



UNIVERSIDAD DE LA REPÚBLICA
FACULTAD DE INGENIERÍA



Monitoreo no intrusivo de cargas eléctricas (NILM)

MEMORIA DE PROYECTO PRESENTADA A LA FACULTAD DE
INGENIERÍA DE LA UNIVERSIDAD DE LA REPÚBLICA POR

Franco Marchesoni, Camilo Mariño, Elías Masquil

EN CUMPLIMIENTO PARCIAL DE LOS REQUERIMIENTOS
PARA LA OBTENCIÓN DEL TÍTULO DE
INGENIERO ELÉCTRICO.

TUTORES

Alicia Fernández Universidad de la Repùblica
Pablo Massaferro Universidad de la Repùblica

TRIBUNAL

Matías Di Martino Universidad de la Repùblica
Álvaro Gómez Universidad de la Repùblica
Leonardo Steinfeld Universidad de la Repùblica
Juan Pablo Maggi UTE

Montevideo
lunes 1 junio, 2020

Monitoreo no intrusivo de cargas eléctricas (NILM), Franco Marchesoni, Camilo Mariño, Elías Masquil.

Esta tesis fue preparada en L^AT_EX usando la clase iietesis (v1.1).

Contiene un total de 187 páginas.

Compilada el lunes 1 junio, 2020.

<http://iie.fing.edu.uy/>

La incertidumbre es la magia

GHIGGIA

Agradecimientos

Muchas personas nos han acompañado a lo largo de este Proyecto. Agradecemos a nuestros tutores, Alicia Fernández y Pablo Masaferro (y su familia, que permitió que invadiéramos su hogar). A nuestra contraparte en UTE, formada por el Ing. Juan Pablo Maggi y todo su equipo, dentro del cual se encuentran Luis Pérez y Fernando Santomauro. Nuestra contraparte nos brindó valioso asesoramiento técnico, además de los indispensables equipos y el apoyo económico. A nuestro colega y amigo Ignacio Hounie, que aportó infraestructura vital para que se realizara el Proyecto. A Roberto Rodríguez y Sergio Beheregaray, que nos ayudaron en la creación del medidor desde el taller del Instituto de Ingeniería Eléctrica (IIE). Al Laboratorio de Energía Solar (LES), por brindarnos infraestructura computacional, almacenamiento y poder de cómputo. A Ignacio Ramírez, que colaboró en la creación del PCB. A Nicolás Pérez, que respondió múltiples consultas cruciales. A Martha Delgado, que hizo posible que llegaran los productos del exterior y siempre estuvo disponible. A los integrantes del Taller Fourier, que colaboraron indirectamente con el Proyecto. A Andrés Cardozo, Martín Avas, Santiago Machado, Gabriel Gómez, Pablo Musé y José Lezama, que atendieron consultas en sus respectivas áreas de expertiz. A Ayelen Iglesias, corregidora de estilo. Agradecemos por último a quienes permitieron que empezáramos el proyecto: a nuestra familia, por obvios motivos. Al Uruguay y su UdeLaR, por tener puertas tan abiertas.

Por otro lado agradecemos el generoso apoyo del Centro de Innovación en Ingeniería (CII), que colaboró económicamente con el Proyecto mediante la Convocatoria a Proyectos de Fin de Carrera. También se agradece al Fondo de Fomento de las Ingenierías, financiado por la Agencia Nacional de Investigación e Innovación, que colaboró con la asignación de una Beca de Egreso para estudiantes de la Facultad de Ingeniería a Franco Marchesoni, uno de los autores del Proyecto.

A nuestras familias, amistades y todos los que nos ayudaron a llegar hasta acá

Resumen

En un contexto donde la sostenibilidad del planeta es una preocupación global, la eficiencia energética toma vital importancia y abarca diversas aristas, dependiendo varias de ellas del uso y generación de información de utilidad.

Non Intrusive Load Monitoring (NILM), o “monitoreo no intrusivo de cargas eléctricas”, es el nombre de una aplicación específica del procesamiento de señales. A partir de la medida de corriente y voltaje consumidas por un conjunto de cargas, se usan técnicas para asignar a cada carga individual una potencia consumida. En este sentido se presenta, en el marco de este proyecto, una revisión del estado del arte sobre los algoritmos y dispositivos para NILM, un prototipo del aparato medidor que toma muestras a alta frecuencia, la implementación de algoritmos de clasificación de electrodomésticos y también de algoritmos de desagregación, entrenados sobre bases de datos públicas. En la clasificación de electrodomésticos, se logran resultados comparables a los del estado del arte y se realiza un estudio minucioso de las distintas características que pueden extraerse de su señal de potencia eléctrica.

En desagregación, se valida la hipótesis de que incorporar datos de alta frecuencia (medidas a más de 7 kHz) permite un mejor ajuste a los datos y se presentan adaptaciones de los algoritmos de referencia en esta línea, que logran en algunos casos mejores resultados que las implementaciones de referencia. Este enfoque de aprendizaje supervisado se desempeña bien al entrenar los modelos sobre los electrodomésticos a monitorear.

Además, se releva una base de datos de validación con medidas en Uruguay recolectada con el prototipo propuesto. Esta base de datos cuenta con medidas individuales de distintos electrodomésticos y medidas de alta frecuencia de la potencia agregada de la casa.

Tabla de contenidos

Agradecimientos	5
Resumen	9
1. Introducción	1
1.1. Motivación	1
1.2. Definición del problema	1
1.3. Contexto del Proyecto	2
1.4. Contribuciones del Proyecto	2
1.5. Organización de contenidos	2
2. Revisión Bibliográfica	5
2.1. Resumen	5
2.2. Introducción	5
2.2.1. Electrodomésticos	6
2.3. Adquisición	7
2.4. Visión general del problema	7
2.5. Algoritmos	8
2.6. Evaluación de resultados y métricas asociadas	9
2.7. Productos comerciales	9
3. Medidor no intrusivo	11
3.1. Resumen	11
3.1.1. Sistema de medida	12
3.1.2. Requerimientos del medidor	12
3.2. Componentes	13
3.2.1. Conversor Analógico Digital (ADC)	14
3.2.2. Circuito analógico acondicionador	15
3.3. Circuito acondicionador	15
3.3.1. Señales de entrada	15
3.3.2. Esquema general del circuito	16
3.3.3. Diseño del circuito	18
3.3.4. Implementación del circuito	21
3.3.5. Otros circuitos evaluados	22
3.4. <i>Software</i> del medidor	22
3.5. Calibración	23
3.6. Conclusiones	23

Tabla de contenidos

4. Sistema de recolección de datos	25
4.1. Resumen	25
4.2. Red de medidores intrusivos	25
4.3. Integración de los sistemas de medida	26
4.3.1. Servidor remoto	28
4.4. Base de datos NILM-UY	28
4.4.1. Motivación	28
4.4.2. Sistema de medida	28
4.4.3. Descripción de los datos	29
4.4.4. Formato de los datos y estructura de carpetas	29
4.4.5. Protocolo de medida	30
5. Ingeniería de características y clasificación de electrodomésticos sobre PLAID	31
5.1. Introducción	31
5.1.1. Resumen	31
5.1.2. Motivación	31
5.1.3. Trabajo relacionado	32
5.1.4. PLAID	34
5.2. Procedimiento	34
5.2.1. Selección de características	34
5.2.2. Información mutua	36
5.2.3. Características sobre series agregadas	36
5.2.4. Extracción del transitorio	37
5.3. Resultados y conclusiones	38
5.3.1. Resultados	38
5.3.2. Características a ser utilizadas para desagregación	42
5.3.3. Conclusión	45
6. Base de datos UK-DALE	47
6.1. Resumen	47
6.2. Elección de la base de datos	47
6.3. Descripción de los datos de baja frecuencia	47
6.3.1. Potencia aparente	48
6.3.2. Rango temporal de las medidas	48
6.3.3. Electrodomésticos medidos	49
6.4. Procesamiento de los datos de baja frecuencia	49
6.4.1. Alineamiento temporal	49
6.4.2. Datos faltantes	50
6.4.3. Activaciones y no activaciones	50
6.4.4. Obtención de datos de entrada y salida	50
6.5. Descripción de los datos de alta frecuencia	53
6.5.1. Rango temporal de las medidas	54
6.6. Procesamiento de los datos de alta frecuencia	54
6.6.1. Descriptores de alta frecuencia	54
6.6.2. Descarga de los datos	55
6.6.3. Procesamiento	56
6.7. Conclusión	59

Tabla de contenidos

7. Desagregación con redes neuronales	61
7.1. Resumen	61
7.2. Introducción a las Redes Neuronales Artificiales	61
7.2.1. Tipos de redes utilizadas	62
7.3. Conjuntos de datos	64
7.3.1. Datos sintéticos	64
7.3.2. Entrenamiento, validación y test	65
7.3.3. Comentario sobre el conjunto test visto	66
7.3.4. Preprocesamiento de entrada a las redes	67
7.4. Redes de desagregación	67
7.4.1. Arquitecturas de referencia	67
7.4.2. Variantes propuestas	69
7.5. Conclusión	72
8. Selección de modelos	73
8.1. Resumen	73
8.2. Métricas de desempeño	73
8.3. Modos de evaluación	75
8.4. Entrenamiento	76
8.4.1. Selección de la grilla	76
8.4.2. Número de épocas	77
8.4.3. Variantes que afectan el entrenamiento	78
8.4.4. Mejores modelos	78
8.5. Evaluación del mejor modelo sobre datos de validación	79
8.5.1. Mejor modelo para cada electrodoméstico	79
8.5.2. Desempeño en desagregación	79
8.5.3. Desempeño en detección o clasificación	80
9. Evaluación sobre UK-DALE	83
9.1. Resumen	83
9.2. Metodología	83
9.3. Comparación sobre test visto usando activaciones	84
9.4. Comparación para test no visto usando activaciones	87
9.5. Influencia del mecanismo de evaluación	88
9.6. Comparabilidad de implementaciones	91
9.7. Funcionamiento del sistema NILM sobre casas desconocidas del Reino Unido	91
9.8. Conclusiones	93
10. Evaluación sobre datos de Uruguay	95
10.1. Resumen	95
10.2. Procesamiento de los datos para la evaluación	95
10.3. Disponibilidad de datos	96
10.4. Evaluación	96
10.4.1. Casa 1 (Franco)	96
10.4.2. Casa 2 (Pablo)	98
10.5. Conclusiones	101
11. Conclusiones y perspectivas	103
11.1. Resumen	103
11.2. Conclusiones	103
11.3. Perspectivas y trabajo a futuro	105
11.4. Conclusiones de la asignatura	106

Tabla de contenidos

Apéndices	107
A. Circuitos de referencia	107
B. Pinza de corriente	109
B.1. Resumen	109
B.2. Introducción	109
B.3. Metodología	110
B.4. Medidas	111
B.5. Impedancia del banco de carga	112
B.6. Conclusiones	115
C. Hojas de datos	119
D. Archivo de Configuración Wget	123
E. Resultados de los algoritmos	125
E.1. Resumen	125
E.2. Resultados de validación sobre UK-DALE	125
E.2.1. AUC de validación para todos los modelos y experimentos propuestos	125
E.3. Resultados de test sobre activaciones en UK-DALE	125
E.3.1. Casas vistas durante el entrenamiento	125
E.3.2. Casas no vistas	131
E.4. Resultados de test sobre ventanas deslizantes en UK-DALE	137
E.4.1. Casas vistas durante el entrenamiento	137
E.4.2. Casas no vistas durante el entrenamiento	142
E.5. Resultado de test en ventanas deslizantes sobre NILM-UY	147
F. Comparación contra resultados de referencia	157
F.1. Resumen	157
F.2. Comparación sobre métricas de desagregación	157
G. Enfoques a futuro	161
G.1. Acerca de los algoritmos de AI	161
G.1.1. <i>“The bitter lesson”</i>	161
G.1.2. Falencias del aprendizaje supervisado	162
G.1.3. Distintos enfoques	162
G.2. Algoritmo no-supervisado	163
G.2.1. Solución genérica	163
G.2.2. Una implementación fallida	165
G.3. Black Box	166
G.3.1. Bloques o subcomponentes	167
Referencias	168
Índice de temas	172

Capítulo 1

Introducción

1.1. Motivación

Según el último reporte del Panel Intergubernamental del Cambio Climático (IPCC) el cambio climático presenta desafíos importantes que requieren de un ambicioso accionar. La mejora de la eficiencia energética es un camino que ya está siendo recorrido pero se requiere “*a greater scale and pace of change*” para limitar el calentamiento a 1,5°C. La eficiencia energética es una preocupación tanto por parte de las empresas distribuidoras como por parte de los consumidores. Por parte de los consumidores existe la motivación del ahorro económico, además de la responsabilidad ambiental que algunos poseen. Por parte de la empresa distribuidora, el despacho óptimo, el control de los picos de demanda, el análisis de los patrones de consumo, y la posibilidad de avanzar hacia el concepto de *smart grid* son factores de gran importancia.

El problema de monitoreo no intrusivo de cargas eléctricas (NILM) está relacionado con una necesidad que ha de satisfacerse antes de avanzar en los aspectos antes nombrados: conocer cuánto, cuándo y cómo un usuario (un nodo en la red) consume energía eléctrica. Estas mediciones han de realizarse a un costo razonable, por lo que un método que permita evitar medir las cargas individualmente es valioso. De eso se trata NILM, de inferir la información de consumo de múltiples cargas a partir de la medida en un solo punto, donde las señales se encuentran agregadas.

1.2. Definición del problema

Dentro del área de NILM, este Proyecto de Fin de Carrera se ocupa del caso particular en que las cargas son electrodomésticos y los usuarios son los habitantes de un hogar con suministro monofásico. El problema consiste entonces en identificar el consumo de algunos electrodomésticos individuales a partir de las mediciones realizadas por un sólo dispositivo. El Proyecto también incluye la evaluación de desempeño de la solución realizada.

El alcance del Proyecto es el de una solución punta a punta. Esta solución implica el desarrollo de un prototipo de medidor y la implementación de algoritmos de desagregación. El desempeño de la solución también será evaluado en hogares del territorio uruguayo.

Capítulo 1. Introducción

1.3. Contexto del Proyecto

Si bien se contiene en sí mismo, el Proyecto se enmarca en un proyecto de investigación que realiza el Departamento de Procesamiento de Señales del Instituto de Ingeniería Eléctrica. Este proyecto de investigación tiene como contraparte a la empresa nacional de distribución de energía eléctrica de Uruguay: La Administración Nacional de Usinas y Trasmisiones Eléctricas (UTE). El proyecto de investigación tiene como objetivo identificar las potencialidades de sistemas de monitoreo no intrusivo, y este Proyecto de Fin de Carrera es uno de las primeras aristas de trabajo desarrolladas en el marco del mismo.

1.4. Contribuciones del Proyecto

Este Proyecto aporta una serie de productos que se espera sean útiles para el cliente (UTE) y la comunidad académica. En particular, se logró:

- La creación de la base de datos NILM-UY (Capítulo 4)
- El estudio y la generación de un algoritmo identificador de electrodomésticos a partir de medidas individuales (Capítulo 5)
- La validación de los algoritmos presentados en *Neural NILM: Deep Neural Networks Applied to Energy Disaggregation* [19] cuyos autores son Jack Kelly y William Knottenbelt
- La propuesta de nuevos métodos que apuntan a incorporar la información de alta frecuencia (Capítulo 7).
- La evaluación del error de generalización a hogares distintos a los presentes en los datos de entrenamiento, pertenecientes tanto al país de los datos de entrenamiento (UK) como a Uruguay (Capítulos 9 y 10).
- Implementación punta a punta de un sistema de NILM y perspectivas para futuras versiones

1.5. Organización de contenidos

El presente documento se divide en tres partes tras la revisión bibliográfica. La primera parte refiere, a grandes rasgos, a la parte de *hardware* y a la base de datos NILM-UY. La segunda parte refiere a los algoritmos, o al *software*. La tercera y última parte expone los resultados obtenidos, las conclusiones y las perspectivas.

Dentro de la primera parte se encuentra el proceso de diseño y construcción de la electrónica utilizada por el medidor de corriente y voltaje de alta frecuencia (14kHz). El circuito en cuestión se encarga de adaptar las señales de voltaje y de la pinza de corriente a la entrada de un convertidor análogo digital (ADC), implementado como una tarjeta de sonido, como se describe en el Capítulo 3. Por otro lado, se analiza el medidor de corriente a utilizar. En particular, se releva su respuesta en frecuencia, para luego definir la frecuencia de muestreo del medidor. La respuesta en frecuencia de la pinza de corriente es estudiada en el Apéndice B. En el Capítulo 4, se presenta el sistema completo de monitoreo intrusivo, que fue utilizado para relevar la base de datos NILM-UY.

