

Herramienta basada en procesamiento digital de señales y aprendizaje automático para la clasificación de trazas sísmicas volcánicas

Yair Mauad Sosa

Ingeniería Electrónica con Orientación en Sistemas Digitales
Departamento de Electrónica - FCFMyN
2023



Directora: Romina Soledad Molina, PhD
Co-Director: Esp. Ing. Alejandro Nuñez Manquez

Temario

- **Objetivos**
- **Caso de estudio**
- **Flujo general de trabajo**
- **Armado del conjunto de datos**
 - Pre-procesamiento
 - Obtención de características
 - Clusterizado de señales
 - Etiquetado
- **Clasificador basado en ML**
 - Señales crudas
 - Características
- **Resultados**
- **Conclusiones y trabajos futuros**



Objetivos

1. Estudio de las características más relevantes para la representación de señales sísmicas volcánicas, para su posterior extracción.

Objetivos

1. Estudio de las características más relevantes para la representación de señales sísmicas volcánicas, para su posterior extracción.
2. Exploración de técnicas de procesamiento digital de señales para la extracción de características de las señales sísmicas volcánicas.

Objetivos

1. Estudio de las características más relevantes para la representación de señales sísmicas volcánicas, para su posterior extracción.
2. Exploración de técnicas de procesamiento digital de señales para la extracción de características de las señales sísmicas volcánicas.
3. Generación de un conjunto de datos a ser empleado por técnicas de aprendizaje automático para clasificación de señales en 1 dimensión.

Objetivos

1. Estudio de las características más relevantes para la representación de señales sísmicas volcánicas, para su posterior extracción.
2. Exploración de técnicas de procesamiento digital de señales para la extracción de características de las señales sísmicas volcánicas.
3. Generación de un conjunto de datos a ser empleado por técnicas de aprendizaje automático para clasificación de señales en 1 dimensión.
4. Facilitar, mediante una herramienta, el procesamiento por lote de señales sísmicas volcánicas, reduciendo los tiempos de análisis manual.

Objetivos

1. Estudio de las características más relevantes para la representación de señales sísmicas volcánicas, para su posterior extracción.
2. Exploración de técnicas de procesamiento digital de señales para la extracción de características de las señales sísmicas volcánicas.
3. Generación de un conjunto de datos a ser empleado por técnicas de aprendizaje automático para clasificación de señales en 1 dimensión.
4. Facilitar, mediante una herramienta, el procesamiento por lote de señales sísmicas volcánicas, reduciendo los tiempos de análisis manual.
5. Aprender a utilizar Python como lenguaje de programación para la implementación de la herramienta, así como librerías asociadas al procesamiento de trazas provenientes de eventos sísmicos (obspy).

Objetivos

1. Estudio de las características más relevantes para la representación de señales sísmicas volcánicas, para su posterior extracción.
2. Exploración de técnicas de procesamiento digital de señales para la extracción de características de las señales sísmicas volcánicas.
3. Generación de un conjunto de datos a ser empleado por técnicas de aprendizaje automático para clasificación de señales en 1 dimensión.
4. Facilitar, mediante una herramienta, el procesamiento por lote de señales sísmicas volcánicas, reduciendo los tiempos de análisis manual.
5. Aprender a utilizar Python como lenguaje de programación para la implementación de la herramienta, así como librerías asociadas al procesamiento de trazas provenientes de eventos sísmicos (obspy).
6. Aprender el uso del paquete hls4ml para el mapeo de redes neuronales en FPGA.

Colaboraciones

Este trabajo se realizó utilizando datos de la Red del Volcán Copahue; instalada y mantenida por el Laboratorio de estudios y seguimientos de volcanes activos (LESVA) de la Universidad Nacional de Río Negro, bajo la dirección de Alberto Caselli.

Colaboraciones

Este trabajo se realizó utilizando datos de la Red del Volcán Copahue; instalada y mantenida por el Laboratorio de estudios y seguimientos de volcanes activos (LESVA) de la Universidad Nacional de Río Negro, bajo la dirección de Alberto Caselli.

Colaboración interdisciplinaria entre los Departamentos de Electrónica y Física de la Universidad Nacional de San Luis, en conjunto con Multidisciplinary Laboratory (MLab)/ STI Unit perteneciente a The International Centre for Theoretical Physics (ICTP)

Colaboraciones

Este trabajo se realizó utilizando datos de la Red del Volcán Copahue; instalada y mantenida por el Laboratorio de estudios y seguimientos de volcanes activos (LESVA) de la Universidad Nacional de Río Negro, bajo la dirección de Alberto Caselli.

Colaboración interdisciplinaria entre los Departamentos de Electrónica y Física de la Universidad Nacional de San Luis, en conjunto con Multidisciplinary Laboratory (MLab)/ STI Unit perteneciente a The International Centre for Theoretical Physics (ICTP)

Consultores externos:

- Dra Silvana Spagnotto (UNSL)
- Dr Ivan Melchor (UNRN)

Caso de Estudio

Volcán Copahue

Volcán activo ubicado en la frontera entre Argentina y Chile.

Su cima está a una altitud de 2,997 metros sobre el nivel del mar. Copahue forma parte del cinturón volcánico de los Andes y es uno de los volcanes más activos del sur de los Andes.



República Argentina



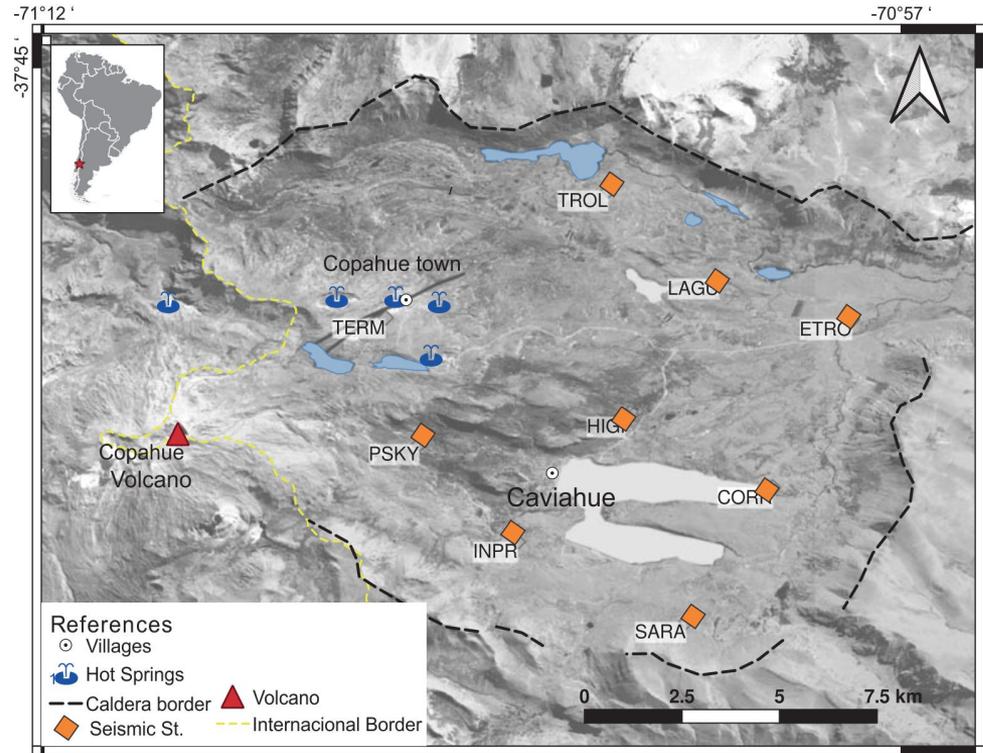
Fuente: www.caviahue-copahue.gob.ar

Caso de Estudio

Estaciones de Monitoreo

Ocho (8) estaciones sísmicas de banda ancha.

Sensores triaxiales de 20 y 30 segundos (Nanometrics Trillium Compact 20s y Guralp 40T 30s) y digitalizadores con sistema de almacenamiento en memorias tipo flash y sistema de tiempo y posicionamiento GPS (Nanometrics Centaur, Guralp DAS DM-24, y Sara SL06).

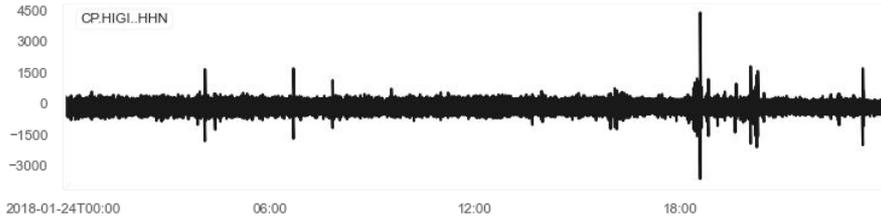


Caso de Estudio

Señal cruda completa

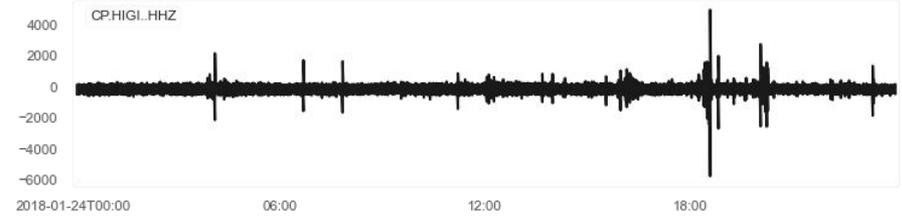
Canal N (Norte-Sur)

2018-01-24T00:00:00 - 2018-01-24T23:59:59.99



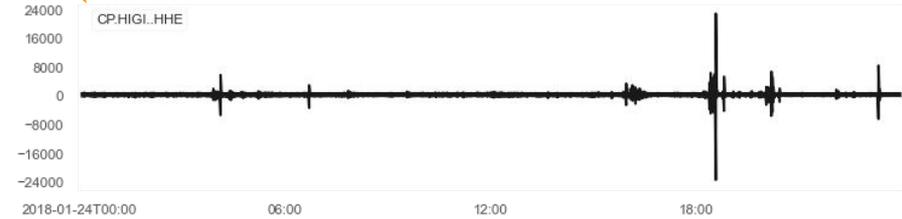
Canal Z (vertical)

2018-01-24T00:00:00 - 2018-01-24T23:59:59.99

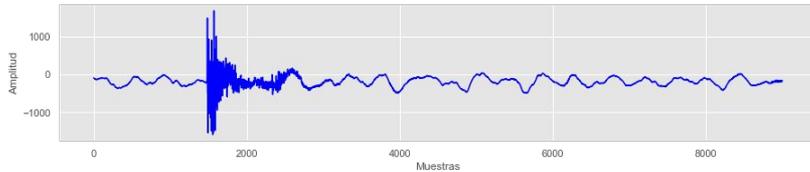


Canal E (Este - Oeste)

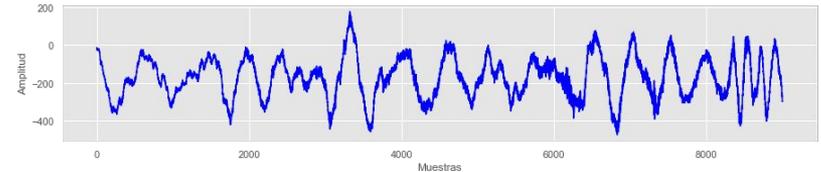
2018-01-24T00:00:00 - 2018-01-24T23:59:59.99



Evento (9,000 muestras)



No evento (9,000 muestras)

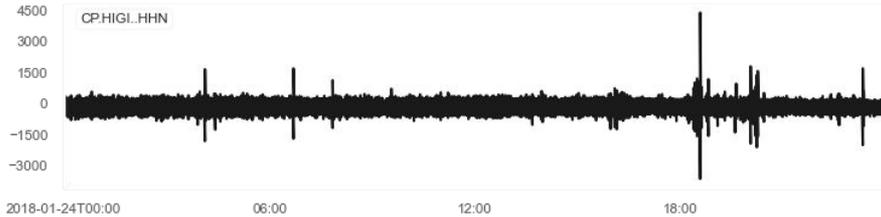


Caso de Estudio

Señal cruda completa

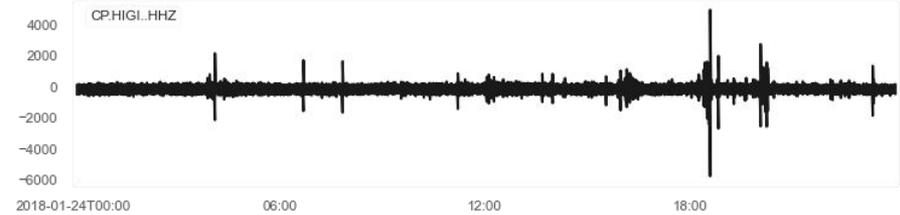
Canal N (Norte-Sur)

2018-01-24T00:00:00 - 2018-01-24T23:59:59.99



Canal Z (vertical)

2018-01-24T00:00:00 - 2018-01-24T23:59:59.99



Canal E (Este - Oeste)

2018-01-24T00:00:00 - 2018-01-24T23:59:59.99



Se trabaja con las señales de los canales E, N y Z de la estación **HIGI**.

- Cercana al volcán y se encuentra dentro del GrupoA descrito en el TF de Montenegro, 2019.
- Funcionó correctamente durante todo el periodo de adquisición de datos.

