. Pividori, M. V. Zanoni, and M. del Valle, "Development and application of an electronic tongue for detection and monitoring of nitrate, nitrite and ammonium levels in waters," *Microchemical Journal*, vol. 110, pp. 273–279, 9 2013.
- [33] L. G. Dias, A. Fernandes, A. C. Veloso, A. A. Machado, J. A. Pereira, and A. M. Peres, "Single-cultivar extra virgin olive oil classification using a potentiometric electronic tongue," *Food Chemistry*, vol. 160, pp. 321–329, 10 2014.
- [34] X. Ceto, M. Gutierrez-Capitan, D. Calvo, and M. D. Valle, "Beer classification by means of a potentiometric electronic tongue," *Food Chemistry*, vol. 141, pp. 2533–2540, 12 2013.
- [35] A. Mimendia, A. Legin, A. Merkoçi, and M. del Valle, "Use of sequential injection analysis to construct a potentiometric electronic tongue: Application to the multidetermination of heavy metals," *Sensors and Actuators B: Chemical*, vol. 146, pp. 420–426, 4 2010.

- [36] N. Major, K. Marković, M. Krpan, G. Šarić, M. Hruškar, and N. Vahčić, “Rapid honey characterization and botanical classification by an electronic tongue,” *Talanta*, vol. 85, pp. 569–574, 7 2011.
- [37] A. C. Lazanas and M. I. Prodromidis, “Electrochemical impedance spectroscopya tutorial,” *Cite This: ACS Meas. Sci. Au*, vol. 2023, pp. 162–193, 2023.
- [38] D. D. MacDonald, “Reflections on the history of electrochemical impedance spectroscopy,” *Electrochimica Acta*, vol. 51, pp. 1376–1388, 1 2006.
- [39] Y. Elamine, P. M. Inácio, B. Lyoussi, O. Anjos, L. M. Estevinho, M. da Graça Miguel, and H. L. Gomes, “Insight into the sensing mechanism of an impedance based electronic tongue for honey botanic origin discrimination,” *Sensors and Actuators B: Chemical*, vol. 285, pp. 24–33, 4 2019.
- [40] A. Riul, A. M. Soto, S. V. Mello, S. Bone, D. M. Taylor, and L. H. Mattoso, “An electronic tongue using polypyrrole and polyaniline,” *Synthetic Metals*, vol. 132, pp. 109–116, 1 2003.
- [41] Ángela González López, “Desarrollo de una lengua electrónica,” 2021.
- [42] A. Kutyła-Olesiuk, M. Nowacka, M. Wesoły, and P. Ciosek, “Evaluation of organoleptic and texture properties of dried apples by hybrid electronic tongue,” *Sensors and Actuators B: Chemical*, vol. 187, pp. 234–240, 10 2013.
- [43] M. Zabada, I. Ufnalska, K. Chreptowicz, J. Mierzejewska, W. Wróblewski, and P. Ciosek-Skibińska, “Performance of hybrid electronic tongue and hplc coupled with chemometric analysis for the monitoring of yeast biotransformation,” *Chemometrics and Intelligent Laboratory Systems*, vol. 167, pp. 69–77, 8 2017.
- [44] A. Kutyła-Olesiuk, M. Zaborowski, P. Prokaryn, and P. Ciosek, “Monitoring of beer fermentation based on hybrid electronic tongue,” *Bioelectrochemistry*, vol. 87, pp. 104–113, 10 2012.
- [45] M. Gutiérrez, C. Domingo, J. Vila-Planas, A. Ipatov, F. Capdevila, S. Demming, S. Büttgenbach, A. Llobera, and C. Jiménez-Jorquera, “Hybrid electronic tongue for the characterization and quantification of grape variety in red wines,” *Sensors and Actuators B: Chemical*, vol. 156, pp. 695–702, 8 2011.
- [46] P. Ivarsson, C. Krantz-Rülcker, F. Winqvist, and I. Lundström, “A voltammetric electronic tongue,” *Chemical Senses*, vol. 30, pp. i258–i259, 1 2005.
- [47] Z. Zhang and Y. Takane, “Multidimensional scaling,” *International Encyclopedia of Education, Third Edition*, pp. 304–311, 1 2010.

-
- [48] A. Antoniadis, J. Bigot, and S. Lambert-Lacroix, “Peaks detection and alignment for mass spectrometry data,” *Journal de la société française de statistique*, vol. 151, pp. 17–37, 2010.
- [49] M. Navarro-Reig, J. Jaumot, T. A. van Beek, G. Vivó-Truyols, and R. Tauler, “Chemometric analysis of comprehensive $lc \times lc$ -ms data: Resolution of triacylglycerol structural isomers in corn oil,” *Talanta*, vol. 160, pp. 624–635, 11 2016.
- [50] X. Cetó, F. Céspedes, and M. del Valle, “Comparison of methods for the processing of voltammetric electronic tongues data,” *Microchimica Acta*, vol. 180, pp. 319–330, 4 2013.
- [51] M. Sun, C. C. Liu, J. Yang, Z. Jin, and J. Yang, “A two-step framework for highly nonlinear data unfolding,” *Neurocomputing*, vol. 73, pp. 1801–1807, 6 2010.
- [52] J. A. Arancibia, A. C. Olivieri, D. B. Gil, A. E. Mansilla, I. Durán-Merás, and A. M. D. L. Peña, “Trilinear least-squares and unfolded-pls coupled to residual trilinearization: New chemometric tools for the analysis of four-way instrumental data,” *Chemometrics and Intelligent Laboratory Systems*, vol. 80, pp. 77–86, 1 2006.
- [53] M. del Valle, “Sensor arrays and electronic tongue systems,” *International Journal of Electrochemistry*, vol. 2012, pp. 1–11, 2012.
- [54] G. Liang, Z. He, J. Zhen, H. Tian, L. Ai, L. Pan, and W. Gong, “Development of the screen-printed electrodes: A mini review on the application for pesticide detection,” *Environmental Technology Innovation*, vol. 28, p. 102922, 11 2022.
- [55] BVT, “Ac1 electrochemical sensor.”
- [56] V. Meiler, J. Pfeiffer, L. Bifano, C. Kandlbinder-Paret, and G. Fischerauer, “Approaches to detect microplastics in water using electrical impedance measurements and support vector machines,” *IEEE Sensors Journal*, vol. 23, pp. 4863–4872, 3 2023.