En la segunda parte se presentan los algoritmos de desagregación utilizados para atacar el problema. Partiendo de la descripción de *The UK-DALE dataset* [18], base de datos que se utiliza para entrenar los modelos, se explican los algoritmos basados en [19], la implementación de los mismos y las mejoras o variantes propuestas por

1.5. Organización de contenidos

los autores. Entre las posibles mejoras propuestas, se incluye la exploración del uso de datos de alta frecuencia, para lo cual se dedica un capítulo en el que se exploran los datos de PLAID [11]. Finalmente, esta parte concluye con la presentación de las métricas para evaluar los modelos y los criterios para elegir los modelos a utilizar de forma definitiva.

La parte tres es la última del documento. La misma presenta los resultados finales en los Capítulos 9 y 10. Estos resultados son los que permiten evaluar el poder de generalización de los algoritmos aprendidos. El Capítulo 11 presenta las conclusiones y perspectivas del Proyecto, dando también una opinión de cuáles son los enfoques que permitirían llegar a un producto escalable de NILM con buen desempeño.

Capítulo 1. Introducción

Capítulo 2

Revisión Bibliográfica

2.1. Resumen

En este capítulo se presenta una revisión del estado del arte, basada en [36] y en los múltiples artículos que fueron consultados a lo largo del Proyecto. Se comienza con el estado del arte en adquisición y las potencialidades de cada solución. Luego se explican los pasos correspondientes al enfoque clásico: detección de flancos y extracción de características. Los distintos algoritmos se presentan al final, diferenciando aquellos que se basan en redes neuronales.

2.2. Introducción

Non Intrusive Load Monitoring (NILM) o el monitoreo no intrusivo de cargas eléctricas, tiene como objetivo detectar estados en electrodomésticos o máquinas (encendido, apagado) y asignar cuánta potencia consume cada uno, a partir de medidas en un solo punto de la instalación eléctrica. Uno de los trabajos fundacionales de NILM [13] ya planteaba las potenciales aplicaciones: en seguridad, detección de fallas, y caracterización de la demanda. El modelado de la red eléctrica de una casa es el de cargas en paralelo. Esto implica que, por el teorema de Tellegen, la potencia que se consume es aditiva.

Una formulación matemática posible es la siguiente: generando un vector binario a , que representa encendido o apagado de cada electrodoméstico con un 1 o un 0 respectivamente, se cumple que la potencia en tiempo t es

$$P(t) = \sum_{i=1}^n a_i(t) P_i + e(t) \quad (2.1)$$

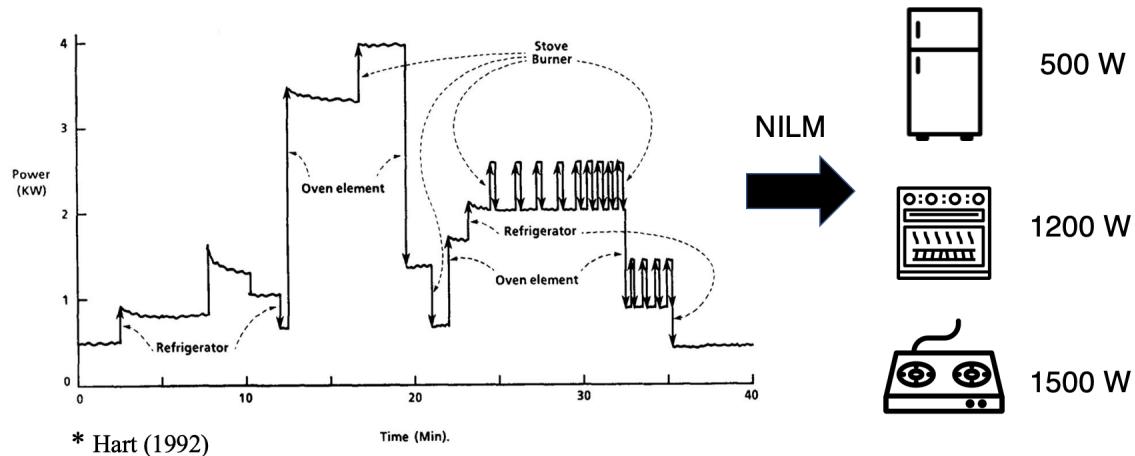
donde $e(t)$ es un error que representa ruido y P_i es la potencia que consume el electrodoméstico en régimen. En esta definición se asume que P_i está bien definida, es decir, no se están considerando electrodomésticos con distintos estados (o se considera cada estado como un P_i distinto).

La solución de la asignación se da, para cada instante t , encontrando aquel a que minimice $|e(t)|$. Esto es un problema NP-completo [13]. No hay soluciones exactas para las dimensiones naturales del problema. La formulación anterior deja implícita la necesidad de conocer P_i , que se infiere a partir de la serie temporal. También hay implícito un prior de continuidad en el tiempo de cada función binaria $a_i(t)$.

Capítulo 2. Revisión Bibliográfica

La formulación matemática deja por fuera otros problemas de ingeniería, como la adquisición de los datos, el procesamiento de los mismos y la creación y evaluación de los algoritmos.

El enfoque clásico de NILM consta de la adquisición de datos, la detección de flancos en la potencia activa, la extracción de características a partir de las curvas de tensión y corriente consumida y por último la identificación de electrodomésticos. Estas etapas serán comentadas por separado en lo que resta del capítulo.



2.2.1. Electrodomésticos

Los electrodomésticos que componen la señal agregada con su consumo de potencia pueden ser clasificados. Una clasificación usualmente utilizada consiste en cuatro grupos:

- Siempre encendidos (ej. router)
- ON/OFF (ej. calefón)
- Máquina de estado finita (FSM) (ej. lavarropas)
- Continuamente variable (ej. dimmer, taladro)

El modelado ON/OFF es el que se asume en muchos trabajos. La máquina de estado finita debe cumplir, por definición, que la suma de las potencias de transición sea 0. Este tipo de electrodomésticos también es trabajado usualmente, aunque se requieren algoritmos más sofisticados. En particular [13] propone un algoritmo de resolución de conflictos basado en el algoritmo Viterbi. En este trabajo, los errores en la asignación de flancos a electrodomésticos llegan al 10 % y no se garantiza que aquellos electrodomésticos que consumen lo mismo sean separados.

Hoy en día se encuentran cada vez más electrodomésticos del tipo *inverter*, que si bien utilizan esa etiqueta muchas veces de manera comercial, es usual que presenten curvas de potencia difíciles de identificar, debido a que presentan un consumo variable a lo largo del tiempo. Un ejemplo es el “microondas *inverter*”, que en lugar de reducir la potencia encendiendo y apagando el magnetrón, como es usual, lo hacen emitiendo continuamente un valor de potencia menor.

2.3. Adquisición

Hoy en día pueden identificarse dos factores que atentan contra el desarrollo de esta área: (i) la caída de costos de medidores intrusivos, aunque este riesgo no es relevante para algunas aplicaciones del lado de la red de distribución, y (ii) la sofisticación y el uso masivo de electrodomésticos de consumo variable, que dificultarían enormemente el problema como se plantea hoy en día.

2.3. Adquisición

Un sistema NILM debe tener un medidor. Este medidor se encuentra normalmente cerca de la acometida del hogar, y usualmente mide, directa o indirectamente, la potencia consumida. Con este fin se puede medir solamente corriente y asumir el voltaje constante, o medir tanto corriente como voltaje. Una solución común es recurrir a los *smart meters* ya existentes en la red eléctrica, como los que se están desplegando en Uruguay. Esta solución es económicamente viable, pero tiene como principal falencia la baja frecuencia de muestreo: en el caso de los medidores siendo instalados en Uruguay, se tiene una medida cada quince minutos. La frecuencia de muestreo limita el desempeño de los algoritmos a ser utilizados sobre los datos. Una solución directa para la recolección de datos es adquirir un dispositivo comercial. Los mismos tienen como principales desventajas la baja frecuencia de muestreo y poca flexibilidad en su uso. Generalmente estos medidores tienen algún sistema de comunicación para transmitir datos a otros puntos, sea por Ethernet, Wi-Fi, o algún otro protocolo como ZigBee o MQTT. La frecuencia de muestreo máxima presentada por los dispositivos comerciales es de 1Hz [33] [14]. Hay consenso en que frecuencias de muestreo más elevadas posibilitan análisis más profundos y que para implementar clasificadores de buen desempeño se requieren al menos 4kHz [9]. La principal diferencia entre frecuencias de muestreo del orden de los Hz y del orden de los kHz, es que la mayor permite obtener información de dentro de un ciclo y poder observar con detalle el transitorio. Sin embargo, el hardware de adquisición suele ser más costoso para este caso, y además se requiere más ancho de banda para transmitir los datos a una estación centralizada. Se han implementado varios medidores diseñados específicamente para la aplicación, especialmente a la hora de relevar bases de datos de altas frecuencias. En cuanto a frecuencia de muestreo se podría indicar de forma gruesa que los medidores comerciales abarcan frecuencias menores a algunos Hz, las placas de adquisición comerciales, como tarjetas de sonido por ejemplo, pueden llegar a poco más de 100kHz, y el hardware específico no tiene limitaciones relevantes en frecuencia de muestreo. Otra característica que han de tener los adquisidores, en especial aquellos con alta frecuencia de muestreo, es contar con suficiente capacidad de cómputo. Para algunas frecuencias de muestreo el ancho de banda necesario para transmitir todas las observaciones es demasiado grande. Es por esto que se suele pasar los datos por una etapa de extracción de características para luego enviar estas últimas en lugar de los datos crudos. Resumiendo, el desempeño y el tipo de hardware utilizado en la recolección de datos determina las opciones disponibles en las siguientes etapas.

2.4. Visión general del problema

Una vez adquiridos los datos, se generan curvas de consumo agregado. Entre corriente y voltaje, la magnitud más importante es la corriente, ya que el voltaje se mantiene relativamente constante. En general, cuando la medida es de alta frecuencia se pasa por una etapa de extracción de características. Dos enfoques destacan a la hora de identificar electrodomésticos: los basados en detección de flancos y los que trabajan directamente sobre la curva de consumo agregado.

Capítulo 2. Revisión Bibliográfica

Los trabajos que no se basan en detección de flancos utilizan diversos enfoques de aprendizaje automático u optimización. Combinatorial Optimization es un algoritmo que resuelve el problema formulado en la Ecuación 2.1. Otro algoritmo clásico se basa en Factorial Hidden Markov Models [4]. Un trabajo interesante que utiliza con otras variables además de la potencia agregada es [27]. En este trabajo se muestra que incluir variables exógenas, como la temperatura, redunda en una mejora del desempeño. El último enfoque, sobre el cual se basa este trabajo, se centra en redes neuronales, que teóricamente tienen la capacidad para aprender a detectar flancos y asignarlos correctamente a cada electrodoméstico, dada una cantidad suficiente de datos etiquetados. La publicación en la que se basa fuertemente este proyecto es [19]. Los aportes más importantes de este trabajo son que las redes neuronales, entrenadas con un conjunto de electrodomésticos presente en ciertas casas, generalizan bien a una casa no vista durante el entrenamiento, bajo un enfoque de aprendizaje supervisado. Las dos arquitecturas con mejor desempeño de ese trabajo serán evaluadas en el presente Proyecto.

El enfoque de identificación que se basa en detectar flancos generalmente consta de dos etapas. La primera es la detección de flancos propiamente dicha. La segunda es el agrupamiento de esos flancos de forma que se identifiquen electrodomésticos. Un ejemplo de este enfoque se muestra en [28]. Este artículo ilustra una implementación de un sistema completo de NILM que obtiene un buen desempeño, basado en un algoritmo de *clustering* de flancos, definidos en aquellos t tales que $\Delta P_t = P_{t+1} - P_t > 5\text{W}$. Más formalmente, un flanco está definido como el período entre dos estados de régimen de las señales medidas. Los métodos de detección de flancos pueden dividirse en métodos heurísticos expertos, modelos probabilísticos y *matched filters* [3]. La detección de eventos se realiza en el caso del Proyecto implícitamente en el algoritmo implementado, basado en redes neuronales. En este Proyecto se implementó un detector de flancos para señales de consumo de electrodomésticos individuales, que se explica en el capítulo 5.

2.5. Algoritmos

Gran parte del esfuerzo de la comunidad académica se ha orientado a la identificación de buenas características que permitan atacar el problema. Un análisis detallado del estado del arte en cuanto a extracción de características se presenta en el Capítulo 5. Según [44], hay dos grandes enfoques para resolver el problema de NILM. El primero intenta buscar una combinación de electrodomésticos tal que su suma dé lo más parecido a la señal agregada que se está registrando. En este caso, la presencia de distintos electrodomésticos se asigna simultáneamente a distintos flancos detectados sobre un período largo de tiempo. En el otro enfoque los cambios de estado se asignan a las características o flancos detectadas una por una. Este enfoque es uno de reconocimiento de patrones, y técnicas como k-NN, clasificador de Bayes, y redes neuronales fueron utilizadas. El trabajo de [13] usa un enfoque intermedio, en donde se hace una asignación uno a uno y luego se implementan reglas de resolución de conflictos. Como el primer enfoque utiliza más información, se espera que tenga mejor desempeño en desagregación. Por otro lado, el segundo enfoque parece ser más robusto para electrodomésticos con patrones desconocidos. Una comparación directa entre ambos enfoques no es posible debido a la ambigüedad de las métricas utilizadas. Otra dirección en el desarrollo de algoritmos consta de la combinación de varios métodos buscando mejorar el desempeño. Esto se conoce como *information fusion*. Según [44], combinar métodos puede ayudar a aumentar el desempeño en un 10%.

2.6. Evaluación de resultados y métricas asociadas

2.6. Evaluación de resultados y métricas asociadas

En [29] se pretende dar un marco unificado a la hora de evaluar algoritmos de NILM. El marco de validación propuesto consiste en establecer que a la hora de describir los experimentos que se llevan a cabo con el algoritmo, sea definida claramente la base de datos utilizada, qué preprocesamiento se le realiza, por ejemplo al completar datos corruptos o faltantes, y se brinde una descripción de esta base de datos. Estos puntos son tenidos en cuenta por los autores a la hora de explicar el proceso de evaluación, haciendo foco en la descripción y procesamiento de los datos disponibles. Aunque la mayoría de los trabajos de NILM usan alguna métrica para reportar la efectividad del algoritmo, no hay acuerdo en qué métrica de clasificación tiene que ser utilizada. En [13] se utilizan tanto la fracción de eventos correctamente reconocidos como la fracción de energía total explicada. En otros trabajos se utiliza la diferencia entre la potencia estimada y real para un electrodoméstico específico. Otras métricas comunes incluyen la *accuracy*, la fracción de energía asignada a cada electrodoméstico, y fracción de intervalos de encendido no estimados. Si bien hay trabajos que proponen métricas que toman en cuenta la detección adecuada y la identificación adecuada, no se hace explícito el compromiso entre más falsas detecciones o ignorar muchas detecciones reales. La curva ROC muestra este compromiso. Según [44] el uso de la ROC puede ser beneficioso para el *benchmarking* de los algoritmos de NILM.

2.7. Productos comerciales

En esta última sección se presenta un resumen de los dispositivos comerciales hoy en día disponibles para el monitoreo no invasivo, enfocados principalmente en el consumidor residencial. Una lista de referencia puede encontrarse en la Wiki de NILM¹. Esta lista, si bien es extensa, no está completamente actualizada, y se observa el fenómeno de que muchas de las empresas listadas ya no se encuentran en actividad.

Los productos cuentan con su propio hardware de fácil instalación para la medida de tensión y corriente en la entrada de energía eléctrica al hogar. Una parte del procesamiento de los datos se realiza de forma local y la otra en la nube. En muchos casos, el preprocesamiento local se realiza para subir a la nube lo necesario y no hacer un uso excesivo del ancho de banda de la conexión a internet.

Algunos productos brindan monitoreo del consumo, reportando en tiempo real los datos para que el consumidor visualice la curva de carga, y pueda configurar alarmas que le permitan controlar el gasto mensual. Este tipo de equipos no incluyen algoritmos de desagregación de cargas. Otras empresas, como Smappee, utilizan algoritmos de desagregación para reportar la potencia consumida por cada electrodoméstico. Esta empresa tiene especial relevancia para este trabajo ya que se enfoca en un sistema de distribución de energía eléctrica que utiliza los mismos valores de voltaje y frecuencia que en Uruguay.