Diagrama de flujo general



Dataset Volcán
Copahue

Diagrama de flujo general

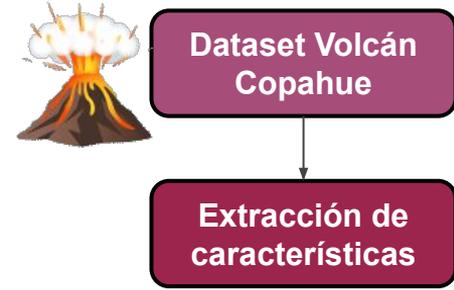


Diagrama de flujo general

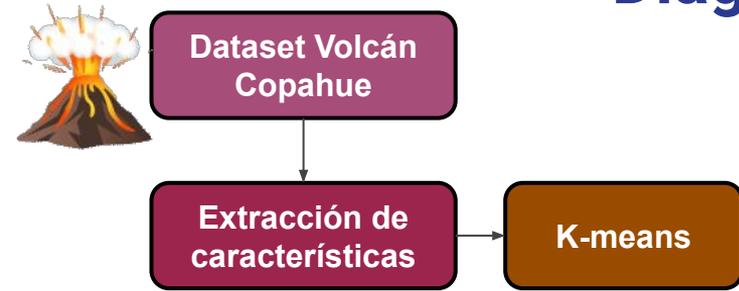


Diagrama de flujo general



Dataset Volcán Copahue

Extracción de características

K-means

Etiquetado del conjunto de datos



Diagrama de flujo general

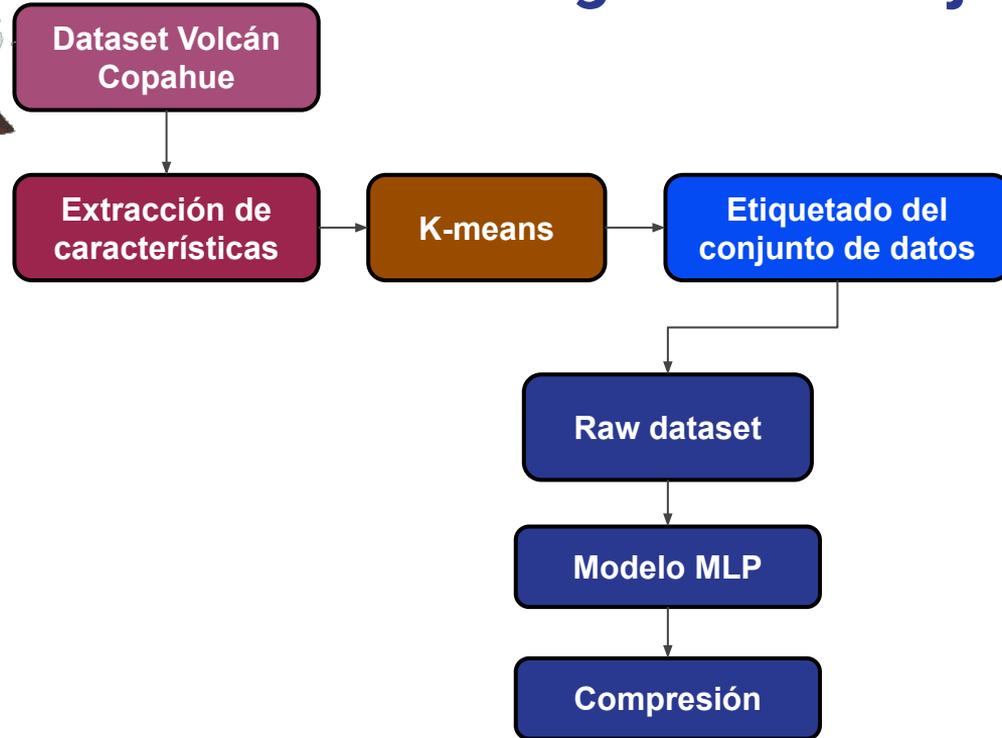


Diagrama de flujo general

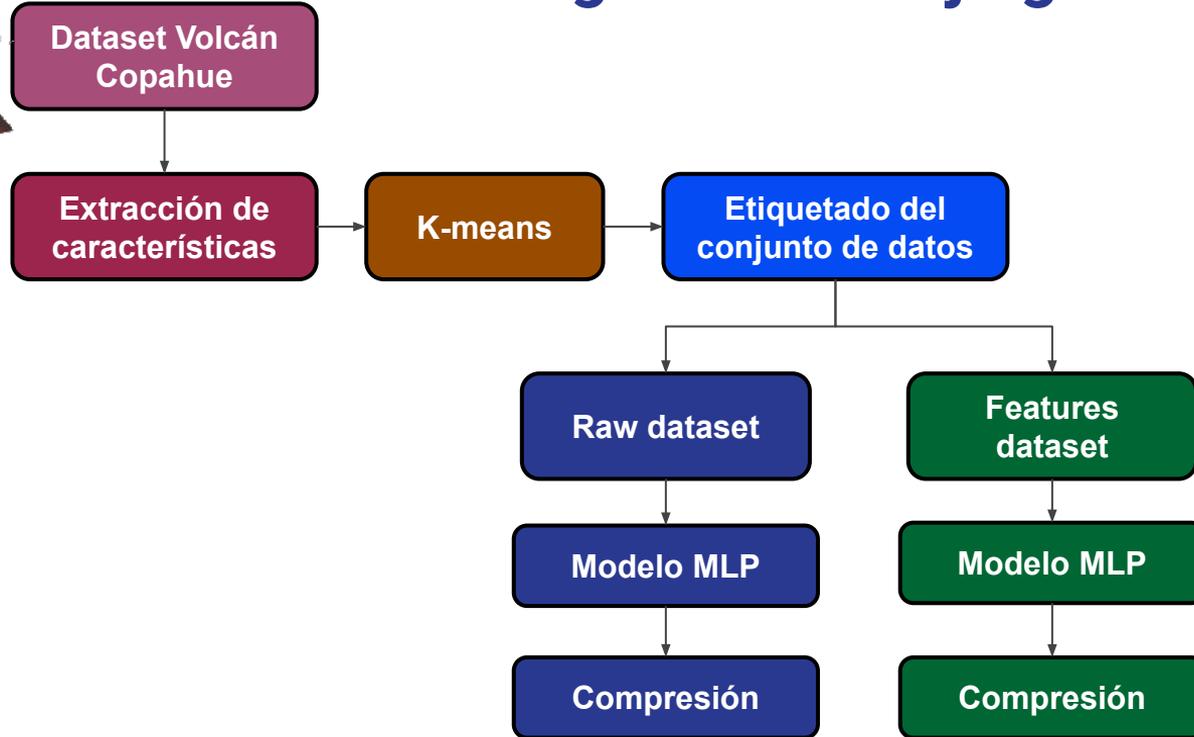
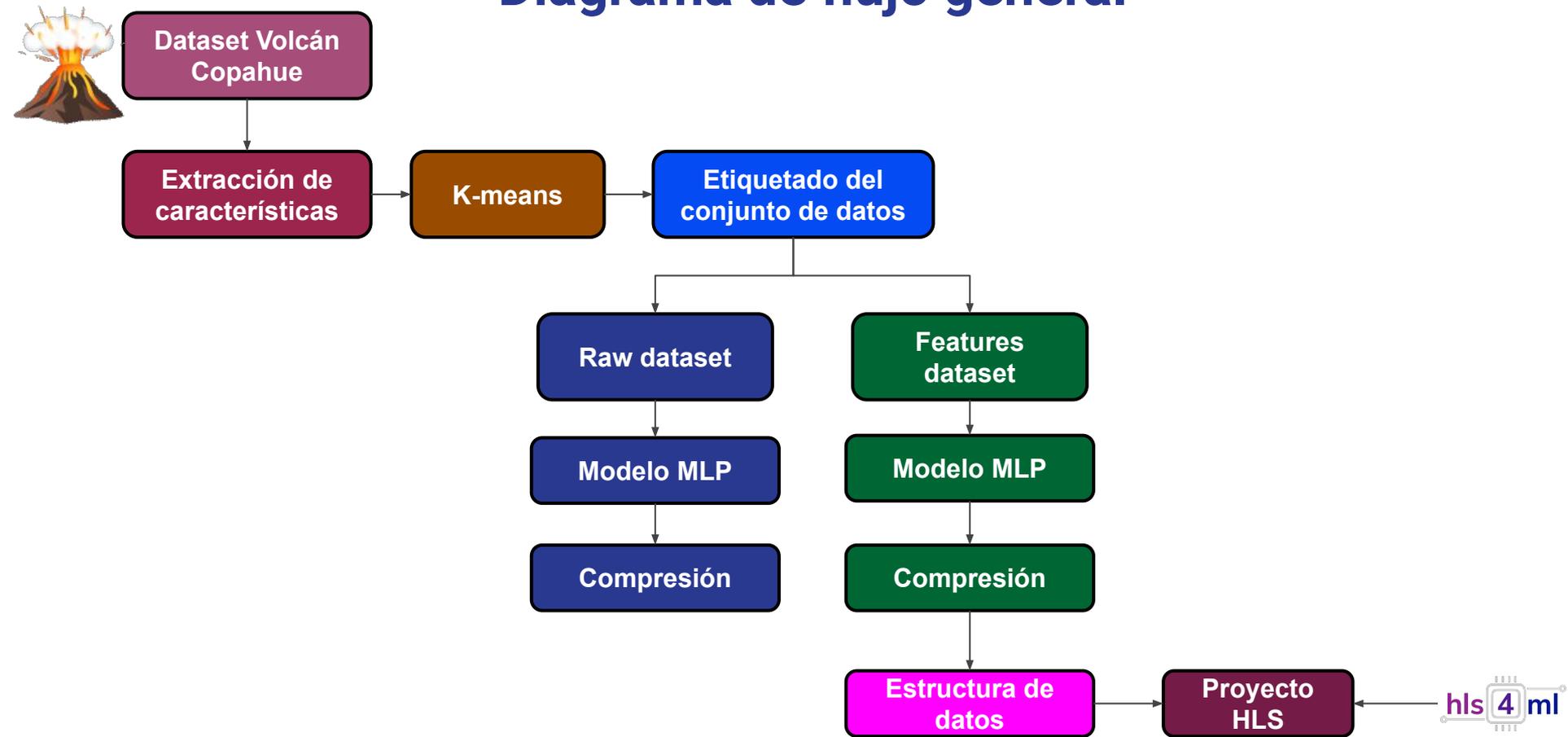


Diagrama de flujo general



Extracción de características y clusterizado



Diagrama de flujo preprocesamiento

Extracción de
características

Diagrama de flujo preprocesamiento

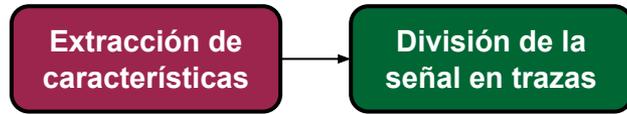


Diagrama de flujo preprocesamiento

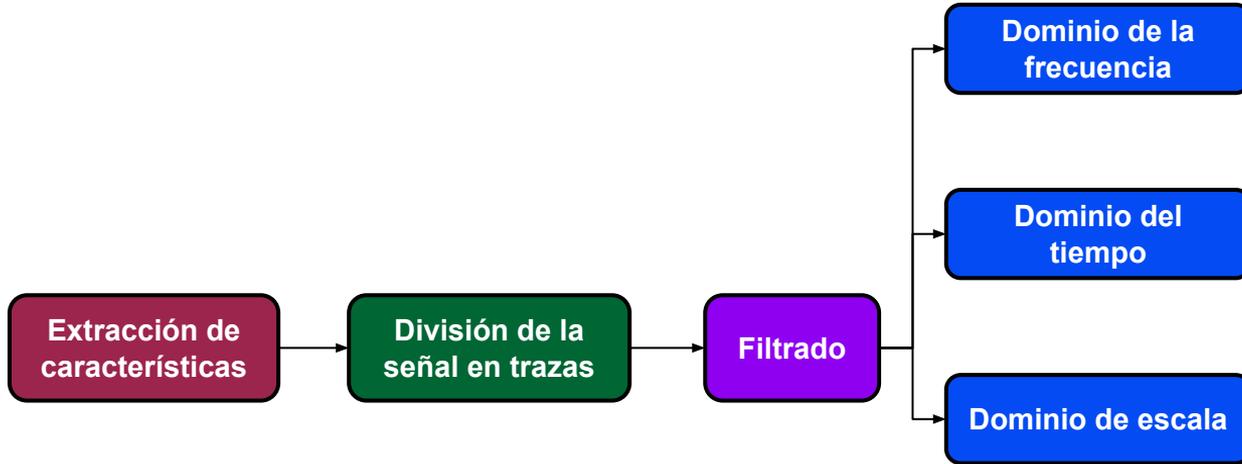


Diagrama de flujo preprocesamiento

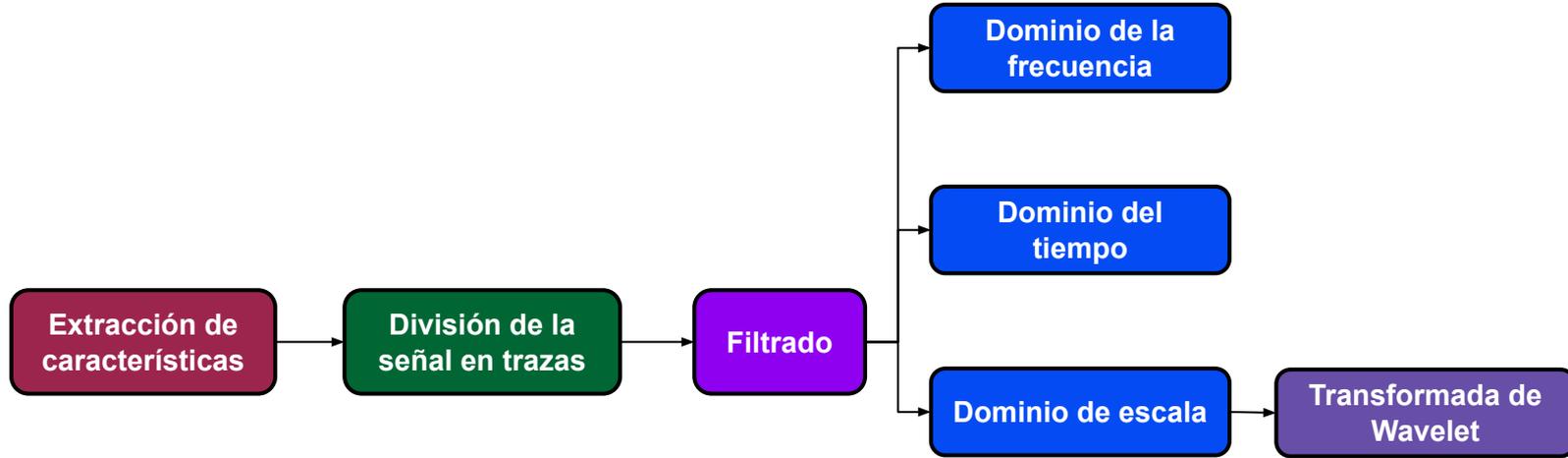


Diagrama de flujo preprocesamiento

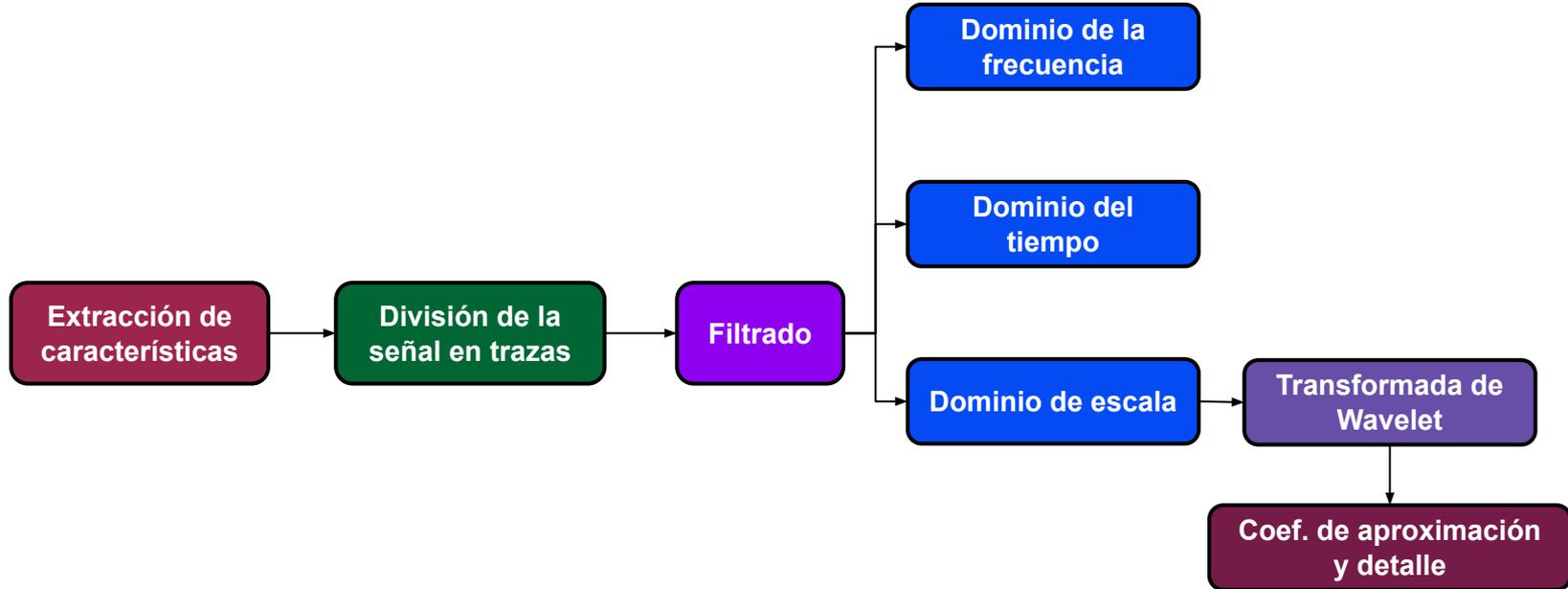
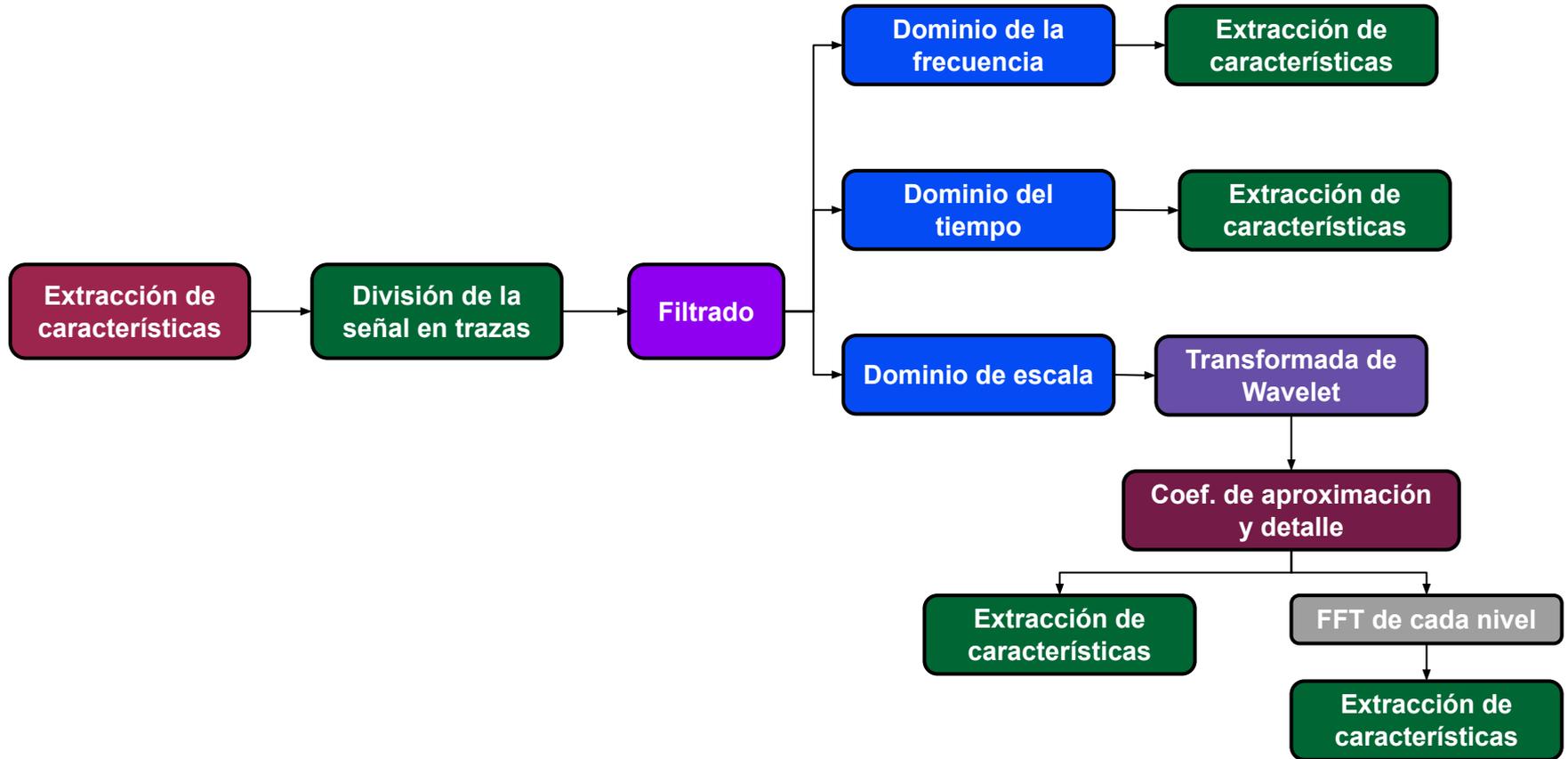


Diagrama de flujo preprocesamiento

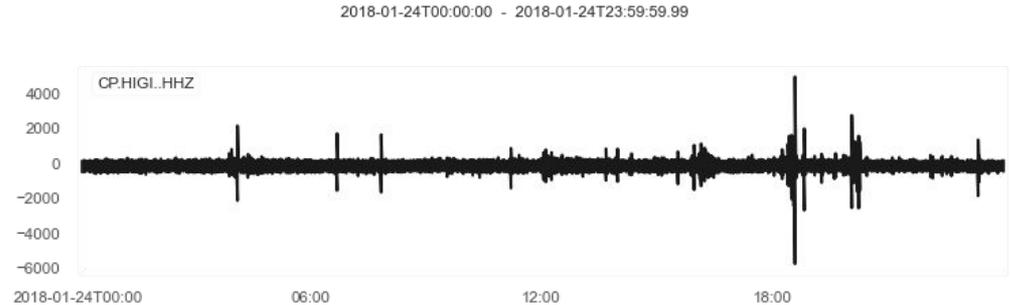


Preprocesamiento

División de la señal original en trazas

- Señal original:
 - 8,640,000 muestras

Señal original completa

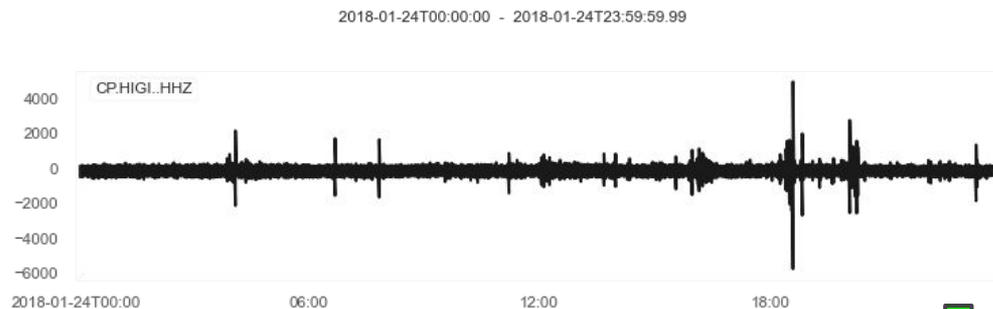


Preprocesamiento

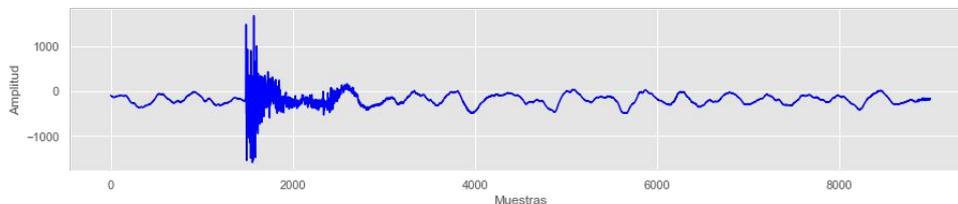
División de la señal original en trazas

- Señal original:
 - 8,640,000 muestras
- Nueva traza:
 - 9,000 muestras
 - 960 trazas resultantes