El producto de Smappee cuenta con conexión WiFi y es interactivo con el usuario para mejorar la precisión de las predicciones. Incluye una aplicación que permite generar eventos por activación, no detectando equipos con consumos menores a 300 W. Los métodos para mejorar las predicciones incluyen cuestionarios en los que se indica cuál es el equipo que se utilizó en cada momento y la posibilidad de registrar ciclos de encendido-apagado de un mismo equipo varias veces para facilitar su identificación. A lo largo de este proyecto, en colaboración con el grupo de investigación de identificación de electrodomésticos del IIE, se pusieron a prueba algunos de sus medidores y no se obtuvieron buenos resultados *out-of-the-box*. Los principales inconvenientes encontra-

¹<http://wiki.nilm.eu/companies.html>

Capítulo 2. Revisión Bibliográfica

dos fueron que los medidores requerían de mucha interacción humana para etiquetar electrodomésticos y que, aún luego de etiquetadas algunas instancias, ocurrían muchas confusiones entre los mismos. En particular, el equipo siempre reportaba el encendido de una aspiradora cuando en realidad estaba funcionando el aire acondicionado. Por la hora de actividad inusual para una aspiradora, cuando en realidad era un aire acondicionado encendido por la noche, se observa que las técnicas de desagregación de Smappee no utilizan variables externas como la hora del día.

Capítulo 3

Medidor no intrusivo

3.1. Resumen

En este capítulo se tratará con detalle el diseño y la implementación del medidor no intrusivo de voltaje y corriente. Este medidor es la componente de bajo nivel de la solución punta a punta del sistema de medida. Partiendo de los requerimientos, se tratarán tanto las componentes de hardware del medidor, circuito acondicionador y pinza de corriente, como los programas que controlan la adquisición, compresión y envío de medidas al servidor externo.

Las medidas tomadas por el dispositivo han de permitir la exploración y evaluación, en forma irrestricta, de algunos algoritmos de desagregación, asignación de potencia, o identificación de electrodomésticos. Esto implica requerimientos en cuanto a la frecuencia de muestreo y resolución utilizadas.

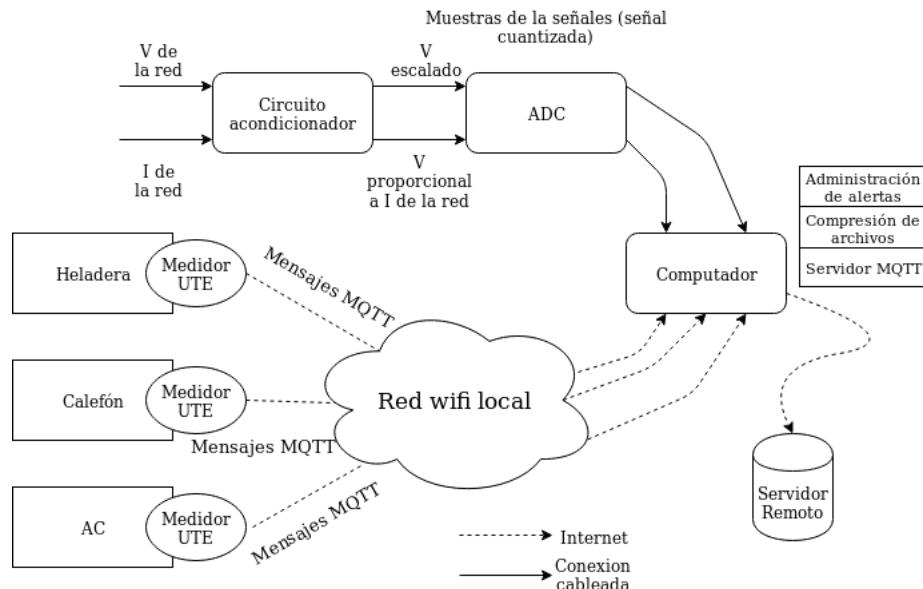


Figura 3.1: Esquema de la arquitectura general del sistema de medida. El medidor no intrusivo comprende al circuito acondicionador, ADC y computador.

Capítulo 3. Medidor no intrusivo

3.1.1. Sistema de medida

El medidor no intrusivo se enmarca dentro de un sistema de medida completo que permite recolectar una base de datos de validación. Este sistema integra el medidor no intrusivo con un conjunto de medidores de potencia eléctrica individuales y envía las medidas a un servidor remoto (Figura 3.1). Este sistema se analizará en profundidad en el Capítulo 4.

3.1.2. Requerimientos del medidor

En esta subsección se busca listar los requerimientos a nivel del medidor de NILM y explicar qué implicancias tienen en la elección de las distintas componentes que lo integran.

Tamaño del sistema y *casing*

El medidor no intrusivo ha de colocarse en el tablero eléctrico de un hogar, por lo tanto debe tener un tamaño adecuado que permita que sea instalado en dicho lugar. Dado que es un dispositivo experimental con el fin principal de recolectar datos (y no un dispositivo comercial) se toma como requerimiento de tamaño que quepa en el tablero de los integrantes del grupo. Este criterio implica una restricción de tamaño que indica que el *casing* debe ser menor a 18x13x6 cm.

A su vez, dados los niveles de voltaje y corriente que se tienen en el tablero eléctrico, el dispositivo debe contar con la debida protección. Esto implica un *casing* de un material aislante, en el que no queden terminales vivos accesibles. Además, es deseable que exista aislación galvánica y medidas de protección eléctrica frente a sobretensiones en la entrada del Conversor Analógico Digital (ADC).

Frecuencia de muestreo

Con el objetivo de evaluar los algoritmos de desagregación, asignación de potencia o identificación de electrodomésticos, se busca confeccionar un medidor que permita adquirir datos a una frecuencia de muestreo similar a las de las bases de datos académicas, por ejemplo 16kHz en el caso de [18]. Se desea, a su vez, que la frecuencia de adquisición no sea una limitante para la exploración de algoritmos. Aunque las bases de datos académicas no superen, en su inmensa mayoría, los 100kHz, se tiene en conocimiento de que la empresa Sense [2] muestrea a 1MHz y realiza extracción de características sobre estas medidas. Además, muestreos a más alta frecuencia (100MHz) pueden ser utilizados por métodos como el presentado en [31].

Dado que los algoritmos implementados se entrenaron sobre la base de datos UK-DALE [18], la base de datos de evaluación, recabada con el sistema presentado en el Capítulo 4, deberá ser de una frecuencia similar. Esta frecuencia es un parámetro en el código del medidor no intrusivo y debe configurarse de forma adecuada para facilitar la comparación con otros trabajos del área.

Resolución en potencia

En la misma línea que el punto anterior, se desea que las medidas realizadas no resulten en una limitante para los algoritmos. Como una de las líneas a explorar es la de asignación de potencia, se define una mínima resolución admisible en su medida de 10W.

3.2. Componentes

Capacidad de procesamiento suficiente

El medidor en sí mismo debe ser capaz de capturar los datos del ADC. En forma paralela debe poder comprimir los datos y enviarlos de forma programada a un servidor remoto. Si bien no son exigencias importantes para la capacidad de procesamiento que se tiene hoy en día, se debe estar seguro de que se pueden cumplir todas estas funciones sin que exista pérdida de datos. Una exigencia extra surge de que el medidor forma parte del sistema utilizado para recolectar la base de datos NILM-UY descrita en el Capítulo 4. Por cómo se diseñó el sistema, el medidor deberá poder tener operativo un servidor MQTT y realizar el procesamiento de datos de baja frecuencia (0,17Hz).

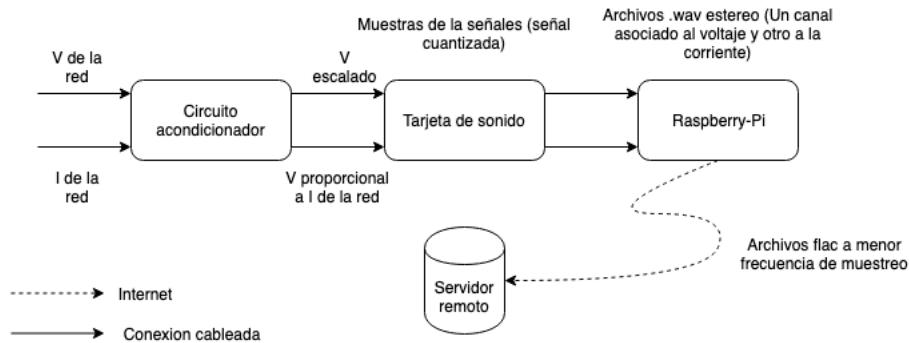


Figura 3.2: Esquema del sistema de medición no intrusivo

3.2. Componentes

El diseño e implementación de un medidor de voltaje y corriente no intrusivo involucra la elección de los componentes que permiten medir las señales de la red eléctrica, el diseño de un circuito acondicionador de las mismas, y la programación del ADC y del computador que se encargan de digitalizar estos datos, preprocesarlos y enviarlos a un servidor remoto. En esta sección se describen los componentes del medidor NILM. En el capítulo siguiente se describirá el sistema completo de recolección de datos, haciendo especial énfasis en la integración del medidor no intrusivo y en el software que controla todo el sistema.

Buscando satisfacer los requerimientos sin incurrir en un costo excesivo, pero sin que esto introduzca limitantes en las funcionalidades, se opta por componentes que permitan el desarrollo del medidor priorizando la agilidad de diseño e implementación.

Ordenador

Dados los requerimientos en cuanto a poder de cómputo para el medidor, y la necesidad de contar con un ordenador en el sistema de medida de la Figura 3.1, se decidió localizar el ordenador también en el medidor no intrusivo. Dado que el ordenador debe caber dentro del tablero de una casa, se optó por una Raspberry Pi 3 B+ por su reducido tamaño y buenas prestaciones de hardware: la misma cuenta con conectividad WiFi, lo cual es clave para satisfacer los requerimientos de conectividad que exige el sistema de recolección de datos presentado en la Sección 3.1.1. Además, al tener sistema operativo, se cuenta con una interfaz amigable que permite un prototipado más ágil. Como sistema operativo se utilizó Raspbian, aunque es posible utilizar otros. No todos tienen interfaz gráfica ni incluyen todos los drivers *out-of-the-box*.

Capítulo 3. Medidor no intrusivo



Figura 3.3: Raspberry Pi 3 B+

3.2.1. Conversor Analógico Digital (ADC)

Como ADC se decidió utilizar una tarjeta de sonido. La principal ventaja es que fue diseñada para muestrear a frecuencias mayores o iguales a 44.1 kHz (para cubrir el ancho de banda audible por el oído) por lo que no impone ninguna restricción práctica en cuanto a la frecuencia de muestreo. Además ya tiene incorporada un filtro pasa bajos (de anti-aliasing) adecuado y presenta una alta impedancia de entrada (prácticamente no toma corriente). En particular se optó por una placa de audio diseñada especialmente para la Raspberry-Pi, que brinda facilidades para su montaje sobre la Raspberry conectándose a través de los pines GPIO (General Purpose Input Output). La placa utilizada es la “Raspberry Pi *stereo sound card*” fabricada por Audio Injector [1], y permite una frecuencia de muestreo de hasta 96kHz y una profundidad en bits de hasta 32 bits por muestra. Posee dos entradas de línea por conexión RCA cuyo rango de entrada nominal es de 3,3V. Al testearse el rango de entrada hasta la saturación se encuentra que es mayor al voltaje nominal, pero ese margen se asume de seguridad y se diseña considerando el valor nominal de 3,3V pico a pico a la entrada.

Se consideraron otras opciones alternativas de tarjetas de sonido, como el modelo ATR2USB y tarjetas de sonido profesionales. Las desventajas de la primera consistían en que el rango de entrada era pequeño (hasta 0,47V), se incluye una etapa de pre-amplificación que podría distorsionar la señal, y es una tarjeta mono, por lo que se necesitan dos tarjetas para adquirir las señales de voltaje y corriente de forma simultánea. De esta forma, surge como requerimiento no excluyente contar con una tarjeta stereo, ya que tiene dos ADC independientes y permite adquirir de forma simultánea voltaje y corriente. A su vez, dado que no se está adquiriendo una señal de voz sino que son señales de voltaje y corriente, es deseable que las mismas no sean amplificadas ni distorsionadas como cuando se tratan señales de voz. Por lo tanto existe el requerimiento de que la tarjeta de sonido utilizada tenga una entrada de línea, que se caracteriza por un rango de entrada mayor y no tiene una etapa de pre-amplificación, en oposición a la entrada de micrófono, más usual, como la que existe en la placa ATR2USB.

Por otro lado, también se consideraron tarjetas de sonido profesionales, la principal desventaja de las mismas es su elevado precio dentro del contexto del proyecto.

3.3. Circuito acondicionador



Figura 3.4: Raspberry Pi stereo sound card

3.2.2. Circuito analógico acondicionador

Para el acondicionamiento de las señales de voltaje y corriente del medidor no intrusivo a ser leídas por el ADC, se diseñó un circuito analógico que transforma estas señales en voltajes dentro del rango adecuado. Este circuito es el que se presenta en la Sección 3.3. Además, el medidor no intrusivo utiliza una pinza de corriente para leer la corriente que pasa por el circuito acondicionador.

3.3. Circuito acondicionador

En esta sección se explica el acondicionamiento de las señales a medir. Se esperan $230V_{RMS}$ de valor eficaz en voltaje y una corriente que depende del consumo, pero que se asume inferior a $40A_{RMS}$. Estas señales han de llevarse al rango de entrada del ADC, buscando el mayor aprovechamiento posible de su rango dinámico. Dado que se muestrea en alta frecuencia (14kHz), es de especial importancia tener un circuito acondicionador con una respuesta en frecuencia coherente con la frecuencia de muestreo. Notar que el circuito se diseña para instalaciones monofásicas.

3.3.1. Señales de entrada

Voltaje

La diferencia de potencial en una instalación monofásica en Uruguay es una señal sinusoidal con un valor *Root Mean Squared* (RMS) de $230V_{RMS}$. La frecuencia de esta señal es de 50Hz. La empresa que ostenta el monopolio de la distribución de energía eléctrica en Uruguay es la Administración Nacional de Usinas y Trasmisiones Eléctricas (UTE).

La Unidad Reguladora de Servicios de Energía y Agua (URSEA) dispone tolerancias sobre estos valores nominales. En particular, la diferencia porcentual ΔV entre el voltaje de la red y el nominal, está acotada por $-12\% \leq \Delta V \leq +6\%$. Fuera de ese rango la empresa distribuidora incurre en costos extra, según el Texto Ordenado de Resoluciones de la URSEA.

Capítulo 3. Medidor no intrusivo

La URSEA establece que la empresa distribuidora deberá otorgar compensaciones si el voltaje supera ciertos niveles establecidos, empezando el rango de mayor costo para la empresa distribuidora a los $(230 + 18)V = 248V$. Es este valor el que se utiliza como el máximo voltaje de entrada esperado. En cuanto a las cuestiones de frecuencia, en ese mismo documento se establece que la distorsión armónica máxima (THD) sea de 8 %.

Corriente

La corriente demandada por un hogar es variable y depende del consumo del mismo. El servicio de suministro que brinda UTE es según la potencia contratada, por lo que definiendo una potencia máxima esperada para un hogar, queda determinada la corriente máxima que demandará, ya que el voltaje es un valor conocido. Esta limitación se hace real con las llaves térmicas que limitan la corriente. Para instalaciones monofásicas UTE brinda cuatro niveles de potencia (kW): 3,7, 4,6, 7,4, 9,2. Si bien valores mayores a 7,4 kW no son tan comunes, se define como máxima potencia esperable en un hogar 9,2 kW, lo que implica una corriente máxima de 40A_{RMS}.

Es pertinente aclarar que esta señal no necesariamente será sinusoidal, ya que depende de las características de la carga del hogar. En la actualidad, debido a la gran presencia de elementos electrónicos en los hogares, que utilizan fuentes conmutadas, la señal de corriente presenta distorsiones apreciables. Esta distorsión se puede observar en la Figura 3.5.

3.3.2. Esquema general del circuito

El circuito, presentado en la figura 3.6, se compone de la pinza de corriente STC-013 (Apéndice B), que genera una corriente de un valor pequeño, un transformador de voltaje de señal (ZMPT101B) que se conecta a la red y entrega al circuito una diferencia de potencial dada por su relación de vueltas, un conjunto de resistencias que se utilizan para llevar las señales de entrada a voltajes en el rango del ADC, y diodos de protección.

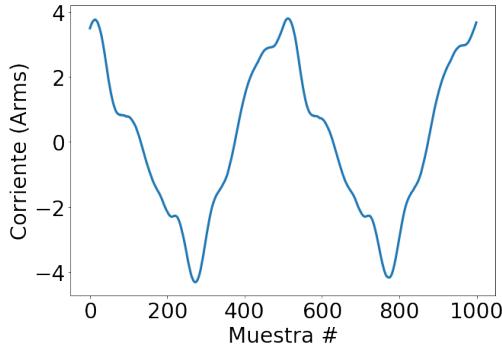
Sub-circuito medidor de voltaje

La rama superior del circuito presentado es la encargada de acondicionar la señal de voltaje a ser leída por el ADC. Consiste en un transformador de señal de relación 1:1 y un divisor resistivo. La ubicación de las resistencias no es casual. Si bien considerar componentes ideales implica, teóricamente, que no importa donde se las coloque, en la realidad sí es importante conseguir una distribución que no sature el núcleo ferromagnético del transformador. Por ejemplo, si todas las resistencias se colocaran en el secundario, se saturaría el núcleo. A pesar de que la corriente que circula por el circuito siempre es la misma (ya que queda fijada por el valor total de la suma de las resistencias), en bornes del primario se tendría el voltaje de la red, implicando allí una mayor potencia. Esto fue comprobado experimentalmente y se observó cómo la señal de voltaje era deformada por la saturación.

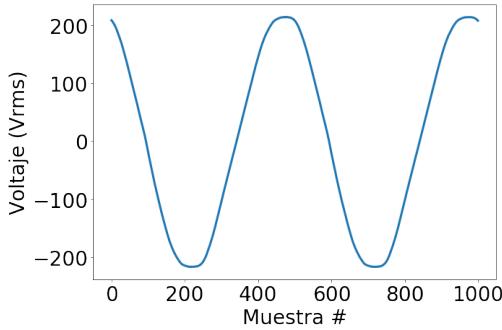
Sub-circuito medidor de corriente

La rama inferior del circuito es la encargada de traducir la señal de corriente de la red eléctrica y convertirla en una señal de voltaje adecuada para el ADC. Para esto, se utiliza un transformador de corriente (denominado comúnmente *Current clamp*) y una resistencia para generar una señal de voltaje proporcional a la corriente. La función de la pinza es obtener una señal de corriente proporcional a la red (sin cargarla) y reducirla

3.3. Circuito acondicionador



(a) Corriente.



(b) Voltaje.

Figura 3.5: Ejemplo de medidas de corriente y voltaje para un aire acondicionado. Medidas recolectadas por estudiantes del curso Taller Fourier EDM 2019. Se observa una considerable deformación en la señal de corriente.

a un valor manejable. En este caso se utiliza una pinza con relación de transformación 100A : 50 mA. La resistencia es simplemente un factor de escala que permite llevar la señal a medir hacia valores en el rango de entrada del ADC, pero respetando el límite en su valor para no saturar al transformador. Para acoplar la pinza a la red, hay que engancharla únicamente a uno de los cables del tablero eléctrico (para medir la corriente por la fase) ya que sino se tendría la corriente entrante y saliente de la casa, generando una medida nula.

Pinza de corriente

La pinza de corriente es uno de los elementos fundamentales en la creación de un medidor de corriente no intrusivo. Dado los requerimientos en frecuencia del medidor, se realizaron experimentos sobre la pinza de corriente YHDC SCT013-000 CT con el objetivo de relevar su respuesta en frecuencia. Los resultados muestran que, experimentando bajo el supuesto de que el banco de resistencias utilizado en los ensayos tiene una respuesta en frecuencia plana, la pinza tiene un comportamiento atenuante en frecuencias altas. Por otro lado, en el Apéndice B se discute y se muestra que esta suposición resulta dudosa.

Capítulo 3. Medidor no intrusivo

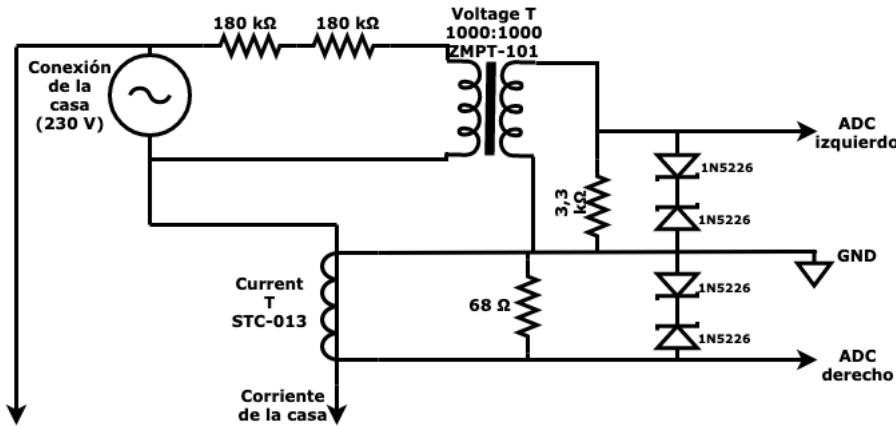


Figura 3.6: Esquemático del circuito acondicionador diseñado



Figura 3.7: Transformador ZMPT101B

Más allá de la observación previa, se concluye que la pinza de corriente permite distinguir del ruido componentes de hasta al menos 7kHz. Si bien esto depende, en general, de la magnitud de la señal a la entrada, se concluye que para una onda cuadrada y una corriente eficaz similar a la consumida en residencias, la pinza de corriente no anula completamente las componentes de altas frecuencias.

Se sugiere al lector a consultar el Apéndice B para ahondar en los experimentos y en las observaciones realizadas.

Protecciones del circuito

Para satisfacer el requerimiento de que exista aislación galvánica con la red eléctrica, se utilizan transformadores para la medida de ambas señales.

Como protección del ADC frente a posibles sobre-tensiones o sobre-corrientes inesperadas, se implementa un bloque limitador de tensión a su entrada mediante diodos Zener en anti-serie.

3.3.3. Diseño del circuito

Valores para corriente

La pinza STC-013 tiene una relación de $N = 100A : 50mA = 2000$. Esto indica que a la entrada de la pinza podrá tenerse un valor máximo, en RMS, de $I_{in\ max}^{ref} = 100A$.

3.3. Circuito acondicionador

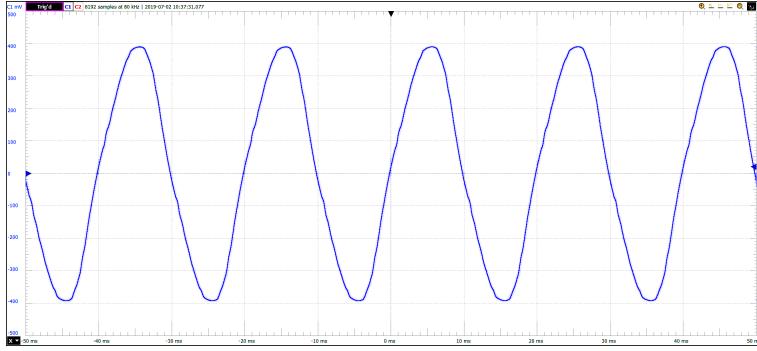


Figura 3.8: Medida de voltaje saturando el transformador

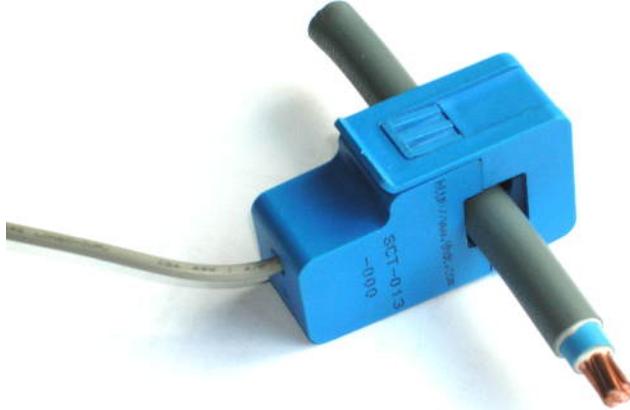


Figura 3.9: Pinza de corriente SCT013

De la hoja de datos adjuntada en el Apéndice C, se tiene que la resistencia máxima a colocar en el secundario ha de ser $R_{\text{sec max}}^{\text{ref}} = 10\Omega$. Esto es una restricción de potencia. Dado que la máxima resistencia admitida a $I_{\text{in max}}^{\text{ref}}$ es $R_{\text{sec max}}^{\text{ref}}$, se tiene que la potencia máxima es:

$$P_{\text{pinza max}} = \left(\frac{I_{\text{in max}}^{\text{ref}}}{N} \right)^2 R_{\text{sec max}}^{\text{ref}} = \left(\frac{100\text{A}}{2000} \right)^2 10\Omega = 25\text{mW} \quad (3.1)$$

Teniendo en cuenta esta restricción, al asumir una corriente máxima a la entrada $I_{\text{in max}}$ menor que 100A, se estará imponiendo una restricción sobre la resistencia máxima a colocar en el secundario $R_{\text{sec max}}$. La corriente máxima para una casa se define como $I_{\text{in max}} = 40\text{A}$. Por lo tanto:

$$\left(\frac{I_{\text{in max}}}{N} \right)^2 R_{\text{sec max}} < P_{\text{pinza max}} \quad (3.2)$$

Capítulo 3. Medidor no intrusivo

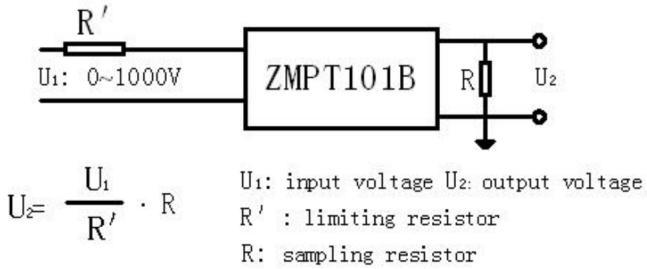


Figura 3.10: Esquema de uso, extraído de la hoja de datos

$$\Rightarrow R_{\text{sec max}} < P_{\text{pinza max}} \left(\frac{N}{I_{\text{in max}}} \right)^2 = 25\text{mW} \left(\frac{2000}{40\text{A}} \right)^2 = 62,5\Omega \quad (3.3)$$

Dada la serie E12 de 1/4 W, se elige un valor de 68Ω para la resistencia de salida. El valor anterior de la serie es de 56Ω . Dado que no se espera trabajar demasiado tiempo sobre el máximo de potencia, se justifica el pequeño exceso de la elección (menos de un 10%).

Valores para voltaje

Para la medida de voltaje, se utilizó el transformador ZMPT101B. Su hoja de datos reporta características y restricciones de funcionamiento, como puede verse en el Apéndice C. Entre estos valores están la máxima carga $R_{\text{max}} = 200\Omega$ y la máxima corriente $I_{\text{max}} = 3\text{mA}$ soportadas, que se traducen en una restricción de potencia:

$$P_{\text{max}} = (I_{\text{max}})^2 R_{\text{max}} = (3\text{mA})^2 200\Omega = 1,8\text{mW} \quad (3.4)$$

Por otro lado, la salida de esta etapa del circuito ha de ser menor al máximo soportado en la entrada del conversor analógico-digital (ADC), que es $V_{\text{in max pp}} = 3,3\text{V}_{\text{pp}}$ para los componentes seleccionados. Eligiendo una disposición como la mostrada en la figura 3.10, la restricción es (en valores eficaces):

$$V_{\text{out max}} = \frac{R}{R'} V_{\text{in max}} \leq \frac{3,3\text{V}_{\text{pp}}}{\sqrt{2}} \quad (3.5)$$

Las dos restricciones restantes vienen dadas, por un lado, por la máxima corriente aceptada por el transformador, y por otro, por las resistencias y potencias nominales de los resistores comerciales. Se eligen resistores de 1/4W de la serie E12.

La lista completa de restricciones, en valores eficaces, es:

$$\frac{V_{\text{in max}}}{R' + R} \leq I_{\text{max}} = 3\text{mA} \quad (3.6)$$

$$R \left(\frac{V_{\text{in max}}}{R' + R} \right)^2 \leq P_{\text{max}} = 1,8\text{mW} \quad (3.7)$$

$$\frac{R}{R'} V_{\text{in max}} \leq \frac{3,3\text{V}_{\text{pp}}}{\sqrt{2}} \quad (3.8)$$

3.3. Circuito acondicionador

Con el objetivo de aprovechar el mayor rango de entrada posible se transforma la última inecuación (3.8) en una igualdad. De ahí que $R' = 106R$. Resolviendo las restantes inecuaciones en R se encuentra que la inecuación 3.7, que involucra a la potencia, es más restrictiva: establece que $R > 3k\Omega \Rightarrow R' > 320k\Omega$.

Los valores finales, correspondientes a resistencias de la serie E12 de 1/4W son $R = 3,3k\Omega$; $R' = 2(180k\Omega)$.

Dimensionamiento de los diodos de protección

Como protección se colocan diodos Zener, modelo 1N4728, en antiserie. En esta configuración, el voltaje de corte viene dado por $V_{\text{corte}} = V_z + V_\gamma$, siendo V_z el voltaje a partir del cual comienza a pasar corriente cuando el diodo se polariza en inverso y V_γ el voltaje a partir del cual empieza a pasar corriente cuando el diodo se polariza en directa. Dado que están en antiserie, siempre que el voltaje en bornes sea mayor que V_{corte} un diodo estará en zona Zener y el otro conduciendo en directo. Teniendo esto en cuenta, diodos Zener de 3,3V de V_z , y con un V_γ de 0,7V aproximadamente, serían adecuados. Sin embargo, el modelo del diodo que considera solamente V_z y V_γ no es lo suficientemente descriptivo (ver Figura 3.11). Usar los valores de los diodos expuestos anteriormente implica dejar pasar una corriente no despreciable que deforma la señal. Usando valores de $V_z = 5,1V$ se resuelve el problema.

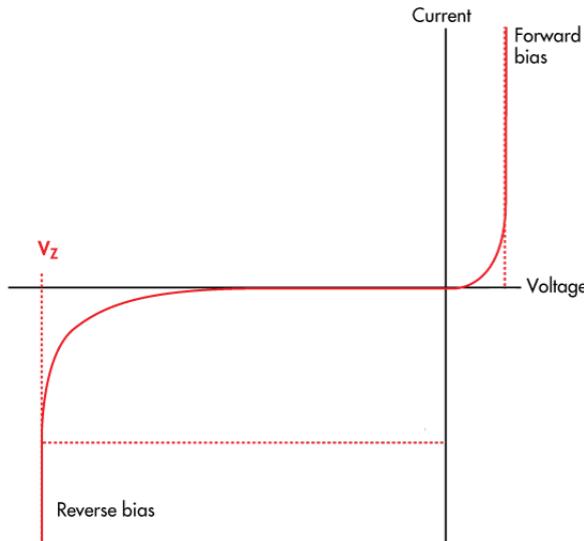


Figura 3.11: Comportamiento del diodo tipo Zener. Un modelo simplificado toma en cuenta solamente V_z y V_γ , las asíntotas verticales en la figura.

3.3.4. Implementación del circuito

Una vez diseñado el circuito y luego de haber dimensionado todas las componentes de acuerdo a los insumos disponibles en plaza, se realizó la implementación del circuito.

La implementación consistió en diseñar, imprimir y confeccionar el circuito en forma impresa (PCB). El diseño se realizó utilizando la herramienta de software libre KiCad, se imprimió el mismo sobre papel fotográfico y se lo pasó a la plancha de cobre

Capítulo 3. Medidor no intrusivo



(a) Proceso de impresión del circuito sobre la placa.



(b) PCB terminado.

Figura 3.12: Implementación del circuito

utilizando una solución de acetona y alcohol. Finalmente, una vez que el diseño en tinta estuvo sobre la plancha de cobre, se sumergió la misma en percloruro de hierro para disolver el cobre de las partes que no conforman el circuito.

3.3.5. Otros circuitos evaluados

Para el diseño e implementación del circuito se investigaron trabajos existentes en el área. Se refiere al lector al Apéndice A donde se presentan los trabajos del Ing. Maggi y del Phd. Jack Kelly.