Señal original completa

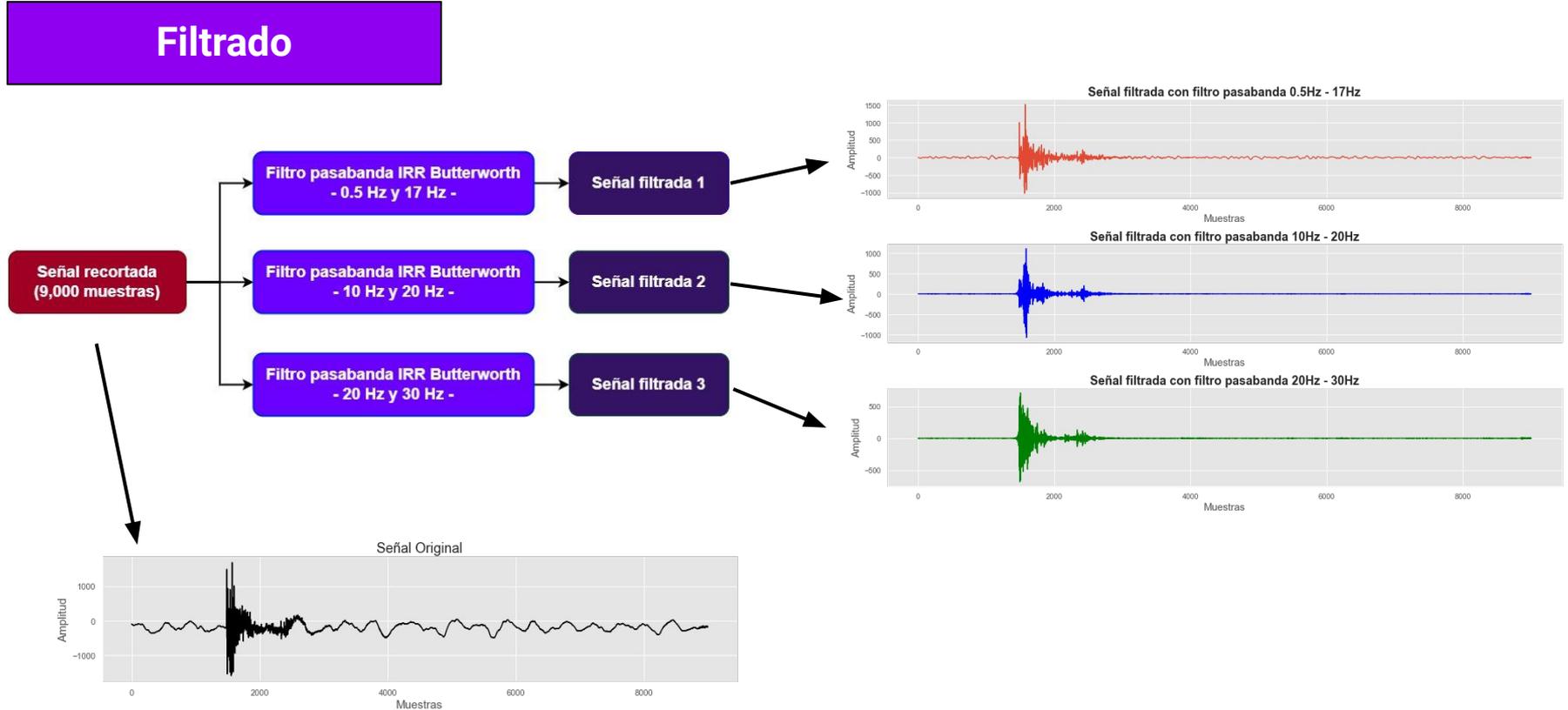


Traza



Preprocesamiento

Filtrado



Preprocesamiento

Dominio del tiempo

- Curtosis
- Media cuadrática (RMS)
- Media
- Mínimo
- Máximo
- Tiempo máximo
- Energía
- Diferencia entre máximo y mínimo
- Diferencia entre máximo y RMS



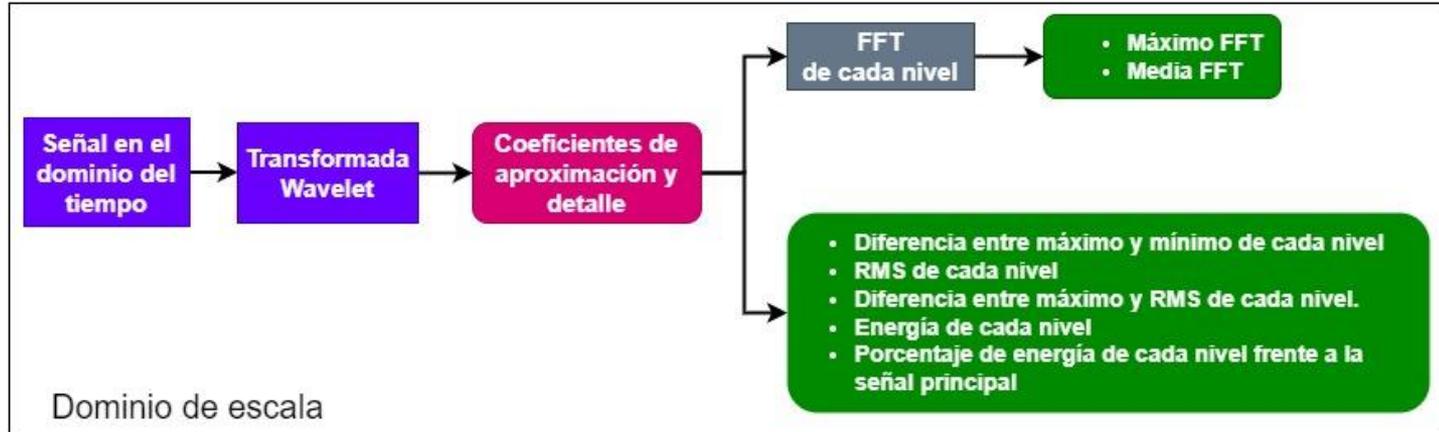
Dominio de la frecuencia

- Amplitud máxima
- Frecuencia máxima
- Media
- Máximo en frec. 10-20Hz
- Máximo en frec. 20-30Hz
- RMS
- Diferencia entre máximo y RMS
- Energía



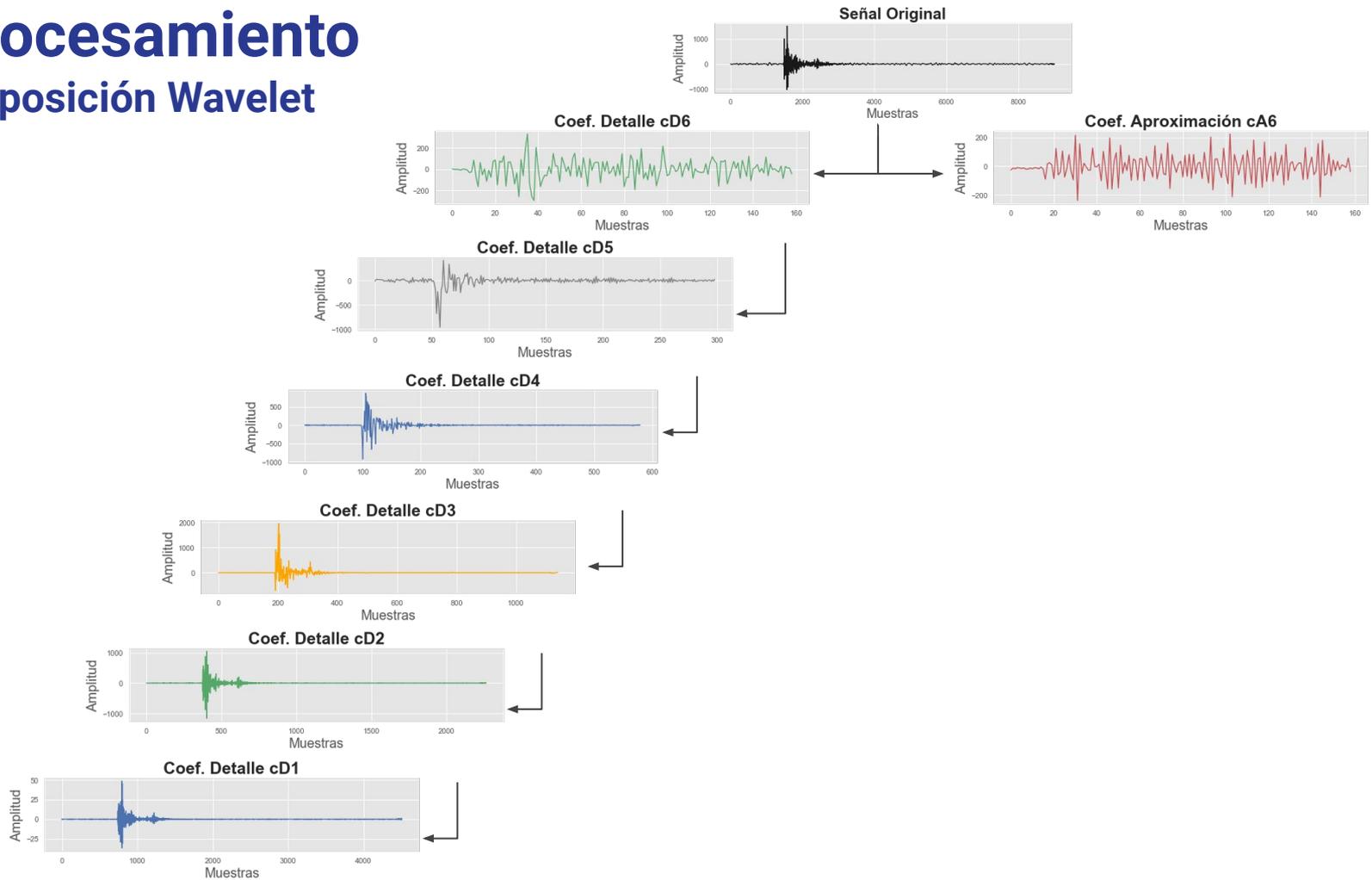
Preprocesamiento

Dominio de escala



Preprocesamiento

Descomposición Wavelet



Características extraídas

Dominio del tiempo

ft1	Curtosis	ft4	Máximo	ft7	Energía
ft2	Media cuadrática (RMS)	ft5	Mínimo	ft8	Diferencia entre máximo y mínimo
ft3	Media	ft6	Tiempo máximo	ft9	Diferencia entre máximo y RMS

Dominio de la frecuencia

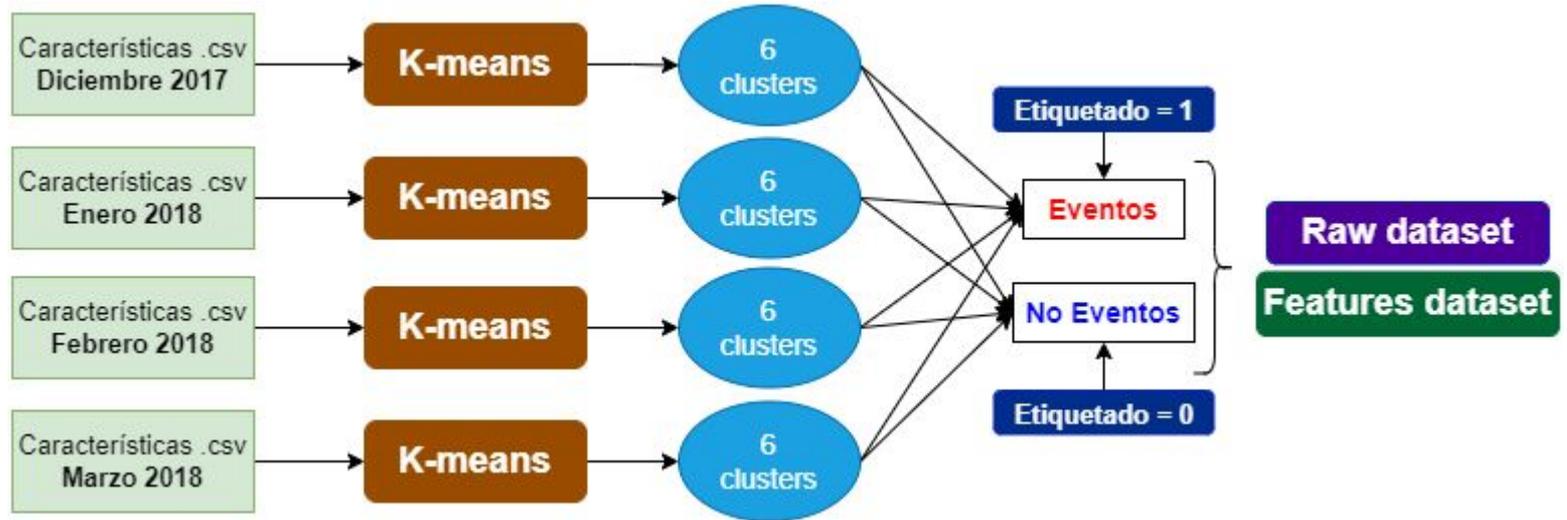
ft10	Amplitud máxima	ft13	Máximo valor en frecuencia 10-20Hz	ft16	Diferencia entre máximo y RMS
ft11	Frecuencia máxima	ft14	Máximo valor en frecuencia 20-30Hz	ft17	Energía
ft12	Media	ft15	RMS		

Dominio de escala

ft18	Diferencia entre máximo y mínimo A6	ft35	Diferencia entre máximo y RMS D4	ft52	Porcentaje de energía D1
ft19	Diferencia entre máximo y mínimo D6	ft36	Diferencia entre máximo y RMS D3	FFT Wavelet	
ft20	Diferencia entre máximo y mínimo D5	ft37	Diferencia entre máximo y RMS D2	ft53	Máximo A6
ft21	Diferencia entre máximo y mínimo D4	ft38	Diferencia entre máximo y RMS D1	ft54	Máximo D6
ft22	Diferencia entre máximo y mínimo D3	ft39	Energía total A6	ft55	Máximo D5
ft23	Diferencia entre máximo y mínimo D2	ft40	Energía D6	ft56	Máximo D4
ft24	Diferencia entre máximo y mínimo D1	ft41	Energía D5	ft57	Máximo D3
ft25	Media cuadrática A6	ft42	Energía D4	ft58	Máximo D2
ft26	Media cuadrática D6	ft43	Energía D3	ft59	Máximo D1
ft27	Media cuadrática D5	ft44	Energía D2	ft60	Media A6
ft28	Media cuadrática D4	ft45	Energía D1	ft61	Media D6
ft29	Media cuadrática D3	ft46	Porcentaje de energía A6	ft62	Media D4
ft30	Media cuadrática D2	ft47	Porcentaje de energía D6	ft63	Media D3
ft31	Media cuadrática D1	ft48	Porcentaje de energía D6	ft64	Media D2
ft32	Diferencia entre máximo y RMS A6	ft49	Porcentaje de energía D4	ft65	Media D1
ft33	Diferencia entre máximo y RMS D6	ft50	Porcentaje de energía D3	ft66	Media D6
ft34	Diferencia entre máximo y RMS D5	ft51	Porcentaje de energía D2		