3.4. Software del medidor

Las señales de voltaje en el rango del ADC llegan desde el circuito acondicionador a la tarjeta de sonido mediante cables del tipo RCA. Se utiliza la entrada de línea estéreo de la placa. En la Raspberry Pi se ejecuta un script de *Python* que consiste en un bucle infinito que lee las muestras capturadas por el ADC. Este bucle dispara un proceso paralelo cada vez que llena un buffer de 4096 muestras. Este proceso se encarga de guardar estas muestras en disco en un archivo .wav. Se trabaja con archivos de audio ya que existe un amplio desarrollo en esta área en cuanto a paquetes y software de preprocesamiento. Un archivo .wav es esencialmente un archivo binario que registra el

3.5. Calibración

valor de la señal cuantizada para cada paso de tiempo discreto. Es vital que el proceso de escritura en disco sea ejecutado en paralelo, así el proceso central que se encarga de recibir las muestras del ADC y alojarlas en memoria no es interrumpido y no se pierde ninguna muestra. Finalmente, un tercer proceso se ejecuta, también en paralelo, al cabo de una hora, donde se llama al software de procesamiento de audio *SoX* [38] para que comprima el archivo .wav correspondiente a esa hora de medición a un archivo .flac y le baje la frecuencia de muestreo a una más adecuada para la aplicación: 14 kHz. Se optó por hacer esta conversión en *SoX* ya que es más eficiente computacionalmente (más veloz), que hacerlo en *Python* usando las librerías *pydub* [34] o *python audio tools* [23] cuyas implementaciones son en *Python* puro y no en C. Se comprime en formato .flac para no tener pérdida y poder reconstruir la señal en cualquier momento posterior. La elección de este formato de archivos también está justificada por la alta eficiencia en compresión que presenta. El formato de compresión fue evaluado y resultó ser el mejor para el almacenamiento de bases de datos de NILM [21]. Es pertinente aclarar que una vez que se dispara este proceso, el proceso central crea un nuevo archivo .wav para comenzar a guardar las muestras de esta nueva hora de forma de que no exista ningún inconveniente ni pérdida de datos entre la conversión de una hora de muestras ya guardadas y la captura de nuevas muestras.

A parte del script que controla al medidor no intrusivo, se implementa otro script que se ejecuta de forma periódica e independiente para enviar todos los archivos .flac al servidor remoto y que además comprueba que el envío haya sido exitoso. En caso de éxito se eliminan los archivos de la tarjeta SD (almacenamiento local) de la Raspberry.

3.5. Calibración

Una vez escogidos todos los componentes del circuito propuesto para adquirir la señal, se realizó su calibración, tanto de voltaje como corriente. Podría realizarse la calibración y conseguir factores para las componentes individuales del medidor, pero por motivos de practicidad se decide calibrar directamente el sistema completo, es decir, el circuito acondicionador, la pinza y el ADC.

Los valores de calibración de voltaje y corriente son constantes a ser utilizadas para obtener las medidas en sus unidades correspondientes (V y A) a partir de las medidas sin unidades generadas por el sistema de medición. La calibración se realizó a la frecuencia y voltaje de la red, 50 Hz y 230 V, y a 4A aproximadamente. Idealmente se podría haber realizado la calibración para más valores de corriente, manteniendo la frecuencia y el voltaje que son los valores de la red, pero esto supondría una complicación práctica innecesaria para el alcance del proyecto. Una idea de un proceso de calibración más profundo puede encontrarse en el estudio de la pinza de corriente presentado en el Apéndice B.

Para calibrar se utilizó el generador de ondas APS-1102 GW Instek y una estufa de 1000 Watts, la cual es una carga puramente resistiva. Se inyecta una señal de voltaje 230V sinusoidal pura de 50Hz a la estufa. Luego se comienza a medir con el sistema por un minuto, tanto voltaje como corriente. Además se adquieren los valores que indica la fuente de voltaje RMS y corriente RMS entregada.

Por último, para obtener ambas constantes de calibración, se divide el valor RMS calculado digitalmente entre el valor real que indica la fuente.

3.6. Conclusiones

El medidor fue implementado con éxito. Es capaz de registrar corriente y voltaje con una profundidad de 32 bits a una frecuencia de 14kHz. El medidor es apto para

Capítulo 3. Medidor no intrusivo

ser utilizado en instalaciones monofásicas de Uruguay de hasta 9,2kW. Además, tiene capacidad de procesamiento suficiente como para comprimir y enviar datos a un servidor remoto vía Wi-Fi. Sirve además como centro de cómputo para el sistema de medición que se plantea en el siguiente capítulo, donde se describe cómo se recaba la base de datos.

Capítulo 4

Sistema de recolección de datos

4.1. Resumen

En este capítulo se presenta el diseño de un sistema de medición que permite integrar el medidor no intrusivo introducido en el Capítulo 3 con un conjunto de medidores de potencia eléctrica individuales (del tipo Smart Plug), con el objetivo de recolectar una base de datos de validación.

El sistema puede dividirse en tres componentes principales: el medidor no intrusivo, los medidores intrusivos, y un computador. Los medidores intrusivos son un conjunto de medidores de voltaje y potencia eléctrica diseñados por UTE. El sistema se completa con un computador que se encarga de cumplir las funciones de recibir los datos enviados por los medidores individuales, capturar los datos adquiridos por el ADC del medidor no intrusivo, y preprocesar y comprimir todo el conjunto de datos antes de enviarlo finalmente a un servidor remoto.

Como producto del sistema de medida, se presenta finalmente la base de datos NILM-UY. Esta base de datos contiene el monitoreo de consumo de energía eléctrica de dos hogares de Montevideo. En la misma, se registran valores agregados en la entrada del hogar y valores individuales de un conjunto de electrodomésticos seleccionados.

4.2. Red de medidores intrusivos

Para realizar las medidas individuales de cada electrodoméstico se utilizan unos medidores comerciales del tipo Smart Plug diseñados por la empresa iTead y adquiridos por UTE, que adaptó su firmware. En vez de enchufar directamente a la red eléctrica el electrodoméstico en cuestión, se lo enchufa al medidor de UTE y éste, utilizando el protocolo MQTT, envía las medidas a un servidor.

Los medidores individuales de cada electrodoméstico registran medidas de potencia activa y voltaje. La medida de potencia se realiza promediando la potencia en un intervalo pequeño alrededor del instante de muestreo, con una frecuencia de una medida por minuto.

Para la recolección simultánea de datos de consumo etiquetado de distintos electrodomésticos se dispuso una red de los medidores previamente descritos. Para este sistema de medición, el computador cumplió la función de broker MQTT utilizando la implementación de *Mosquitto* [25]. En este protocolo el broker es una entidad que recibe mensajes de las distintas unidades de la red con un tópico asociado y las reenvía a todas las unidades que estén suscritas a este tópico. Mediante un script de

Capítulo 4. Sistema de recolección de datos



Figura 4.1: Medidor smart plug desarrollado por UTE

Python se implementó un programa que permanece en un bucle infinito y se suscribe al tópico de los medidores, y cada vez que recibe un mensaje MQTT agrega los datos recibidos a un archivo de texto separado por comas (csv). Este programa, crea, para cada medidor intrusivo, un archivo .csv distinto por día. La comunicación entre la red de medidores y el servidor MQTT se configura asignando direcciones IP fijas dentro de una red local. Todos los equipos que forman parte del sistema de medida deben estar conectados a la red de área local (LAN) de la casa. Además es necesario que la Raspberry del medidor tenga acceso a internet tanto para sincronizar su fecha y hora (ya que estos computadores no poseen un reloj físico) como para poder enviar los datos al servidor remoto.

Para el correcto funcionamiento de los medidores, es necesario que el servidor se inicialice antes de que los mismos se enciendan (requerimiento impuesto por el diseño de UTE). Es por esto que ante un eventual corte del suministro eléctrico o corte de la red WiFi se deben volver a encender los medidores luego de que el servidor haya comenzado a ejecutarse. Para atender esta eventualidad, se implementó un sistema de alertas vía mail que avisa al usuario cada vez que comienza a estar activo el servidor. Una vez que el servidor está activo, simplemente hay que reiniciar todos los medidores (desenchufándolos y volviéndolos a enchufar) y la comunicación comenzará sin problemas. Como se explicará en la Sección 4.3, el medidor no intrusivo está programado para encenderse automáticamente al ser conectado.

Para enviar los archivos .csv correspondientes a días completos se implementa un script que se ejecuta de forma periódica (una vez al día) e independiente de los programas del sistema de medida. Este script comprueba que el envío haya sido exitoso y en ese caso libera el almacenamiento de la Raspberry. Si el envío no es exitoso, el archivo no se borra, a la espera de ser enviado la próxima vez en que el script se ejecute automáticamente.

4.3. Integración de los sistemas de medida

En esta sección se describe cómo se controlan el medidor no intrusivo y la red de medidores no intrusivos desde la Raspberry, de forma de que midan en forma simultánea, sin perder datos y que envíen de modo periódico los datos a un servidor remoto para liberar el almacenamiento local. Es deseable que el sistema sea robusto frente a fallas, por lo que se busca que, si cualquiera de los procesos de medición se ven interrumpidos, comiencen a ejecutarse nuevamente sin acción del usuario.

La Raspberry es capaz de alojar un servidor MQTT que reciba los mensajes de

4.3. Integración de los sistemas de medida

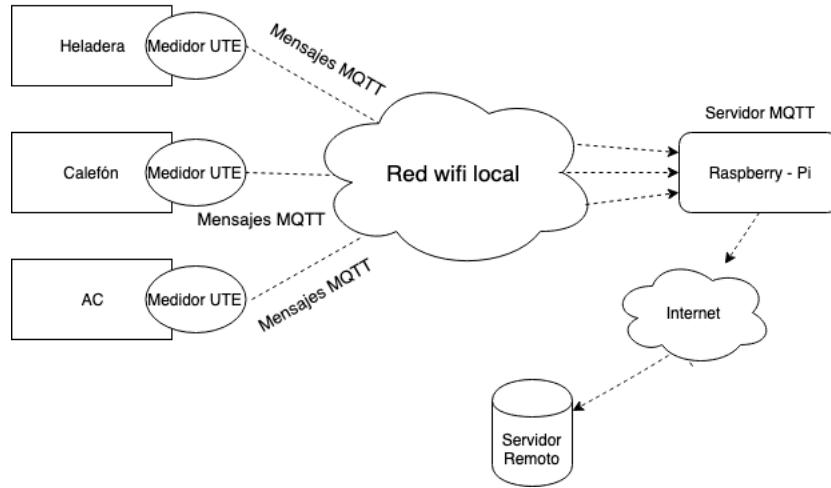


Figura 4.2: Esquema del sistema de medición de consumo etiquetado

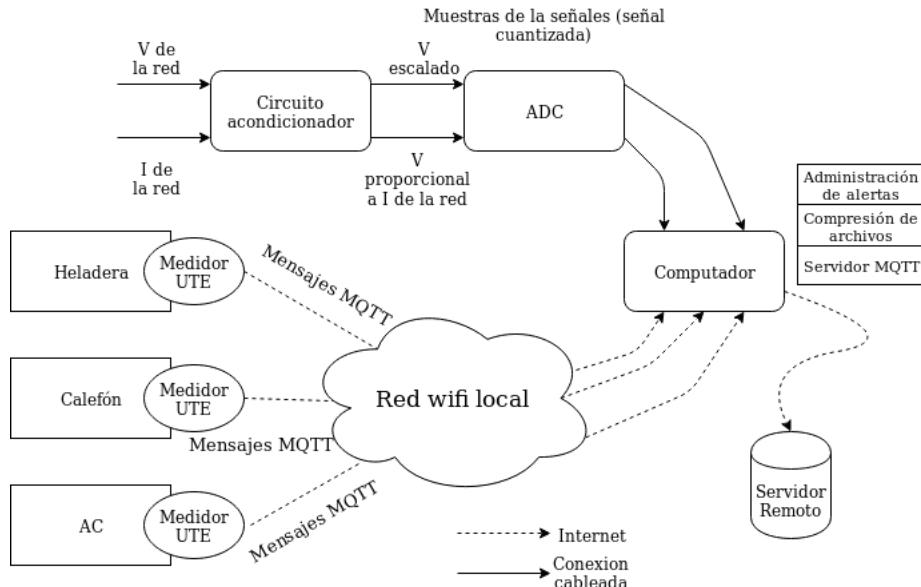


Figura 4.3: Esquema de la arquitectura general del sistema

los medidores individuales, cuenta con conectividad WiFi, lo cual es clave para que pueda funcionar como servidor MQTT. Todo el desarrollo de software relacionado con el sistema de medición fue hecho en *Python*. Se escogió este lenguaje por su versatilidad, agilidad en el desarrollo y por la experiencia existente en el grupo. Para tareas puntuales, se utilizó también *bash scripting* y *SoX* para al procesamiento de las señales provenientes del medidor no intrusivo.

La coordinación y ejecución de ambos sistemas de medida se realizó utilizando la herramienta de control de procesos *Supervisord* [5]. Lo que se ejecuta como sistema de medida es un proceso de *Supervisord* que se encarga de levantar automáticamente

Capítulo 4. Sistema de recolección de datos

ambos programas de medida en paralelo, lleva un log de los sucesos que se producen, y supervisa su estado. En caso de que alguno de los dos programas de medida fallen, *Supervisord* se encarga de reiniciarlos automáticamente.

Por otro lado, es necesario que los scripts de envío de archivos hacia el servidor remoto se ejecuten de forma periódica. Cada una hora para las medidas no intrusivas y una vez por día para las medidas etiquetadas. Lo anterior se logró utilizando la herramienta *cron*, donde se programa que se ejecuten estos scripts cada una hora en el caso de NILM y al inicio de cada día para las medidas intrusivas.

Finalmente, se tomó la decisión de diseño de que el sistema completo de medición comience a funcionar al enchufarse (al encenderse la Raspberry), sin otra interacción por parte del usuario. Para esto se utilizó *cron*, programando que la ejecución de *Supervisord* se dispare cada vez que se encienda la placa. Se deja habilitado el uso de *Secure Shell* (SSH), protocolo que permite el acceso remoto a un servidor por medio de un canal seguro, para acceder al sistema desde otro ordenador, con el objetivo de monitorizarlo sin tener que acceder físicamente a él.

4.3.1. Servidor remoto

Para la implementación de un servidor remoto de alojamiento de archivos se dispuso de un computador a la cual se le asignó un *Fully Qualified Domain Name* de forma que sea accesible a través del DNS utilizando internet. Como opción de servidor se optó por montar un servidor de transferencia de archivos sobre el protocolo FTP, utilizando *vsftpd* [8], por su velocidad y su popularidad, lo cual redunda en que exista mucha documentación y soluciones a los posibles problemas que puedan aparecer. Al ser un servidor liviano en cuanto al consumo de recursos del ordenador existe la posibilidad de montarlo tanto en una Raspberry Pi como en un PC sin grandes prestaciones.

4.4. Base de datos NILM-UY

4.4.1. Motivación

La base de datos NILM-UY se presenta en un contexto tecnológico en el cual comienza a existir interés respecto al consumo inteligente de energía eléctrica (medidores inteligentes, gestión de la demanda, etc.) en el Uruguay. Se espera que no solamente sirva para fines de monitoreo no intrusivo, sino también para otras posibles aplicaciones, como podrían ser optimización de despacho dadas señales de precios o análisis de consumo. Se espera que sea la primera base de datos orientada a NILM en el país.

4.4.2. Sistema de medida

Para todas las medidas se utilizó el sistema previamente detallado tomando las medidas no intrusivas a 14kHz. Para no ocupar excesivo ancho de banda en las casas donde se instala el sistema, se opta por tener el servidor de datos de forma local en cada hogar, ya que de esta manera, los datos se envían a través de la LAN y no se consume ancho de banda del servicio domiciliario de internet. Esta decisión se toma para evitar gastos innecesarios y no por falta de infraestructura. El sistema de medida se enfoca en la recolección de una base de datos para fines académicos, no para una aplicación comercial. En este último caso, la comunicación hacia el servidor central debería replantearse tomando como ejemplo lo que hace la empresa Sense. En lugar de enviar los datos crudos, se podrían calcular los descriptores de alta frecuencia en la Raspberry y solo enviar estos datos procesados que tendrán una menor cadencia.

4.4. Base de datos NILM-UY

4.4.3. Descripción de los datos

Casas monitoreadas

Se recolectaron datos de dos hogares de Montevideo con características disímiles:

- Casa 1: apartamento de dos dormitorios, dos adultos que pasan la mayor parte del día fuera del hogar. Se instalaron 8 medidores correspondientes a: dos aires acondicionados, calefón, heladera, lavarropas, microondas y dos tomas de corriente de propósito general.
- Casa 2: casa de tres dormitorios, dos adultos y tres niños. Dado que se instaló durante las vacaciones, es presumible que la familia se encontró parte significativa del día en su casa. Se instalaron 9 medidores: dos aires acondicionados, calefón, lavavajillas, heladera, jarra eléctrica, lavarropas, microondas y secarropas.

Período temporal

Las medidas para ambas casas no fueron tomadas de forma simultánea. En la Casa 1 se relevaron medidas desde el 05/09/2019 hasta el 03/11/2019, contabilizándose 59 días de medición ininterrumpidos. Por otro lado, en la Casa 2 el medidor fue instalado el 19/12/2019 y se cuenta con dos períodos de medida: del 19/12/2019 al 01/01/2020, y del 16/01/2019 al 05/02/2020.

Cadencia de las medidas

Las medidas tomadas con los medidores individuales y con el medidor NILM tienen distinta frecuencia de muestreo. Los medidores individuales toman una medida por minuto y el medidor NILM muestrea a una tasa de 44,1kHz (estándar de las tarjetas de sonido). En el programa adquisidor de datos se procesan luego de adquirirlos y se guardan con una frecuencia de muestreo de 14kHz.

4.4.4. Formato de los datos y estructura de carpetas

Los datos de cada casa se organizan en carpetas separadas, como se ilustra en la Figura 4.4. En el directorio raíz también se incluye un archivo de calibración para las medidas de alta frecuencia.

Dentro de la carpeta de archivos flac de cada casa, se tienen todas las medidas de voltaje y corriente del medidor NILM, siendo uno de los canales del archivo el voltaje y el otro la corriente. Las señales de voltaje y corriente no tienen unidades, sino que son cuentas del adquisidor. Para convertir estas señales en los valores reales de voltaje y corriente deben multiplicarse por el valor de calibración. El nombre de cada archivo es la fecha y hora en zona horaria GMT+0 a la cual se creó ese archivo, cada medida incluida en el mismo tiene una frecuencia de 14kHz, de esta forma se puede saber exactamente a qué instante corresponde cada medida.

Los datos de los medidores individuales están en formato .csv, se tiene un archivo por día de medición y por electrodoméstico. Dentro de la carpeta “csv” hay una carpeta para cada electrodoméstico contenido los archivos de medida. Los archivos de medidas individuales tienen cuatro columnas: potencia activa, hora de la medición en GMT, identificador del dispositivo de medida y voltaje medido. El nombre de cada archivo está compuesto por el nombre del electrodoméstico y el día correspondiente a la medida.

Capítulo 4. Sistema de recolección de datos

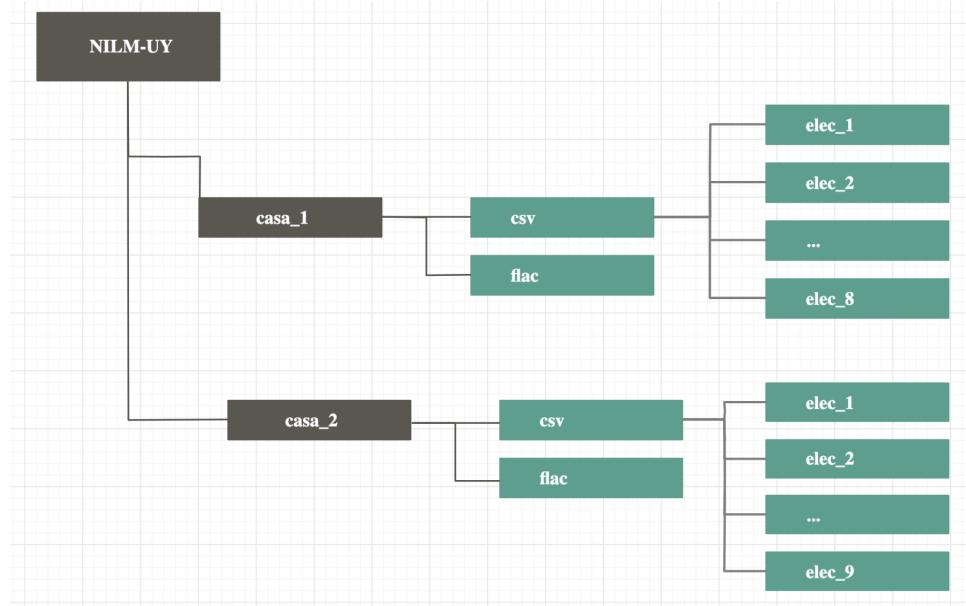


Figura 4.4: Estructura de carpetas de NILM-UY

Archivo de Calibración

Dada la calibración que se realiza (ver Capítulo 3), se obtuvieron valores para reajustar la serie de voltaje e intensidad a la escala correcta. Se cuenta con un único archivo de calibración, ya que se usó el mismo medidor no intrusivo en ambas casas, de nombre “calibration.cfg” y se encuentra dentro del directorio raíz de cada casa. El archivo esta compuesto por dos parámetros: *volts_calibration* y *amps_calibration*. Estos parámetros representan el valor por el cual deben ser multiplicada la señal de voltaje y señal de corriente respectivamente.

4.4.5. Protocolo de medida

Por ser un prototipo, el sistema de medición se instala por el grupo del proyecto en ambas casas. El proceso de instalación consiste en configurar la red local de la casa para que trabaje con un rango de direcciones IPs válidas para los medidores individuales, configurar la conexión de todos los medidores en los electrodomésticos de interés, y poner a disposición un servidor FTP de forma local, además de la instalación del sistema NILM en el tablero de la casa.

Una vez se tiene todo el sistema conectado, se verifica a través de SSH que los datos se estén generando de forma adecuada en la Raspberry y se da por concluida la instalación. Habiendo configurado una red VPN para acceder de forma segura a la red de la casa desde el exterior de la misma, de forma esporádica se realiza un monitoreo de todo el sistema, accediendo por SSH a la Raspberry y al servidor FTP.

Capítulo 5

Ingeniería de características y clasificación de electrodomésticos sobre PLAID

5.1. Introducción

5.1.1. Resumen

La mayor parte de la literatura sobre desagregación utiliza, como datos de entrada, series temporales de potencia consumida, con una frecuencia del orden de una medida cada pocos segundos. Incluso hay trabajos que recaban datos en mayor frecuencia para después procesarlos y obtener la serie de potencia consumida de baja frecuencia. Es razonable suponer que hay información contenida en las medidas de alta frecuencia a ser aprovechada. Con el objetivo de integrarla a algún algoritmo de desagregación, es útil saber cómo extraerla en forma de descriptoros o características.

Este capítulo se enmarca entonces en el área de ingeniería de características: se busca el conjunto de descriptoros de series temporales de corriente y voltaje que sea más discriminativo o informativo. Más específicamente, se busca un conjunto de descriptoros de los transitorios de una señal agregada, y otro conjunto de descriptoros de los estados de régimen de la misma señal.

En el Plug Load Appliance Identification Dataset (PLAID) [11] se pueden encontrar medidas de corriente y voltaje de los primeros segundos de funcionamiento, incluyendo el encendido, de diversos electrodomésticos. La selección de características presentada se basa fuertemente en el proceso de mejora de la *accuracy* en la clasificación de electrodomésticos sobre PLAID. Se utilizan criterios de información mutua (MI), junto con los clasificadores Random Forest (RF) y k-Nearest Neighbours (kNN), para escoger características y evaluar su desempeño, respectivamente.

5.1.2. Motivación

Uno de los principales aportes de este trabajo es la incorporación de datos de alta frecuencia para la desagregación utilizando redes neuronales. Con base en [19], se estudiará si el desempeño de los algoritmos propuestos mejora al incorporar datos de mayor resolución temporal. Se propone entonces estudiar qué características de los transitorios aportan mayor información a la identificación individual de electrodomésticos,

Capítulo 5. Ingeniería de características y clasificación de electrodomésticos sobre PLAID

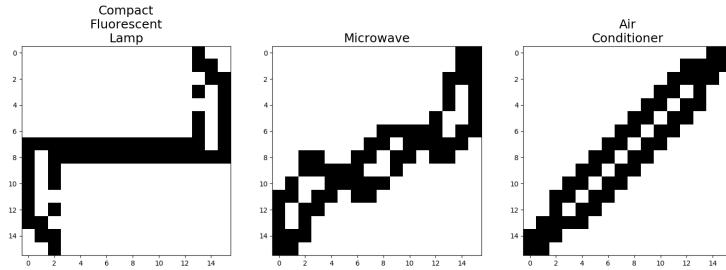


Figura 5.1: Imágenes V-I de electrodomésticos en PLAID

para luego ser utilizadas como descriptores adicionales en series de consumo agregado, con el fin de lograr la desagregación de electrodomésticos, como se verá en los algoritmos de los capítulos posteriores.

Se parte de la siguiente asunción: aquellas características que resulten inútiles para etiquetar medidas de electrodomésticos individuales, permanecerán inútiles para el problema agregado.

5.1.3. Trabajo relacionado

La identificación de electrodomésticos ha sido un problema usualmente atacado con aprendizaje supervisado. Esto es, se utilizan instancias etiquetadas sobre las que se entrena un clasificador. Este enfoque debe utilizar características diseñadas a mano, o extraerlas automáticamente con algún método (ej: *deep learning*). En general, los electrodomésticos muestran valores distintos en características que no se reducen a la **potencia activa, reactiva y aparente** que consumen. En 1992, Hart [13], describió por primera vez otras características, incluyendo **armónicos** de la señal y descriptores de los transitorios. Distintas métricas en el dominio del tiempo o la frecuencia han sido aplicadas a lo largo de la historia al problema de NILM. Otros enfoques fueron implementados en 1995 por [24], como el análisis espectral usando **transformadas de wavelet**. La **trayectoria V-I** fue usada para clasificar electrodomésticos por [39] en 2005, y por [22] en 2007. La **imagen V-I** es una característica usada principalmente para visualizar deformación en las señales. Las cargas resistivas dibujan una línea casi recta, mientras que electrodomésticos con motores muestran trayectorias más complicadas. En la Figura 5.1 se muestran imágenes V-I calculadas a partir de tres instancias de PLAID. Métricas como la **Total Harmonic Distortion** (THD) y el **factor de cresta**, comunes en procesamiento de señales, fueron usadas en 2007 por [42]. En 2011, [26], utilizó nuevamente el **factor de cresta** junto a otras características clásicas.

Nótese que las características que recurren solamente al voltaje no contribuyen de forma relevante a las de la corriente. Esto es debido a que las fluctuaciones en la señal de voltaje generadas por un electrodoméstico específico dependen directamente de la corriente consumida por él. El **corrimiento en fase** entre la corriente y el voltaje es otro importante descriptor. A su vez, hay un conjunto de características relacionado

5.1. Introducción

con la representación de la señal en frecuencia: la **distribución de energía en los armónicos**, la **spectral flatness**, la **relación entre armónicos pares e impares**, y otras relaciones más complejas entre la magnitud de los armónicos (ej: *tristimulus*). El **factor de forma**, el **factor de cresta**, y la **THD**, muchas veces están correlacionados. La **resistencia** (V_{rms}/I_{rms}) es una característica que es más robusta que la potencia a las variaciones de voltaje en la red. Su recíproco, la **admitancia** (I_{rms}/V_{rms}), ha sido también recomendado por [13] para observar los flancos que ocurren sobre esta “señal” y caracterizar a los distintos electrodomésticos según dicha huella o firma. Algunos descriptores usados en audio pueden llevarse a las series temporales usadas en NILM. El **LogAttackTime** describe el tiempo que le lleva a un electrodoméstico llegar al valor máximo de corriente desde el encendido. El **Centroide Temporal** describe el punto de balance temporal de la energía. El **Centroide Espectral** describe el punto de balance del espectro de una señal. Una métrica equivalente pero que ignora el ruido es el **Centroide Armónico Espectral**, que solo toma en cuenta los armónicos de la frecuencia fundamental. Otras características fueron desarrolladas en 2017 por [16], como la **relación entre la magnitud del armónico fundamental con la media espectral**, y el **Inrush Current Ratio**, que relaciona el RMS final e inicial de la señal de corriente en el rango de interés. Dado que algunos electrodomésticos presentan ciclos no simétricos, también se introdujo en ese trabajo el **Positive-Negative Half Cycle Ratio** y el **Max-Min Ratio**, definido siempre menor a 1. Además se presentó el **ratio entre el pico y la media de la señal**. La relación entre el pico y el valor eficaz de la corriente es denominada **Max Inrush Ratio**. La relación entre los **momentos estadísticos** de la señal también puede ser usada como descriptor. Al agregar un filtro pasa-altos se puede llegar al **centroide espectral de alta frecuencia** (S_C_hf) y al **High Frequency Spectral Flatness** (S_F_hf).

El grupo de investigación que recabó PLAID [11] también sugiere como descriptores, ya no escalares, los **histogramas de corriente y voltaje**, además de la **imagen V-I**, la **potencia activa y reactiva** y los **armónicos** de cada señal.

Otro grupo [37] integró, además de las características que se indican arriba y de una caracterización del transitorio, varios descriptores de la **imagen V-I**. Los descriptores usados sobre la imagen fueron la **asimetría**, el **número de intersecciones**, el **área** (con y sin orientación), la **distortion of mean line of VI trajectory from a straight line**, y la **pendiente del segmento central** de la trayectoria. Como descriptores del transitorio se utilizaron cuatro relacionados con la **wavelet transform**, además del **valor y ubicación del mínimo** y el **valor y ubicación del máximo** de la corriente y también de la potencia del transitorio (en el eje temporal).

A lo largo de la historia, se ha discutido cuál es la correcta forma de calcular distintos descriptores del consumo eléctrico cuando este presenta distorsión armónica no nula. Recientemente, con la publicación de [32], se estandarizó el cálculo de varias de estas características clásicas como se explica en [43]. Entre ellas se encuentra la **potencia reactiva fundamental** (Q 11), **potencia aparente fundamental** (S 11), **potencia activa fundamental** (P 11), cuya interpretación es la clásica considerando únicamente el consumo a la frecuencia de la red. Además, se incluye **potencia activa armónica** (P_H), **potencia aparente armónica** (S_H), que no son más que la potencia de los armónicos; **potencia aparente no fundamental** (S_N), **potencia no activa** (N), **poder de distorsión en corriente** (D_I), **poder de distorsión en voltaje** (D_V) y **distorsión armónica en potencia** (D_H). Por ultimo también se estandariza características que mantienen su interpretación clásica, estas son **potencia aparente total** (P), **potencia activa total** (S), **distorsión armónica en corriente y voltaje** (THD_I y THD_V).

Capítulo 5. Ingeniería de características y clasificación de electrodomésticos sobre PLAID

5.1.4. PLAID

Esta base de datos consiste de 55 casas, dentro de las cuales se midieron individualmente los primeros segundos de activación de varios electrodomésticos. Su primera versión cuenta con 1074 instancias. Cada instancia es una medida, habiendo electrodomésticos medidos en repetidas ocasiones. Cada medida contiene muestras de corriente y voltaje, y se realiza a una frecuencia de 30kHz, sobre una red cuya frecuencia es de 60Hz. Esta base de datos fue extendida en otras versiones, conteniendo la versión utilizada para este capítulo 1793 instancias.

Los autores de PLAID evaluaron los descriptores expuestos arriba. Estos vectores se usan en su artículo, individualmente y combinados, con y sin pasar por reducción de dimensionalidad mediante PCA (Principal Component Analysis), para evaluar el desempeño de los siguientes algoritmos de clasificación:

- k-Nearest Neighbours
- Gaussian Naive Bayes
- Logistic Regression Classifier
- Support Vector Machines
- Linear/Quadratic Discriminant Analysis
- Decision Tree
- Random Forest
- Adaptive Boosting

La conclusión de esa investigación es que la trayectoria V-I, junto con el histograma de corriente y voltaje, son las características más relevantes, y que el mejor clasificador, para esta base de datos y dadas las características, es Random Forest (RF). Se reporta la *accuracy* obtenida con el mejor clasificador (RF) y las mejores características (una combinación de todas), sobre la versión de la base de datos con 1074 instancias. Esta *accuracy* fue de 86,03.

Un artículo posterior [7], del mismo grupo de investigación, se ocupa de una dificultad que introduce esta base de datos: su desbalance. Las conclusiones que puedan extraerse de los valores de métricas como la *accuracy* están sujetas a este desbalance. Aún así, todos los resultados comparados están afectados por el mismo desbalance, por lo que las comparaciones son justas.

En [37] se presenta más trabajo sobre la misma base de datos, buscando maximizar la *accuracy*. Sus resultados alcanzaron una *accuracy* de 92,1, 93,2 y 92,5 con los 10, 20 y 30 descriptores más significativos respectivamente. Estos resultados son sobre la primera versión de PLAID, por lo que se cuenta con 1074 electrodomésticos. Sin embargo, se descartan los que no tienen transitorio, por lo que los resultados no son directamente comparables.