Etapa de clusterizado mediante K-means

K-means



Clasificador basado en aprendizaje automático

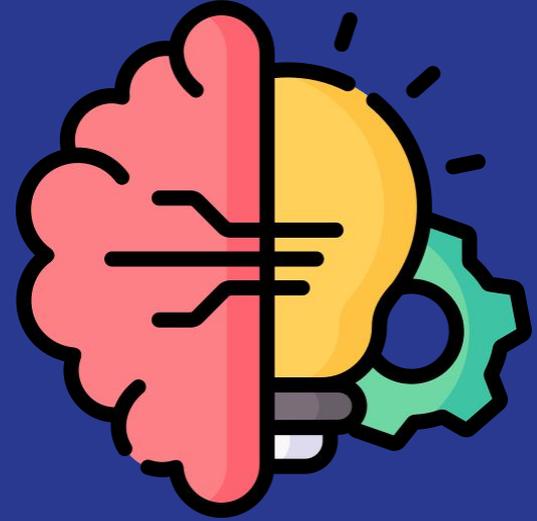
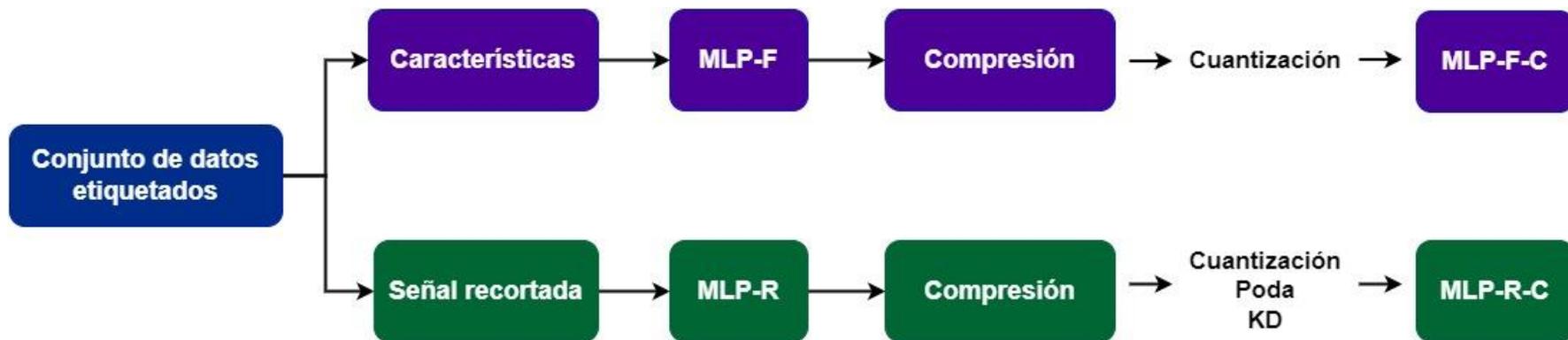


Diagrama de flujo del entrenamiento



Entrenamiento con el dataset de características

Características



```
graph LR; A[Características] --> B[ ]
```

Entrenamiento con el dataset de características

Características

Correlación

CARACTERÍSTICAS CORRELACIONADAS

Dominio del tiempo

ft1 Curtosis
ft3 Media
ft5 Mínimo
ft6 Tiempo máximo

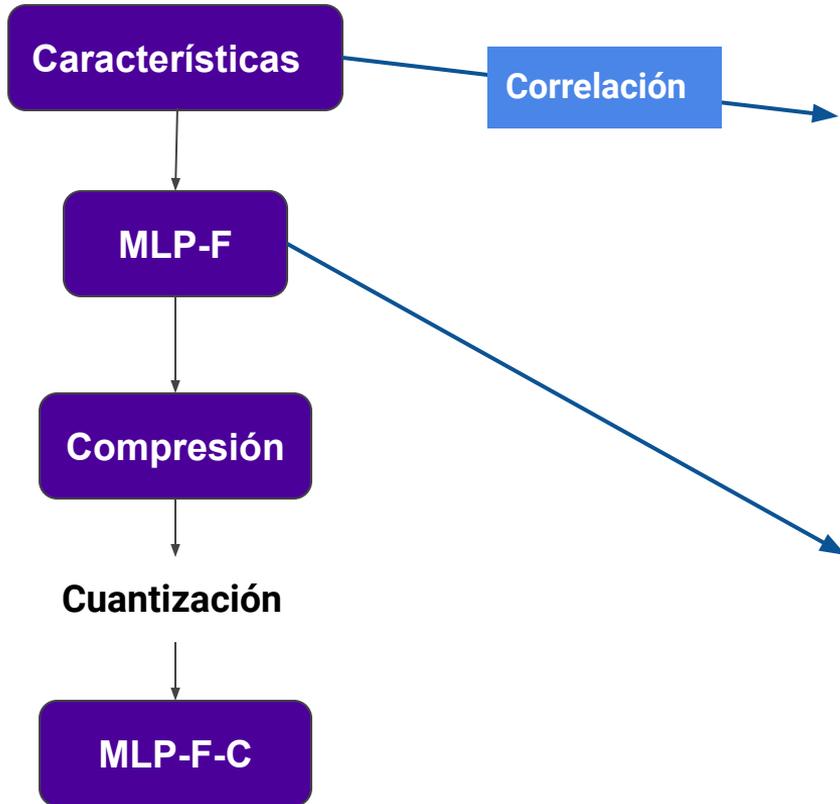
Dominio de la frecuencia

ft11 Frecuencia máxima
ft13 Máximo valor en frecuencia 10-20Hz
ft14 Máximo valor en frecuencia 20-30Hz

Dominio de escala

ft25 Media cuadrática A6
ft46 Porcentaje de energía A6
ft47 Porcentaje de energía D6
ft48 Porcentaje de energía D6
ft49 Porcentaje de energía D4
ft50 Porcentaje de energía D3
ft51 Porcentaje de energía D2

Entrenamiento con el dataset de características



CARACTERÍSTICAS CORRELACIONADAS

Dominio del tiempo

ft1 Curtosis
ft3 Media
ft5 Mínimo
ft6 Tiempo máximo

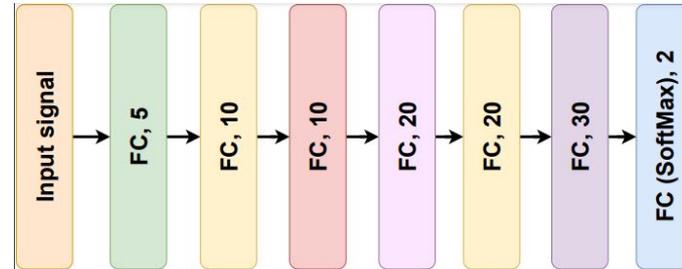
Dominio de escala

ft25 Media cuadrática A6
ft46 Porcentaje de energía A6
ft47 Porcentaje de energía D6
ft48 Porcentaje de energía D6
ft49 Porcentaje de energía D4
ft50 Porcentaje de energía D3
ft51 Porcentaje de energía D2

Dominio de la frecuencia

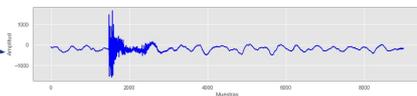
ft11 Frecuencia máxima
ft13 Máximo valor en frecuencia 10-20Hz
ft14 Máximo valor en frecuencia 20-30Hz

ARQUITECTURA MLP-F

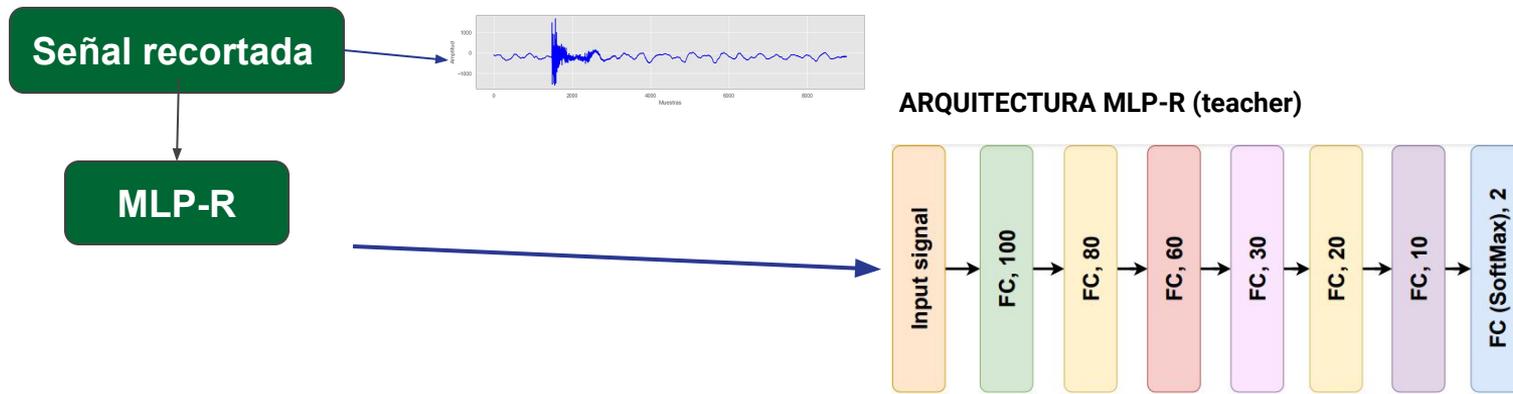


Entrenamiento con el dataset de señales

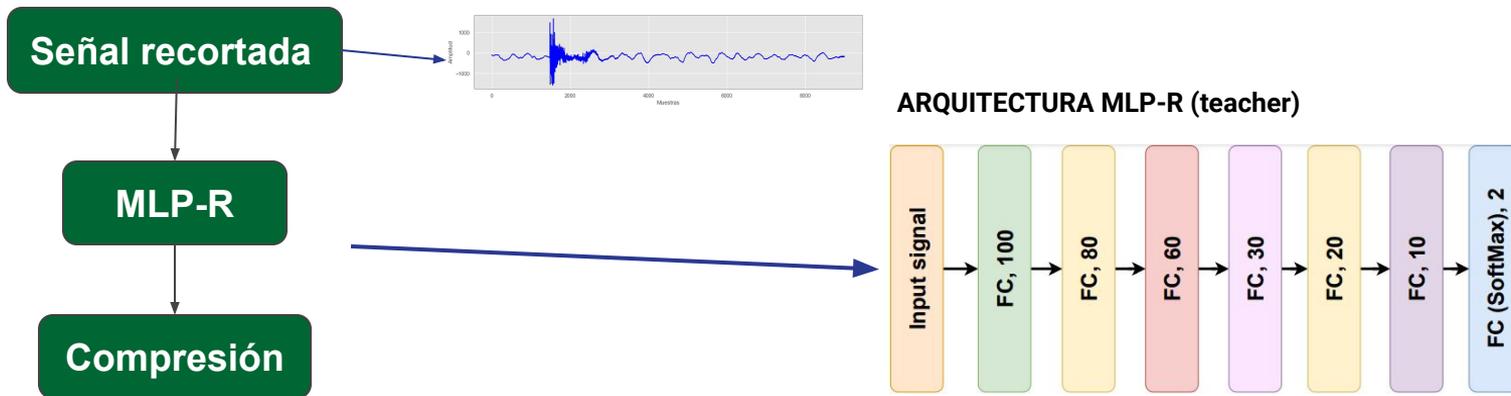
Señal recortada



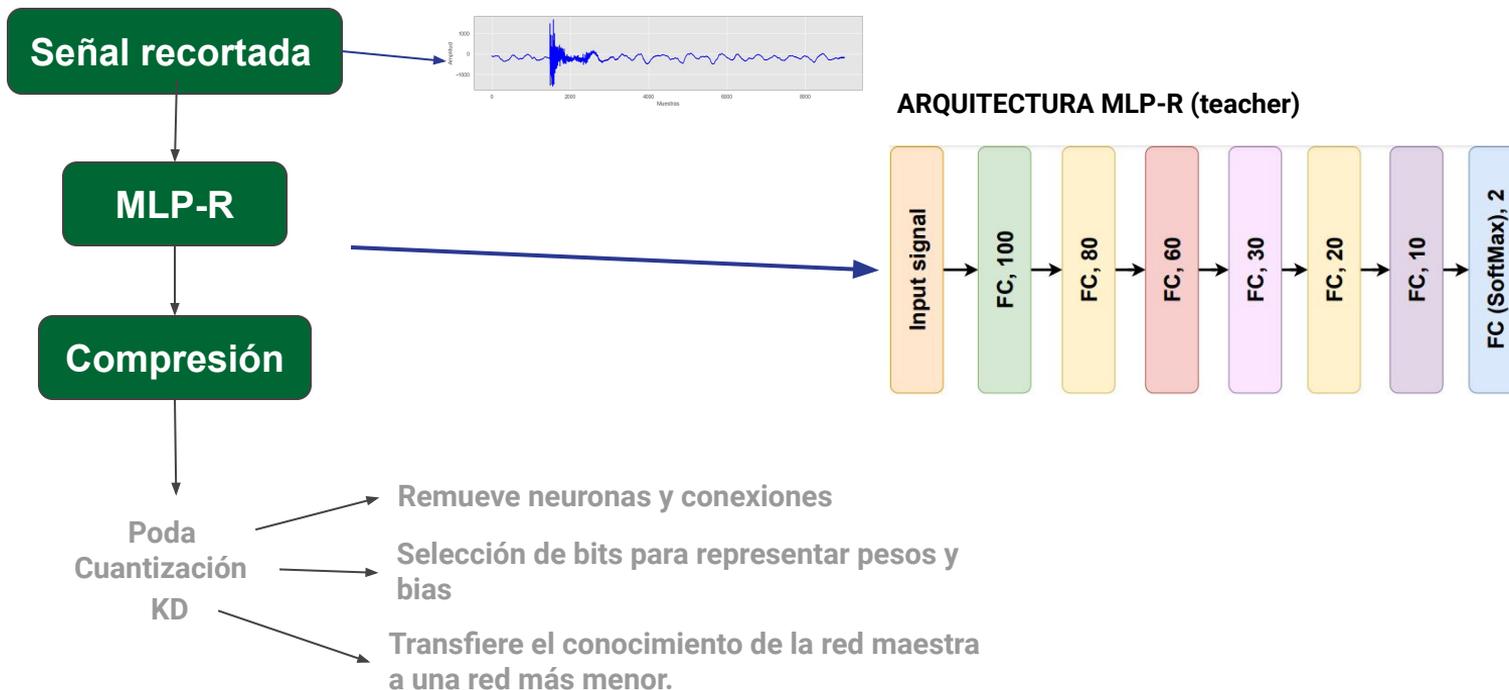
Entrenamiento con el dataset de señales



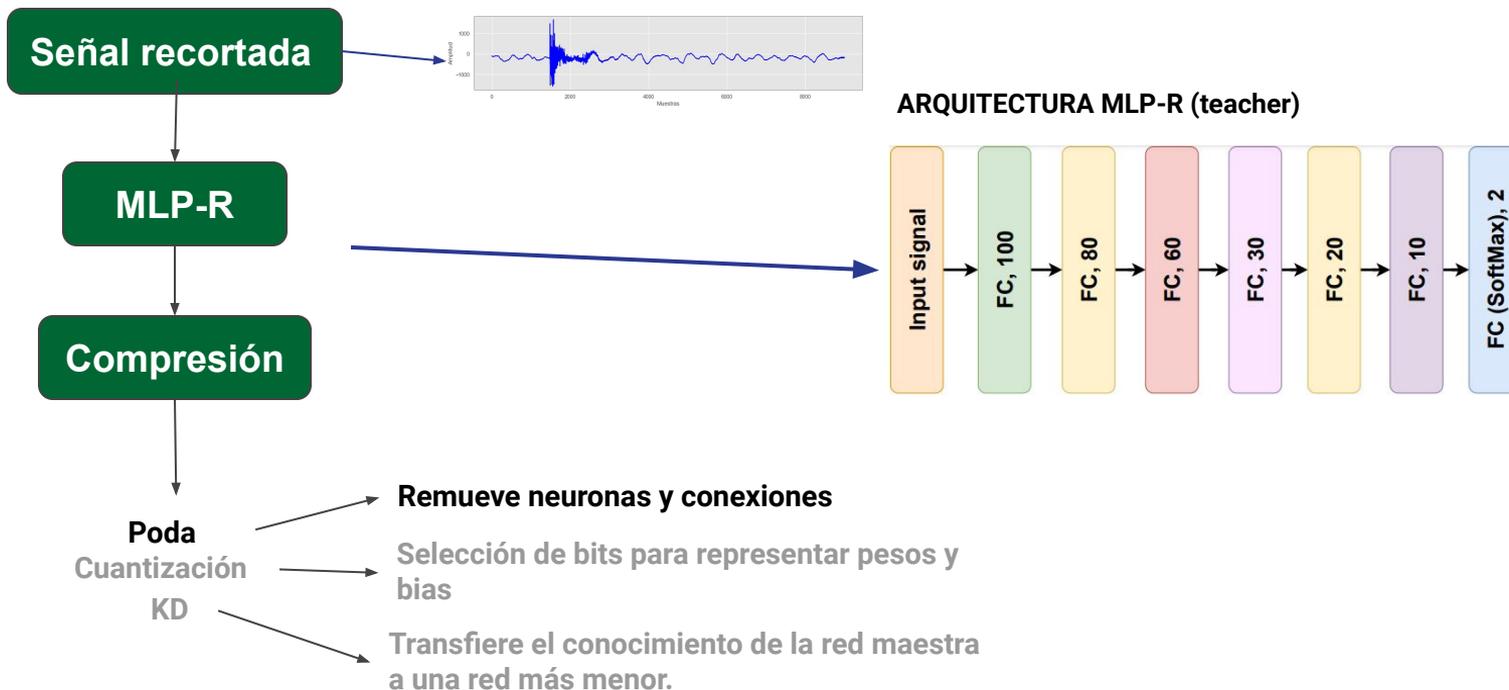
Entrenamiento con el dataset de señales



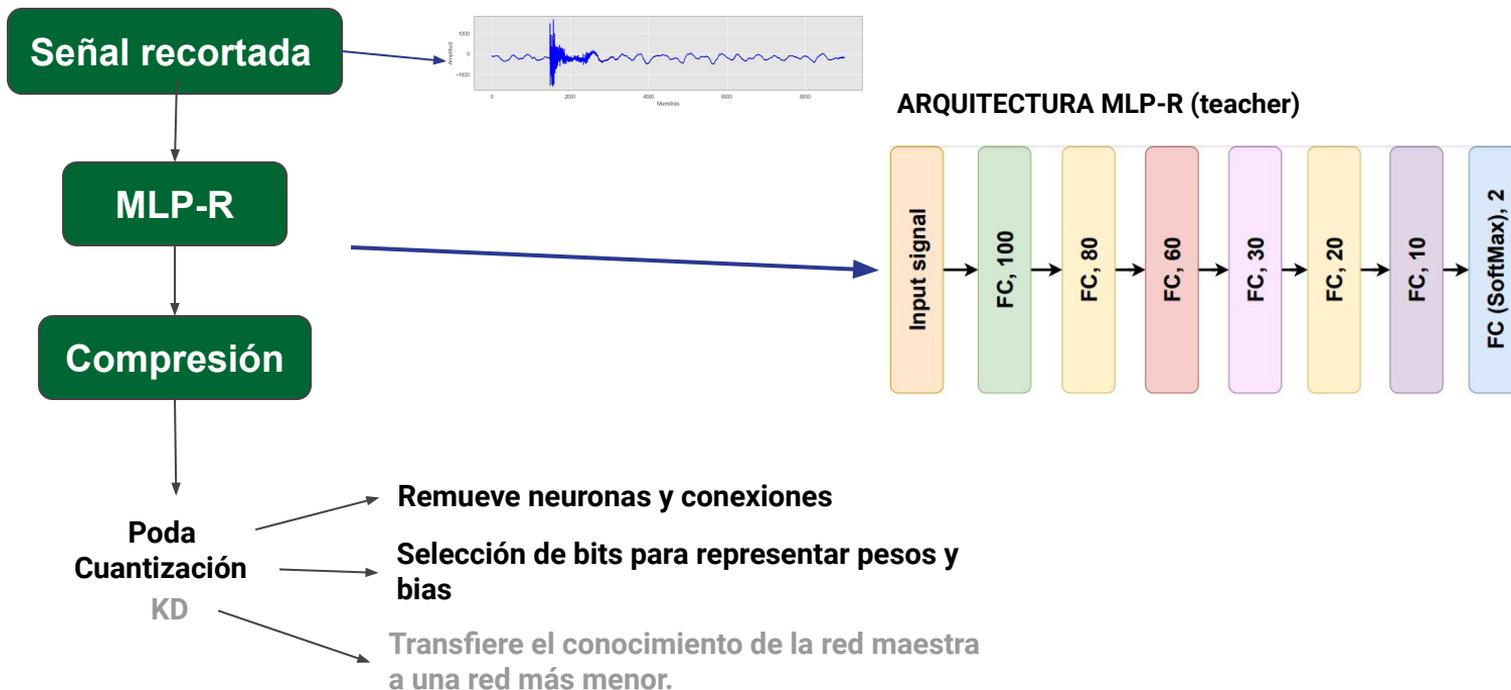
Entrenamiento con el dataset de señales



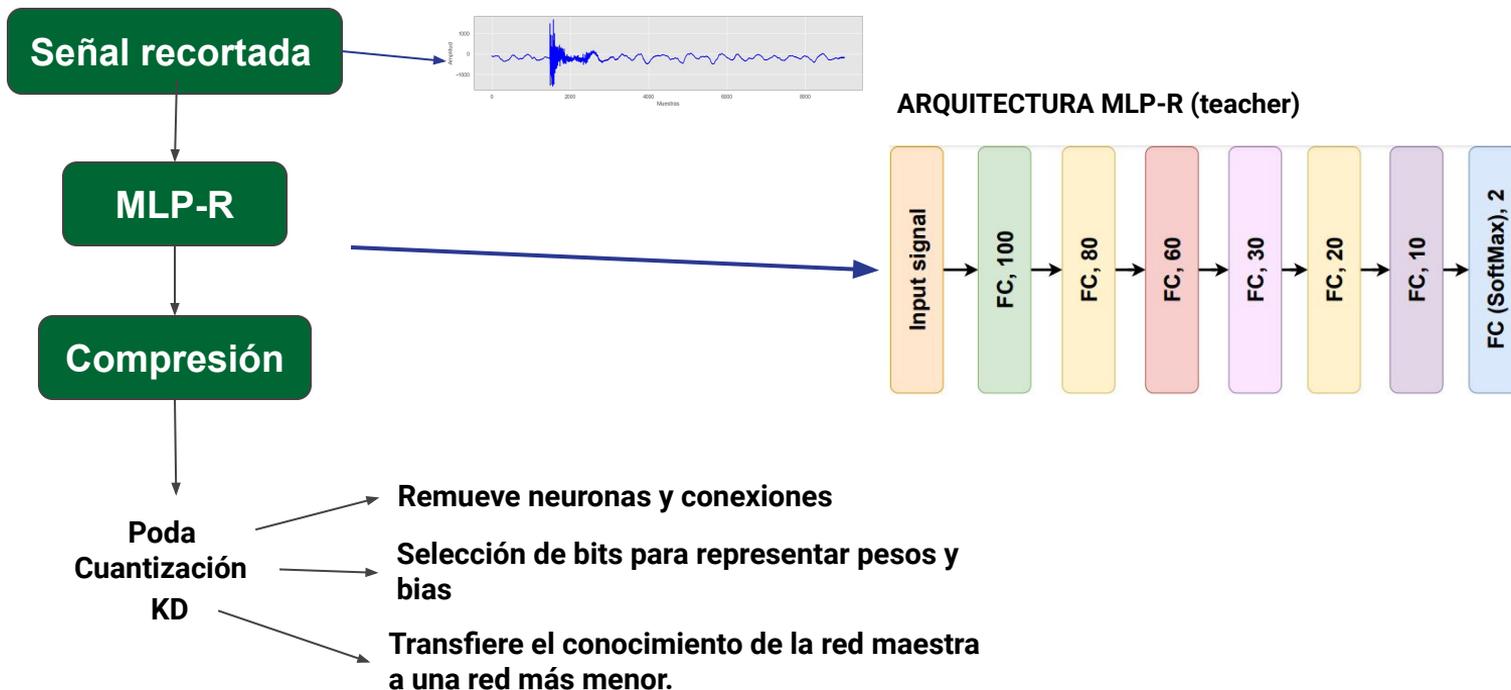
Entrenamiento con el dataset de señales



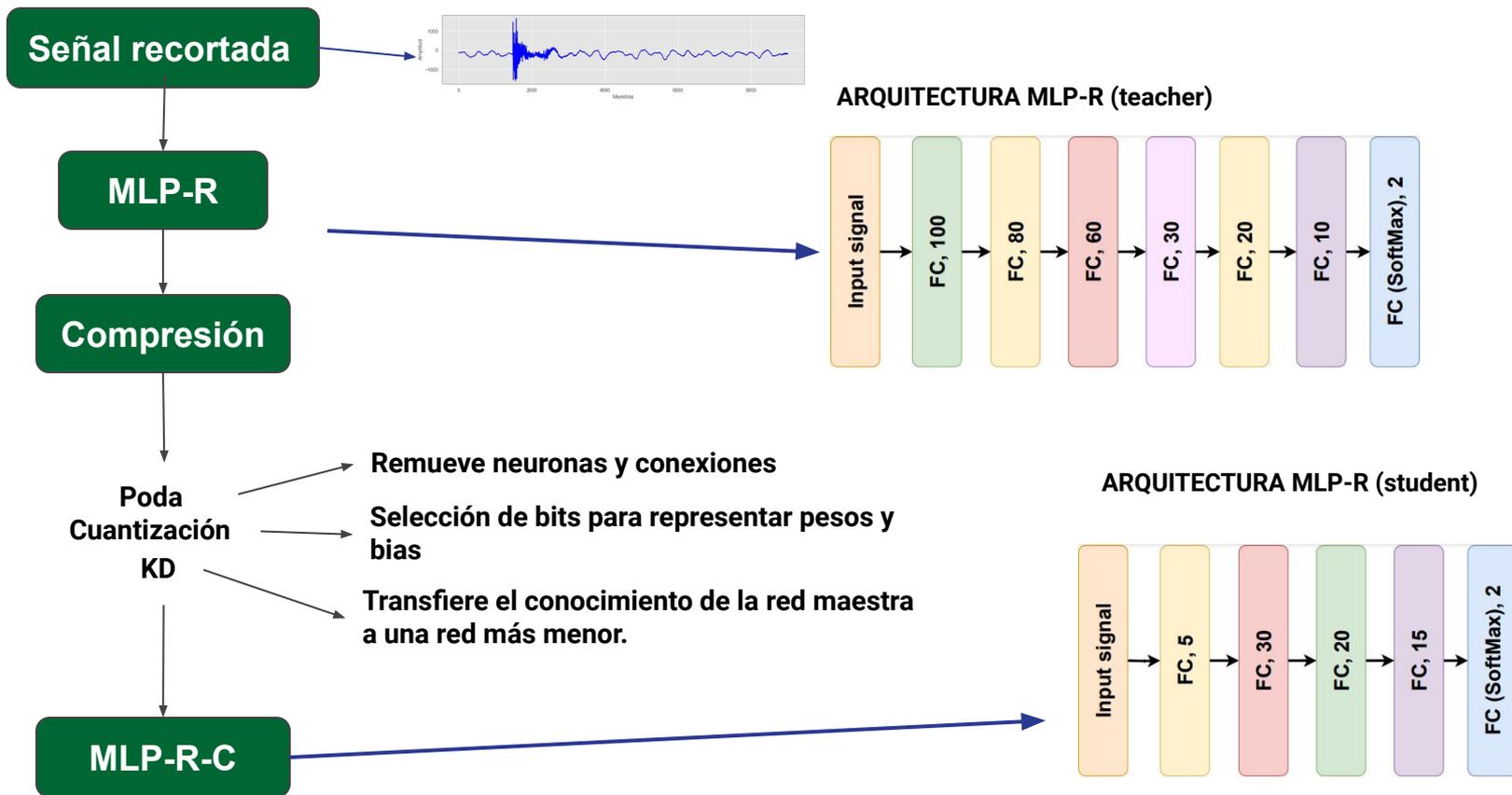
Entrenamiento con el dataset de señales



Entrenamiento con el dataset de señales



Entrenamiento con el dataset de señales



Resultados



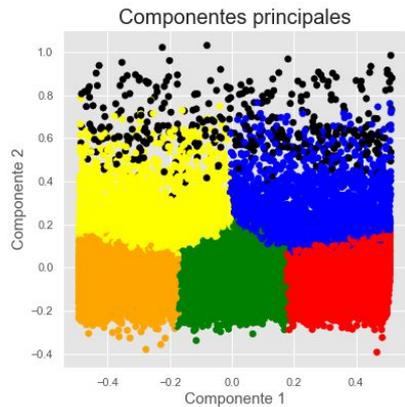
—

Tabla con las características extraídas de las trazas

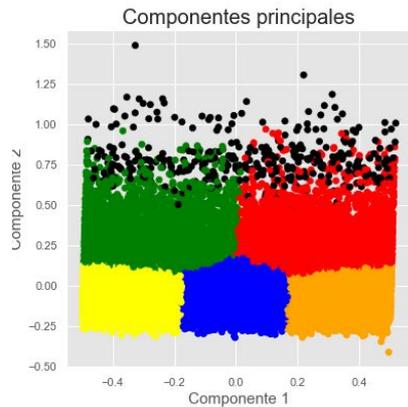
	Kurtosis	Media	Minimo	MaximoTiempo	PeakDiff	MaxMin	RMS	...Maximo_fr_CD4	Media_fr_CD3	Maximo_fr_CD3	Media_fr_CD2	Maximo_fr_CD2	Media_fr_CD1	Maximo_fr_CD1	
0	0.602749	0.005456	-92.764928	91.388954	78.54	184.153882	18.627213	...	1264.063116	507.348359	2078.333277	289.291074	2100.328044	14.917386	118.069276
1	0.671182	0.062621	-95.686359	96.171801	18.62	191.858159	18.631957	...	1047.262982	550.072670	2029.181955	285.269697	2324.703981	12.452816	148.777128
2	0.008840	0.067262	-54.400542	54.862518	49.11	109.263060	16.704589	...	1108.959895	177.388726	624.481978	96.336805	782.251224	4.574654	61.879301
3	-0.051290	0.033543	-73.066079	59.518276	89.20	132.584355	19.169342	...	982.243757	229.870111	883.083239	147.099312	1167.993476	6.826159	69.996192
4	2.199865	-0.075887	-209.206012	102.078327	75.05	311.284338	25.628951	...	1366.007521	706.471633	3002.187525	314.610605	2276.025574	14.584543	128.536873
...
29724	0.104210	0.148899	-106.369400	74.132933	84.73	180.502333	23.126344	...	1654.847604	364.428355	1834.509047	269.826151	2343.163673	10.671191	93.653018