5.2. Procedimiento

5.2.1. Selección de características

La ingeniería de selección de características puede realizarse, en general, a través de tres grandes grupos de métodos: *wrappers*, filtros, y métodos embebidos. El primero se basa en reportar desempeño, sobre un clasificador particular, para un subconjunto de las características consideradas. Por búsqueda exhaustiva, u otro mecanismo, puede encontrarse el subconjunto de características óptimo. El enfoque de filtros no se basa en un clasificador particular, sino que usa algún tipo de métrica distinta al error de

5.2. Procedimiento

clasificación, siendo un ejemplo de este enfoque la correlación entre una característica y una salida. El último enfoque implica insertar en la función de costo algún término que provoque un resultado similar a hacer selección de características, como sucede al utilizar Least Absolute Shrinkage and Selection Operator (LASSO). Si bien puede hacerse selección de características usando árboles de decisión o Random Forest como clasificador, en verdad este tipo de métodos asignan automáticamente importancia a las características de entrada, respondiendo las más importantes a las primeras bifurcaciones.

En este capítulo se utiliza el segundo método, siendo la métrica seleccionada la información mutua. Además se reporta el desempeño sobre dos clasificadores: 1-Nearest Neighbours y Random Forest. Se sabe que la utilización de todas y cada una las características podría llevar a desempeños inferiores. En particular, se podría sufrir la maldición de la dimensionalidad, que implica la reducción de la densidad de los datos debido al aumento del volumen del espacio de características.

Las características utilizadas por los experimentos en este capítulo no abarcan todas las descritas anteriormente, pero sí una gran mayoría. En particular se analizan las siguientes características escalares:

- Para régimen y transitorio
 - form factor
 - crest factor
 - pos neg ratio
 - max min ratio
 - moments (5)
 - Potencia aparente (S)
 - Potencia aparente fundamental (S 11)
 - Potencia aparente armónica (S H)
 - Potencia aparente no fundamental (S N)
 - Potencia activa (P)
 - Potencia activa fundamental (P11)
 - Potencia activa armónica (H)
 - Potencia reactiva armónica (Q11)
 - Poder de distorsión en corriente (D I)
 - Poder de distorsión en voltaje (D V)
 - Distorsión armónica en potencia (D H)
 - Potencia no activa (N)
 - Distorsión armónica en voltaje (THD V)
 - Distorsión armónica en corriente (THD I)
 - Desfasaje de la frecuencia fundamental (F F)
 - Relación entre armónicos pares e impares de I (R I P)
 - Spectral Flat de I (S F)
 - Spectral centroid de I (S C)
 - Fundamental/media espectral de I (R F M)
 - Spectral Flat HighFreq (S F hf)
 - Spectral centroid HighFreq (S C hf)
- Solamente régimen
 - resistance
 - admittance
- Solamente transitorio

Capítulo 5. Ingeniería de características y clasificación de electrodomésticos sobre PLAID

- temporal centroid
- duration
- log attack time
- polynomial (3th grade)

Notar que en este Proyecto se introducen nuevos descriptores de transitorios, dados por los coeficientes del polinomio de tercer grado que mejor ajustan a la curva de corriente. Se experimentó con órdenes más grandes y más chicos, y se concluyó que el tercer grado proporcionaba un compromiso conveniente. Dichos experimentos también permitieron notar el comportamiento del RF frente a las características de entrada, analizado en la Sección 5.3.1.

Además se utiliza la imagen VI, siendo el valor de cada píxel una característica. Debido a que hay características que no varían (los valores de los píxeles de las esquinas, por ejemplo), se realizó un filtrado de características mediante un criterio de varianza. Aquellas características con varianza nula fueron descartadas. No se utilizaron descriptores sobre la imagen VI ni transformadas de wavelet para mantener la simpleza en la implementación.

5.2.2. Información mutua

La información mutua entre dos variables aleatorias X e Y se define como:

$$I(X; Y) = D_{\text{KL}}(P_{(X,Y)} \| P_X \otimes P_Y) \quad (5.1)$$

donde D_{KL} es la divergencia de Kullback-Leibler. La divergencia de Kullback-Leibler mide la similitud entre dos distribuciones de probabilidad. La información mutua es, entonces, una medida de cuán distinta es la distribución conjunta $P_{(X,Y)}$ del producto de las distribuciones marginales. Con variables independientes, la distribución $P_{(X,Y)}$ es exactamente igual a $P_X \otimes P_Y$, y la información mutua es cero. En el problema de clasificación se tienen observaciones de las variables X e Y , siendo x una característica específica, y siendo y la etiqueta correspondiente. Intuitivamente se está calculando cuán dependiente es la etiqueta (tipo de electrodoméstico) de una característica específica (como la potencia). Para realizar este cálculo se utiliza una implementación del método presentado en [35], que estima densidades de probabilidad con un algoritmo que recurre al análisis de los vecinos de cada punto. Una ventaja de la información mutua frente a la correlación es que la información mutua revela cualquier tipo de dependencia, mientras que la correlación solamente revela dependencias lineales. Una desventaja del método es que no modela, como fue usado, dependencias vectoriales: dos características que separadas no son útiles pueden ser muy útiles en conjunto. Este caso no es contemplado. Es por esto que también se utiliza un *wrapper*: RF o 1-NN.

5.2.3. Características sobre series agregadas

La asunción es que las características inútiles en este problema (desagregado) permanecerán inútiles en el problema agregado. No así a la inversa. Hay características que probablemente tengan una varianza desmedida al agregar otros electrodomésticos. Es una incógnita, por ejemplo, hasta qué punto la representación en frecuencia de una señal es útil cuando hay otras señales superpuestas (y por ende, espectros superpuestos). Se debe tener en cuenta, al usar características del transitorio, cuáles pueden ver su utilidad destruida en el problema agregado. Una de las ideas interesantes a considerar es que la información de una activación o estado es útil cuando se compara con la del estado anterior y no por sí sola. El camino de intentar clasificar electrodomésticos y no estados, es ineludible: los posibles estados serán tantos como la cantidad de combinaciones de electrodomésticos posibles (siempre que no se utilice información a priori).

5.2. Procedimiento

La utilidad de las características al pasar de la identificación individual a la agregada se mantendrá cuando las características sean una función lineal sobre las series. Por ejemplo, si se toma a la potencia como característica, como se cumple que la potencia de la suma es la suma de las potencias individuales, su utilidad se mantendrá ya que la diferencia entre un estado y otro sería igual al cambio correspondiente a la adición de (al menos) un electrodoméstico. Esto no quita que características no lineales, como el defasaje, sean de utilidad. Los algoritmos a ser utilizados en el problema de desagregación son capaces de tratar con este tipo de dependencias. De hecho, la inclusión de características no lineales colabora con el desempeño, como se verá en el Capítulo 9.

La evaluación se realiza sobre las características del transitorio, y sobre las características de régimen, por separado, para facilitar la interpretabilidad de los resultados. Los resultados de información mutua son independientes a esta separación, no así los resultados de los clasificadores. Se reporta el desempeño utilizando solamente las características del transitorio, solamente las del régimen (con y sin la inclusión de la imagen VI), y el desempeño al juntar todas las características (con y sin imagen VI). También se modifica el número de árboles (de 100 a 300). No se calcula la información mutua correspondiente a cada píxel de la imagen.

5.2.4. Extracción del transitorio

Es de interés separar el transitorio de la señal en régimen para cada instancia debido a la naturaleza del problema agregado. La discusión sobre la posibilidad de desagregación es compleja: el transitorio aislado ocasionado por el encendido del electrodoméstico A no tiene por qué ser igual al medido en la señal agregada. La señal de corriente del electrodoméstico B podría o no ser obtenida como la diferencia entre las señales de corriente previa y posterior a su encendido. Lo que sí puede afirmarse es que el comportamiento del régimen y del transitorio de una señal agregada no se tratará, en general, como un continuo transitorio+régimen, como se ha hecho sobre PLAID en algunos trabajos. He aquí la motivación de evaluar desempeños sobre las medidas de transitorio y de régimen por separado, sin perjuicio de una evaluación conjunta.

El transitorio se ve a simple vista, pero no hay una definición exacta del mismo, más que la intuición de que es el período entre (*steady states*). Se creó una función que lo extraía, considerando ciertos puntos. Primero, que el transitorio se encuentra entre el régimen y el estado inicial (apagado). Se asume que el estado inicial es de corriente nula. Por otro lado, se asume también que los últimos ciclos de la serie corresponden al comportamiento en régimen. Si bien estas afirmaciones no son ciertas en general, se cumplen en la mayor parte de las instancias. Las laptops de la base de datos, en particular, tienen un comportamiento que difícilmente podría catalogarse de régimen. Además, hay instancias dentro de la base de datos en donde las medidas no contienen al encendido del electrodoméstico.

La construcción de la función que extrae los transitorios de las instancias y la elección de los parámetros fue *ad-hoc*. Dada una serie temporal de corriente de entrada, se calcula, para cada instante t , el valor $\text{RMS}(t)$, que es el valor RMS de la corriente $i(t)$ con $t \in [t - 500, t + 500]$. Por otro lado, se calcula el valor en régimen (RMS_∞) como el RMS calculado sobre el último 10% de las observaciones de corriente. Se define un entorno B_0 alrededor del cero con radio $0,4\text{RMS}_\infty$. A su vez, se define otro entorno B_∞ alrededor de RMS_∞ de radio $0,15\text{RMS}_\infty$. El transitorio se define como $i(t)$ con $t \in [t_i, t_f]$. El índice t_i se calcula como el último índice t para el cual $i(t) \in B_0 \quad \forall t \in (t - n, t]$. Análogamente, t_f se calcula como el primer índice t para el cual $i(t) \in B_\infty \quad \forall t \in [t, t + n]$. Se utiliza $n = 1000$ y se agrega al transitorio un margen de 1000 observaciones hacia adelante y 1000 hacia atrás (si es posible).

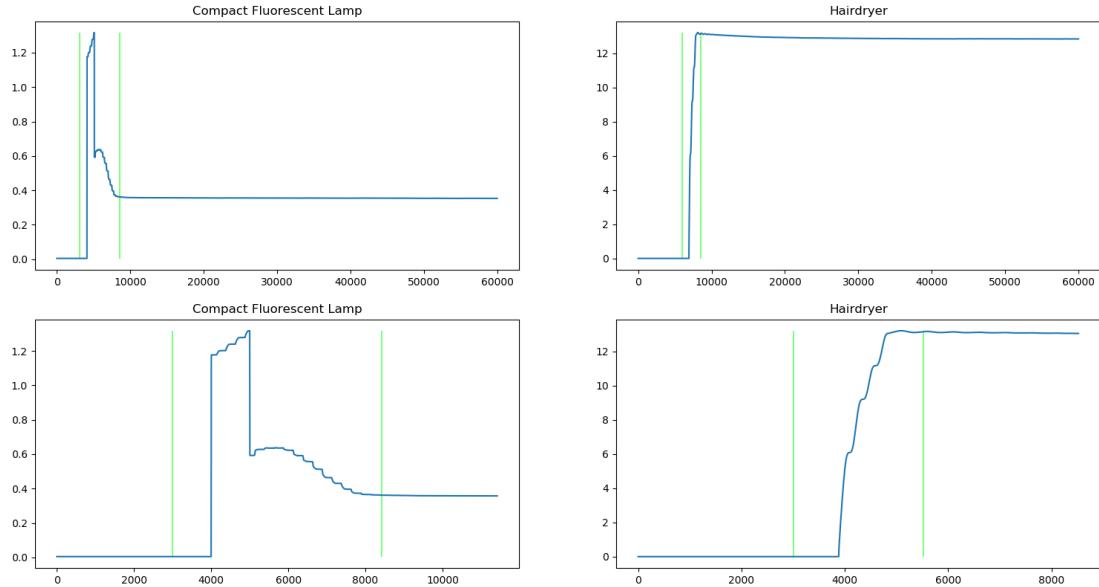


Figura 5.2: Extracción del transitorio de la medida de corriente de dos electrodomésticos. A la izquierda, lámpara fluorescente. A la derecha, un secador de pelo. Arriba, la corriente RMS completa, y debajo, un zoom sobre el transitorio extraído.

Hay instancias para las cuales no hay transitorio y la función detecta correctamente secuencias de largo $1000+1000 = 2000$. Estas instancias no se consideran para el cálculo de descriptores. También se encuentra que en ciertas instancias el estado inicial no es cero, o hay mucha variabilidad en el valor RMS de la corriente, por lo que se excluyen también los transitorios que ocupan más que el 70 % de la señal. Hay que destacar que en las instancias donde no se detecta transitorio, los descriptores del transitorio tienen valores NaN. Para pasarle estos datos al clasificador, se rellenan los valores con la media de las entradas que sí tienen transitorio del conjunto de entrenamiento.

5.3. Resultados y conclusiones

5.3.1. Resultados

Los resultados presentados aquí corresponden al valor de *accuracy* (acc), es decir, la proporción de instancias que fueron correctamente asignadas. La evaluación se hizo mediante validación cruzada (CV) sobre las 55 casas que conforman la base de datos (54 a 1). También se reporta la importancia de cada característica, según el clasificador RF y según el criterio de información mutua. Cabe destacar que la acc se calcula en repetidas ocasiones, estimando su media y su desviación estándar en cada caso, para tomar en cuenta la aleatoriedad inherente al RF. Cada corrida se repite 5 veces. El número de árboles del RF se fijó en $N = 100$ en primera instancia, y en $N = 300$ después.

Durante los experimentos se notó que el algoritmo de Random Forest es sensible

5.3. Resultados y conclusiones

al desbalance de características. Cuando $\sqrt{N} < f$, es un hecho que en cada árbol hay características que no se toman en cuenta (siendo f el número de características). Este comportamiento fue detectado al ver una mejora en el desempeño al aumentar el grado del polinomio que describe al transitorio, cuando se tenía entre el vector de características todos los píxeles de la imagen V-I. También se vió una modificación en el desempeño al agregar al vector de características los coeficientes del polinomio en forma repetida. La conclusión es que al aumentar el número de veces que aparecen ciertas características en el vector de entrada se logra tener impacto en el desempeño, ya que es menos probable que haya características que queden por fuera del entrenamiento de los árboles del clasificador. Sin ir en contra de lo anterior, se nota que al realizar ajustes polinómicos de órdenes distintos hay mejores desempeños que los obtenidos al simplemente repetir el mismo polinomio. Esto implica que hay cierta diferencia en la información incluida en cada coeficiente polinómico.

Los desempeños se muestran en forma de matrices de confusión en las figuras 5.6 5.5 5.4 5.3. El desempeño del RF sobre características de transitorio se presenta en la Figura 5.3.

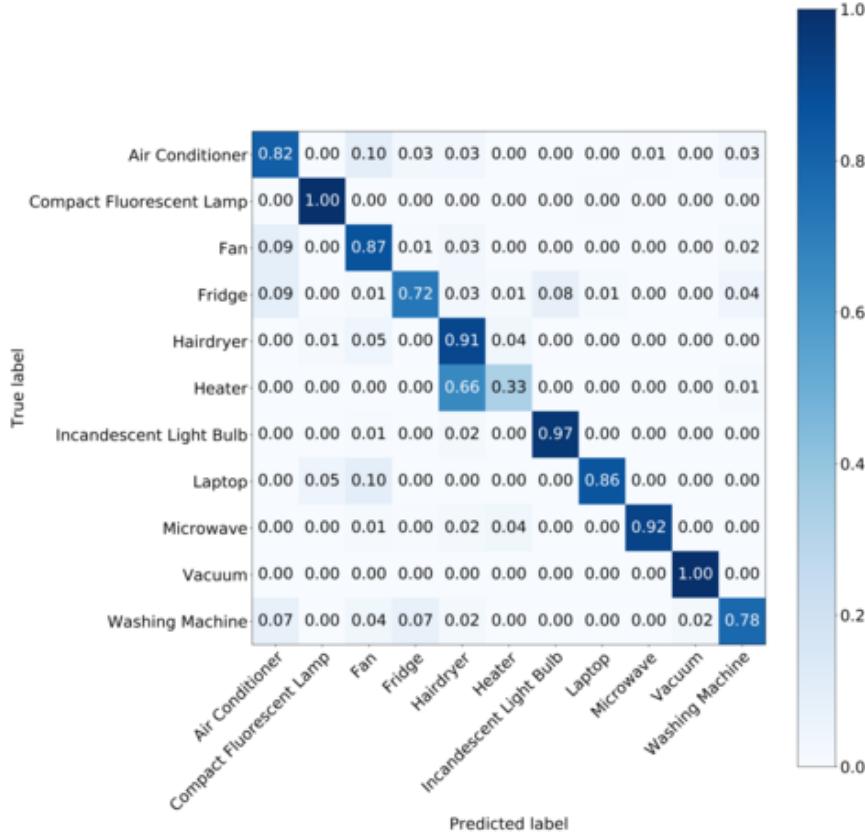


Figura 5.3: Matriz de confusión con características del transitorio

Esta primera matriz de confusión muestra un desempeño razonable para varios elec-

Capítulo 5. Ingeniería de características y clasificación de electrodomésticos sobre PLAID

electrodomésticos, aunque hay una confusión grande entre el Heater (estufa) y el Hairdryer (secador de pelo).