29725	-0.184838	0.048908	-86.054537	85.461158	4.77	171.515695	24.267426	...	1082.517040	366.344770	1175.202339	239.488586	1933.400206	9.128568	127.115237
29726	0.274136	0.013879	-101.880877	108.784512	36.77	210.665389	25.072764	...	1285.080212	723.429570	2743.649271	367.919543	3528.789077	13.754575	124.128719
29727	0.251969	0.011711	-81.328203	84.355748	82.77	165.683952	21.281526	...	1331.622859	474.349280	2278.857652	398.969871	2538.683793	20.705821	277.709207
29728	0.351899	-0.173244	-104.953443	111.756575	18.85	216.710019	27.281867	...	1468.929909	444.762196	1953.079293	657.423055	4741.167705	40.014206	461.464902

Clusterizado con K-means

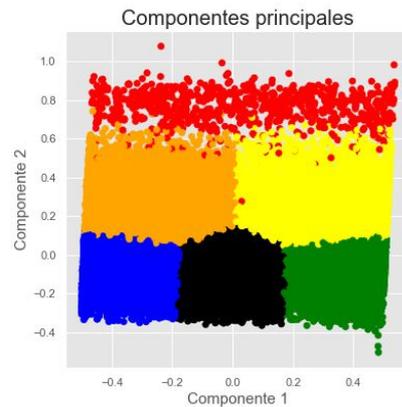
Diciembre



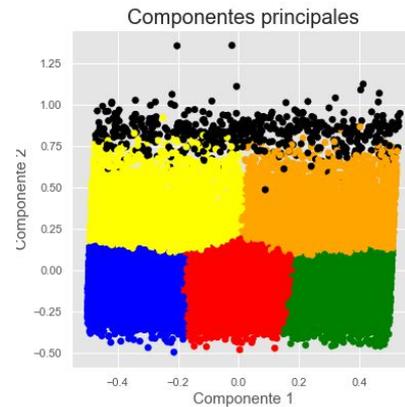
Enero



Febrero

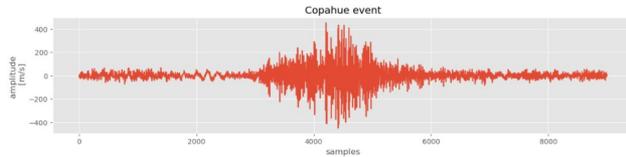
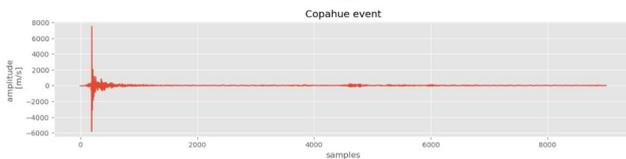
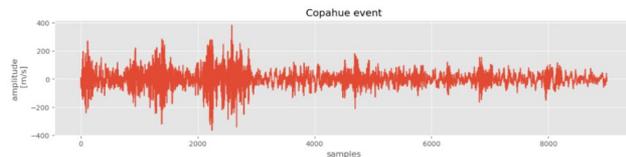
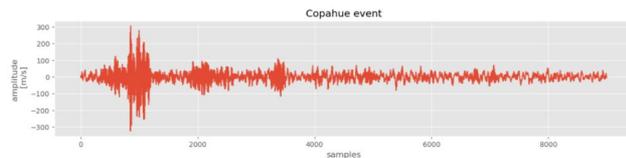


Marzo

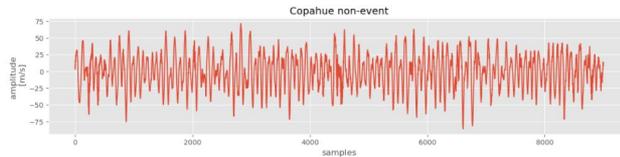
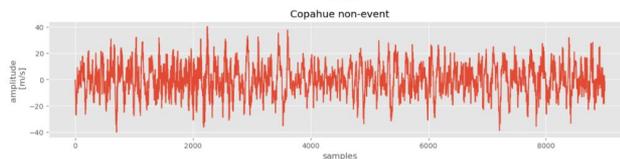
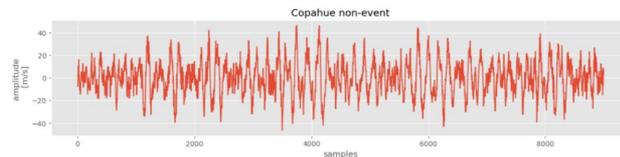
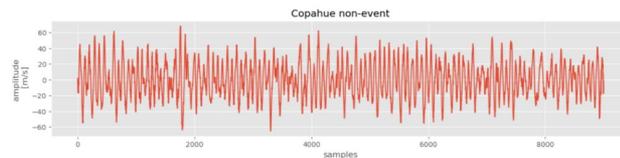


Señales evento y no evento

Evento (9,000 muestras)

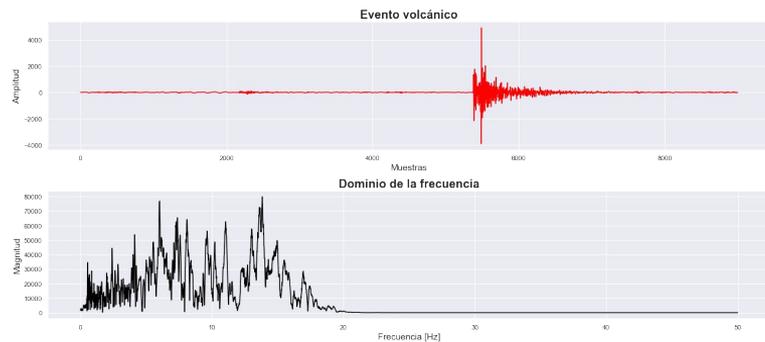


No evento (9,000 muestras)

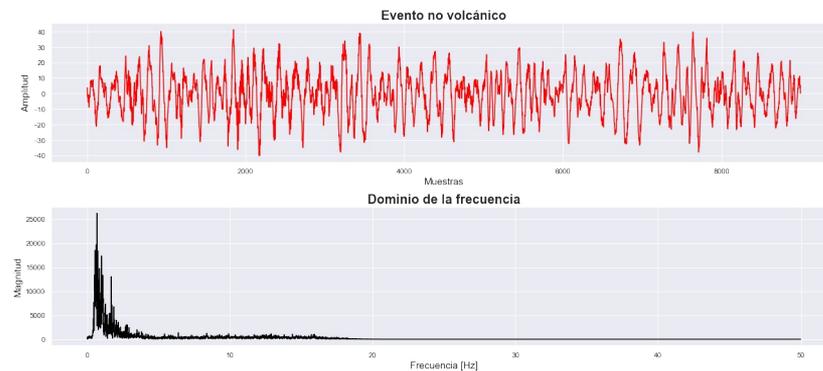


Señales evento y no evento

Evento (9,000 muestras)



No evento (9,000 muestras)



Conjunto de datos etiquetado

Resultados clusterización

Canal	Cluster	Diciembre 2017	Enero 2018	Febrero 2018	Marzo 2018
HIGI E	0	8076	6688	5921	6797
	1	8516	3998	5730	3790
	2	2252	6675	4362	6821
	3	380	6718	4348	6848
	4	2442	505	874	3542
	5	8063	4186	5617	535
HIGI N	0	7513	6486	5518	1648
	1	3202	595	5635	9225
	2	3211	6500	4310	3666
	3	709	4343	4282	9208
	4	7564	6486	5925	699
	5	7530	4360	1182	3887
HIGI Z	0	2250	7119	6511	7305
	1	8514	3214	6390	7222
	2	8077	7419	3474	3136
	3	8062	3309	778	7131
	4	2446	7361	3495	2998
	5	380	348	6204	541

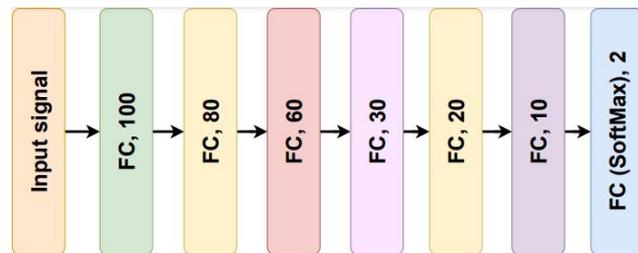
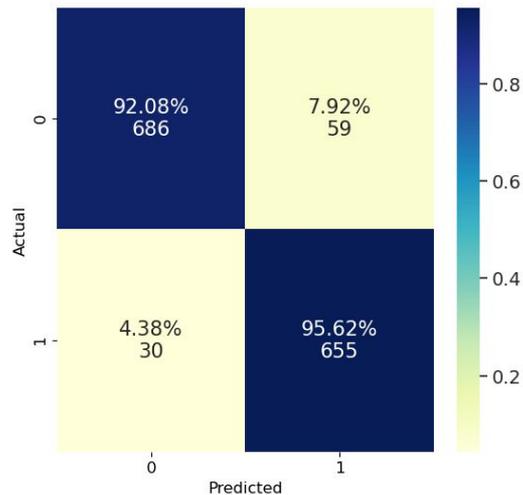
Eventos por mes en cada canal de HIGI

Canal	Diciembre 2017	Enero 2018	Febrero 2018	Marzo 2018	Total por canal
HIGI-E	380	505	874	535	2,294
HIGI-N	709	595	1,182	699	3,185
HIGI-Z	380	348	778	541	2,047
					Total: 7,526

No eventos por mes en cada canal de HIGI

Canal	Diciembre 2017	Enero 2018	Febrero 2018	Marzo 2018	Total por canal
HIGI-E	29,349	28,265	25,978	20,950	104,542
HIGI-N	29,020	28,175	25,670	27,634	110,499
HIGI-Z	21,287	28,422	26,074	27,792	103,575
					Total: 318,616

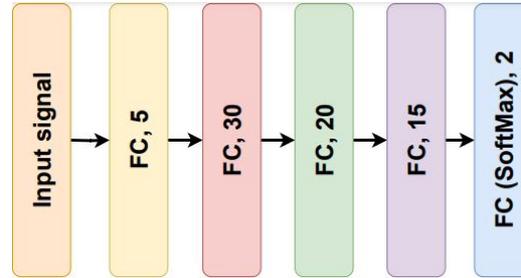
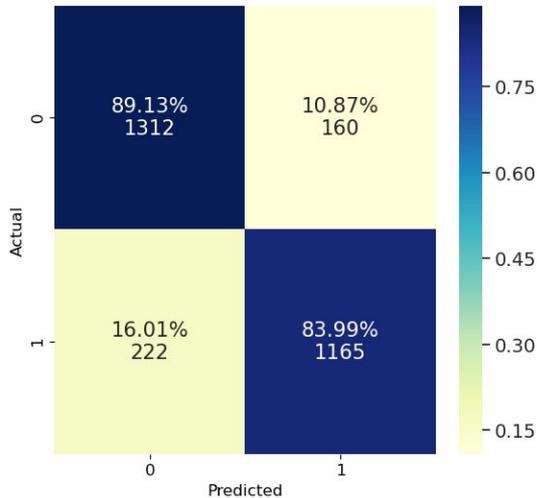
Arquitectura MLP-R



Métricas	Clases	
	0	1
Tasa de acierto	92.08%	95.62%
Precisión	95.81%	91.74%
Recall	92.08%	95.62%

Matriz de confusión para entrenamiento de la red MLP-R

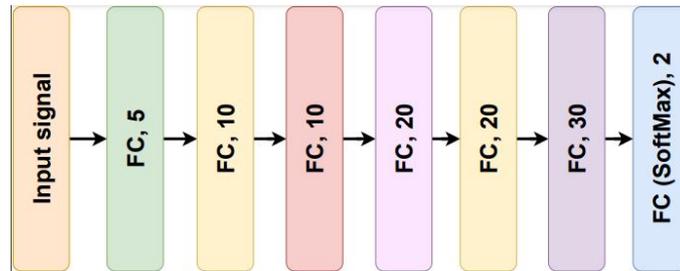
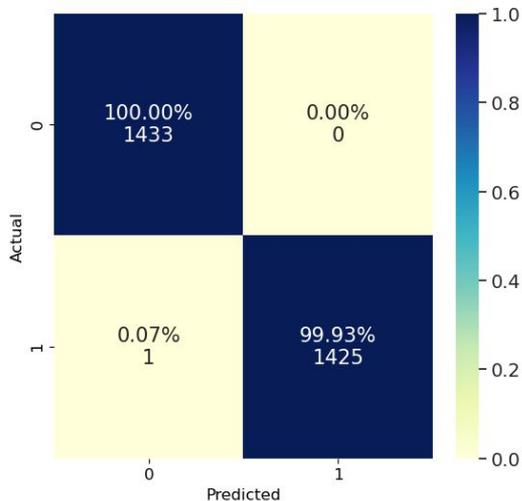
Arquitectura MLP-R-C



Métricas	Clases	
	0	1
Tasa de acierto	89.13%	83.99%
Precisión	85.53%	87.92 %
Recall	83.13%	83.99%

Matriz de confusión obtenida luego del proceso de compresión para generar la red MLP-R-C, considerando 24-bit punto fijo y 70 % de dispersión.

Arquitectura MLP-F



Métricas	Clases	
	0	1
Tasa de acierto	99.93%	99.93%
Precisión	99.86%	100 %
Recall	100 %	99.86%

Matriz de Confusión luego del entrenamiento de la arquitectura MLP-F

Resultados con dataset de prueba

Red	Señal	Resultados
MLP-R	Eventos	63.44 %
	No eventos	90.67 %
MLP-R-C	Eventos	64 %
	No eventos	86 %
MLP-F	Eventos	76 %
	No eventos	79 %

Integración hls4ml y HLS

Placa de desarrollo: PYNQ-Z1 @ 100 MHz

Implementación	BRAM	DSP	FF	LUT	Latencia
4-bits	0 %	2 %	4 %	60 %	0.29 μ s
8-bits	0 %	2 %	6 %	73 %	0.28 μ s
16-bits	0 %	2 %	10 %	88 %	0.31 μ s
24-bits	0 %	2 %	13 %	105 %	0.30 μ s

Conclusiones y Trabajos futuros



Conclusiones

El objetivo de este trabajo final fue generar una herramienta capaz de clasificar señales sísmicas provenientes del volcán Copahue y obtenidas mediante el sistema de adquisición de la red CP. Las técnicas de procesamiento digital de señales y aprendizaje automático permitieron, en primer lugar, lograr una clasificación automática de señales sísmicas volcánicas, a través de las características extraídas, separando eventos y no eventos volcánicos. De esta manera, la herramienta propuesta evita realizar el proceso de clasificación de manera manual, reduciendo los tiempos del procesamiento de los datos adquiridos por las estaciones.

Por otro lado, se entrenaron redes basadas en arquitecturas MLP para efectuar la clasificación binaria, tanto cuando se trabaja con las trazas crudas como con las características que las definen en los dominios temporal, frecuencial y de escala.

Considerando que la ubicación de los volcanes es muchas veces de difícil acceso, se efectuó la compresión de los modelos basados en MLP para su futura implementación en un sistema embebido. En este contexto, los métodos de cuantización, poda y KD se vuelven fundamentales para reducir la huella de memoria, las operaciones de cómputo y, en consecuencia, los tiempos de inferencia. Diversos tipos de datos fueron considerados variando la cantidad de bits para su representación, considerando una plataforma basada en FPGA, analizando el impacto en la utilización de recursos y latencia a través de HLS.

Conclusiones

El objetivo de este trabajo final fue **generar una herramienta capaz de clasificar señales sísmicas provenientes del volcán Copahue** y obtenidas mediante el sistema de adquisición de la red CP. **Las técnicas de procesamiento digital de señales y aprendizaje automático permitieron**, en primer lugar, **lograr una clasificación automática de señales sísmicas volcánicas**, a través de las características extraídas, separando eventos y no eventos volcánicos. De esta manera, **la herramienta propuesta evita realizar el proceso de clasificación de manera manual, reduciendo los tiempos del procesamiento de los datos adquiridos por las estaciones.**

Por otro lado, se entrenaron redes basadas en arquitecturas MLP para efectuar la clasificación binaria, tanto cuando se trabaja con las trazas crudas como con las características que las definen en los dominios temporal, frecuencial y de escala.

Considerando que la ubicación de los volcanes es muchas veces de difícil acceso, se efectuó la compresión de los modelos basados en MLP para su futura implementación en un sistema embebido. En este contexto, los métodos de cuantización, poda y KD se vuelven fundamentales para reducir la huella de memoria, las operaciones de cómputo y, en consecuencia, los tiempos de inferencia. Diversos tipos de datos fueron considerados variando la cantidad de bits para su representación, considerando una plataforma basada en FPGA, analizando el impacto en la utilización de recursos y latencia a través de HLS.

Conclusiones

El objetivo de este trabajo final fue **generar una herramienta capaz de clasificar señales sísmicas provenientes del volcán Copahue** y obtenidas mediante el sistema de adquisición de la red CP. **Las técnicas de procesamiento digital de señales y aprendizaje automático permitieron**, en primer lugar, **lograr una clasificación automática de señales sísmicas volcánicas**, a través de las características extraídas, separando eventos y no eventos volcánicos. De esta manera, **la herramienta propuesta evita realizar el proceso de clasificación de manera manual, reduciendo los tiempos del procesamiento de los datos adquiridos por las estaciones.**

Por otro lado, **se entrenaron redes basadas en arquitecturas MLP** para efectuar la clasificación binaria, tanto cuando **se trabaja con las trazas crudas como con las características que las definen en los dominios temporal, frecuencial y de escala.**

Considerando que la ubicación de los volcanes es muchas veces de difícil acceso, se efectuó la compresión de los modelos basados en MLP para su futura implementación en un sistema embebido. En este contexto, los métodos de cuantización, poda y KD se vuelven fundamentales para reducir la huella de memoria, las operaciones de cómputo y, en consecuencia, los tiempos de inferencia. Diversos tipos de datos fueron considerados variando la cantidad de bits para su representación, considerando una plataforma basada en FPGA, analizando el impacto en la utilización de recursos y latencia a través de HLS.

Conclusiones

El objetivo de este trabajo final fue **generar una herramienta capaz de clasificar señales sísmicas provenientes del volcán Copahue** y obtenidas mediante el sistema de adquisición de la red CP. **Las técnicas de procesamiento digital de señales y aprendizaje automático permitieron**, en primer lugar, **lograr una clasificación automática de señales sísmicas volcánicas**, a través de las características extraídas, separando eventos y no eventos volcánicos. De esta manera, **la herramienta propuesta evita realizar el proceso de clasificación de manera manual, reduciendo los tiempos del procesamiento de los datos adquiridos por las estaciones.**

Por otro lado, **se entrenaron redes basadas en arquitecturas MLP** para efectuar la clasificación binaria, tanto cuando **se trabaja con las trazas crudas como con las características que las definen en los dominios temporal, frecuencial y de escala.**

Considerando que la ubicación de los volcanes es muchas veces de difícil acceso, **se efectuó la compresión de los modelos basados en MLP para su futura implementación en un sistema embebido.** En este contexto, **los métodos de cuantización, poda y KD se vuelven fundamentales** para reducir la huella de memoria, las operaciones de cómputo y, en consecuencia, los tiempos de inferencia. Diversos tipos de datos fueron considerados variando la cantidad de bits para su representación, considerando una plataforma basada en FPGA, **analizando el impacto en la utilización de recursos y latencia a través de HLS.**

Trabajos futuros

- Aplicar ventanas corredizas a las señales de HIGI, para aumentar el dataset de entrenamiento y entrenar con diferentes posiciones de los eventos/no eventos.
- Entrenar otro tipo de arquitecturas, tales como RNN, CNN-1Dconv.
- Integración en el framework PYNQ.
- Identificar entre los tipos de evento volcanotectónicos (VT), de largo periodo (LP), híbridos.
- Extender el análisis a todas las estaciones, con sus correspondientes canales.
- Recortar las trazas a 6,000 muestras (1 minuto).
- Aplicar filtrados en bandas específicas para encontrar eventos VT, LP o tremor.
- Entrenar la red a partir de las imágenes de los eventos.

Agradecimientos



Este trabajo final se enmarca en una colaboración interdisciplinaria entre los Departamentos de Electrónica y Física de la Universidad Nacional de San Luis, en conjunto con Multidisciplinary Laboratory (MLab)/ STI Unit perteneciente a The International Centre for Theoretical Physics (ICTP).

Además, este trabajo se realizó utilizando datos de la Red del Volcán Copahue; instalada y mantenida por el Laboratorio de estudios y seguimientos de volcanes activos (LESVA) de la Universidad Nacional de Río Negro, bajo la dirección de Alberto Caselli.

A Silvana Spagnotto por la disposición y el tiempo. A Iván Melchor por sus conocimientos en el tema. A Alejandro Nuñez por la co-dirección y el nexo fundamental con Silvana.

A Romina Molina, por cada palabra, por la paciencia y la buena energía en cada momento.

A mi familia y amigos.

Workshop "Seismology and Artificial Intelligence"

desde el 13 al 15 de Septiembre de 2023 en Frankfurt.

10:50 - 11:30

Copahue volcano seismic event detection based on digital signal processing and machine learning techniques: towards an on-the-edge implementation

Romina Molina
ICTP, Italy



Copahue volcano is an active stratovolcano located on the border between Argentina and Chile, is part of the volcanic belt of the Andes, and is one of the most active volcanoes for both Chilean and Argentinian government institutions. Owing to the complexity of seismic events, the latest advancements in Machine Learning (ML) offer powerful solutions for handling large datasets and effectively extracting desired features.

Furthermore, in situations where the number of seismic stations is sparse, either because of low financial resources or difficult access, an on-the-edge detector can improve the monitoring because it does not depend on the records of the other stations.

In this presentation, Copahue volcano-seismic event detection based on digital signal processing techniques and ML is presented, which can be considered as an event-triggering system based on ML to operate on the edge.

Preguntas?