En la Figura 5.4 se muestra la matriz de confusión obtenida al utilizar características de régimen. Nuevamente se observa un comportamiento similar en los electrodomésticos, con pequeñas diferencias.

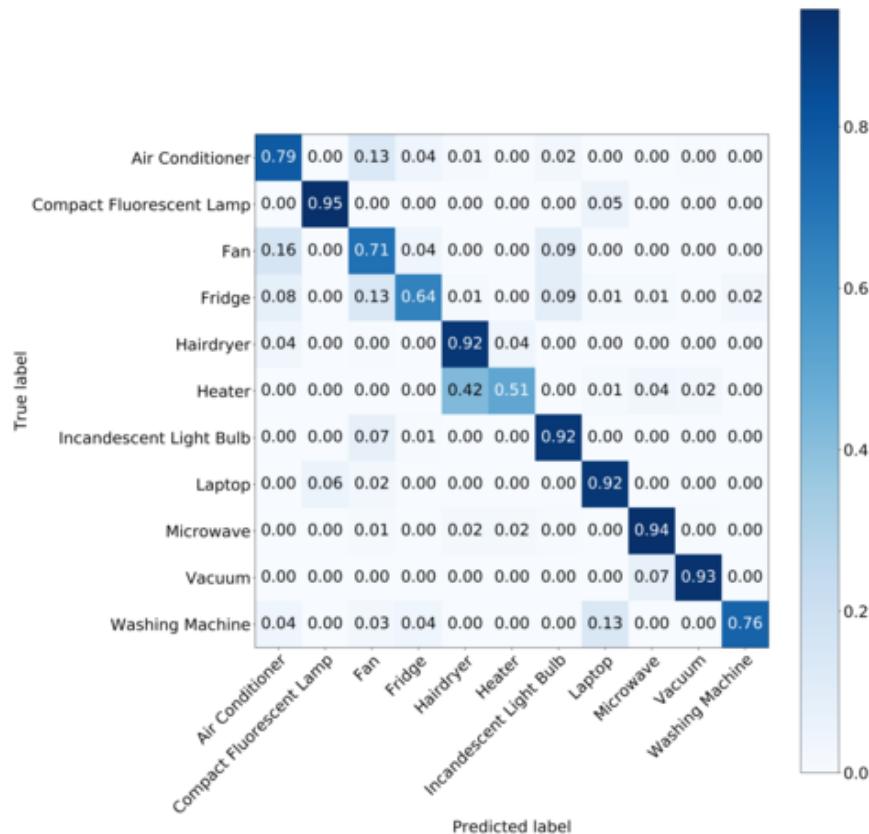


Figura 5.4: Matriz de confusión con características de régimen

En la Figura 5.5 se observa la matriz obtenida al utilizar los descriptores de régimen junto con los píxeles de la imagen VI. Utilizar los píxeles de la imagen provoca un deterioro apenas perceptible.

5.3. Resultados y conclusiones

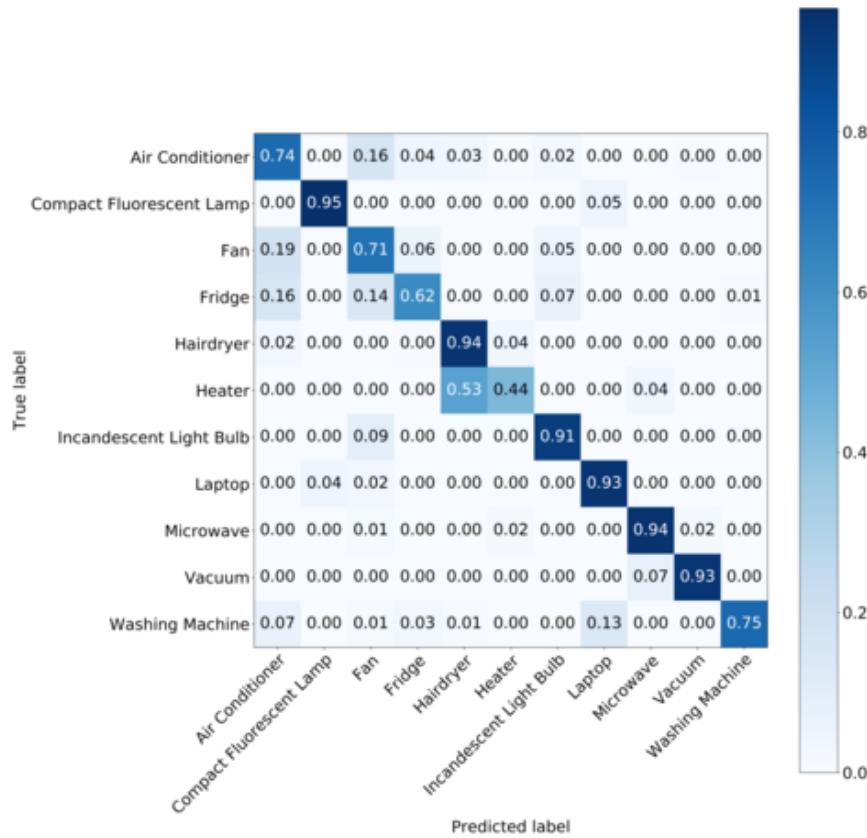


Figura 5.5: Régimen + VI

Por último se muestra en la Figura 5.6 el desempeño utilizando todas las características. Este desempeño es mejor que en cualquiera de los casos anteriormente analizados.

Capítulo 5. Ingeniería de características y clasificación de electrodomésticos sobre PLAID

Instancias	1074		1793	
Características / Clasificador	KNN	RF	KNN	RF
Transitorio	61.70	88.68 ± 0.17	59.35	87.06 ± 0.06
Régimen	75.88	87.24 ± 0.28	66.76	84.23 ± 0.25
Régimen+Transitorio	-	91.47 ± 0.09	-	88.33 ± 0.25
Régimen+VI	75.97	86.67 ± 0.49	66.82	84.14 ± 0.43
Todas	-	92.79 ± 0.13	-	89.08 ± 0.38

Tabla 5.1: *Accuracies con N = 300*

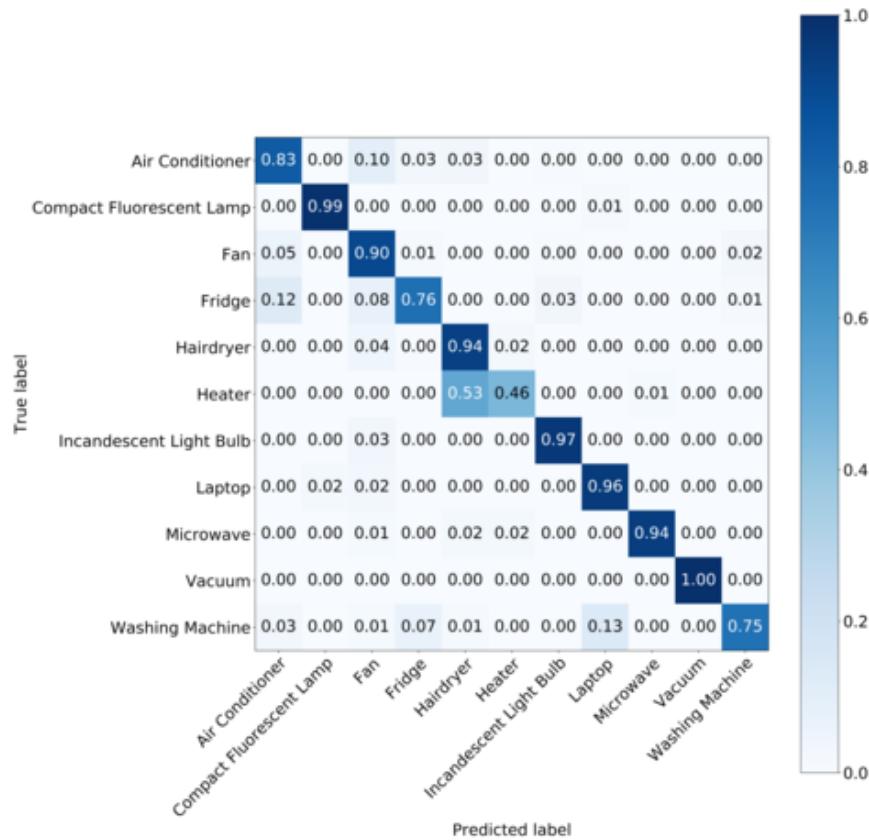


Figura 5.6: Todas las características

Por último, la accuracy y su desviación estándar para cada caso se presentan en la Tabla 5.1.

5.3.2. Características a ser utilizadas para desagregación

La información mutua y la importancia dada por RF a las características del régimen se presenta en la Figura 5.7, y las del transitorio en la Figura 5.8.

5.3. Resultados y conclusiones

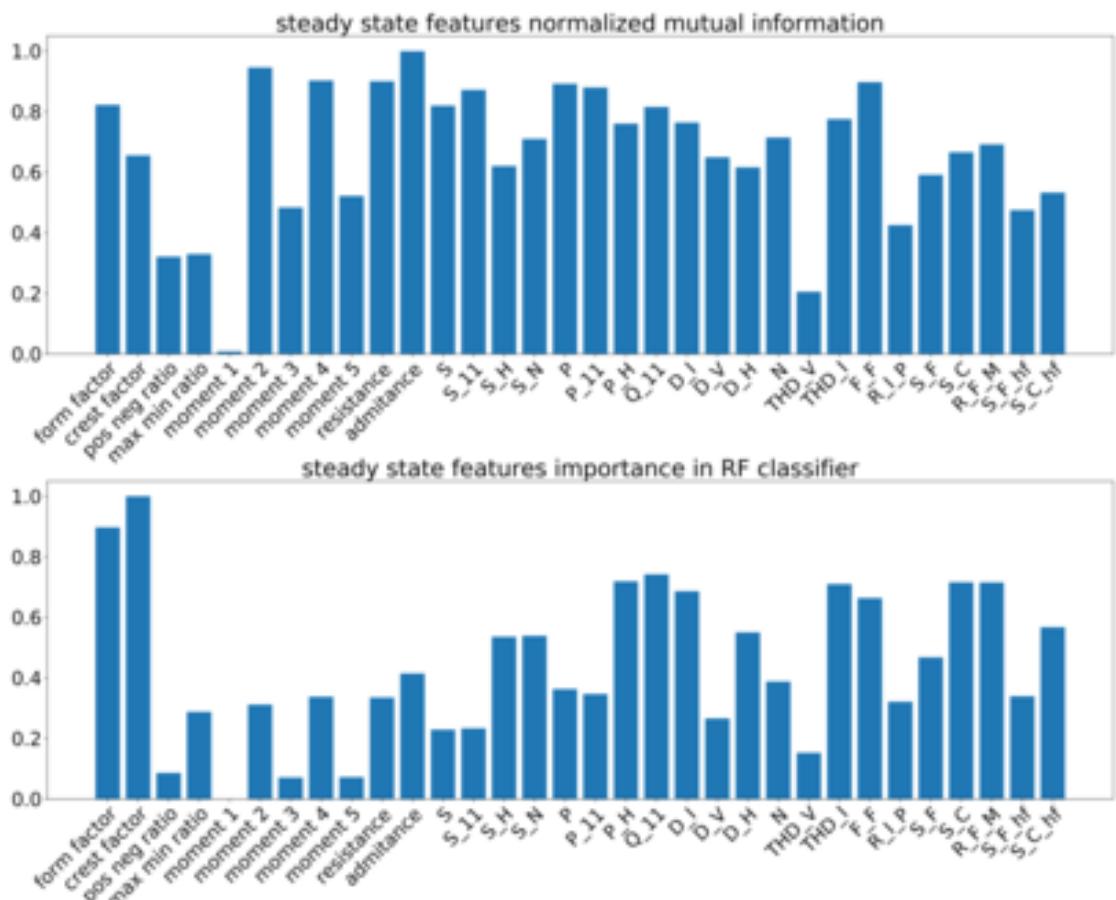


Figura 5.7: Relevancia de las características de régimen

Capítulo 5. Ingeniería de características y clasificación de electrodomésticos sobre PLAID

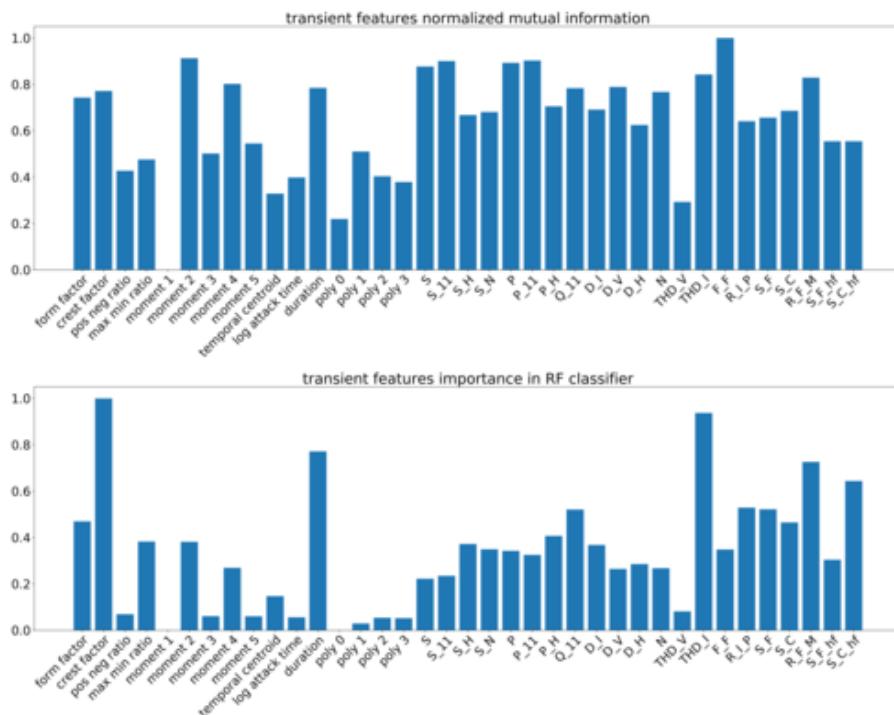


Figura 5.8: Relevancia de las características del transitorio

Como se verá en capítulos posteriores, se requerirá de series multivariadas que contengan información en forma de descriptores además de la potencia agregada de un hogar para servir de entrada a un algoritmo y obtener la serie desagregada de la potencia consumida por un electrodoméstico individual. En particular, como algoritmo se utilizarán redes neuronales e interesa estudiar el impacto de agregar características que pueden obtenerse si la señal fue muestreada a suficientemente alta frecuencia.

Estas series de descriptores, por motivos que se verán más adelante, tendrán un período de 6 segundos entre cada muestra. Para obtener cada una de las series de descriptores de alta frecuencia, se utiliza una ventana deslizante con *STRIDE* apropiado (6s) sobre la serie original (de potencia agregada de alta frecuencia). Las características de una ventana se asignan al tiempo correspondiente al punto medio de la ventana.

Ya que el número de parámetros de la red neuronal aumenta bastante con más características (lo cual no es bueno), se eligen las mismas teniendo en cuenta ciertos criterios excluyentes:

1. que haya funcionado bien con Random Forests (RF)
 2. que la Mutual Information (MI) sea suficientemente alta
 3. que tenga sentido calcularla en ventana deslizante
 4. que tenga sentido usarla en la serie agregada
 - y otros solamente deseables:
 5. que tenga poca redundancia con las demás
1. Las que cumplen haber funcionado bien con RF (top 10) son: factor de forma, factor de cresta, D_I, S_C, R_F_M, THD_I, Q_11, F_F, S_C_hf, S_N, para 1074

5.3. Resultados y conclusiones

muestras. Para +1700 muestras, las características son: factor de forma, factor de cresta, D_I, S_C, R_F_M, THD_I, Q_11, F_F, S_C_hf, P_H. Se toman en cuenta aquellas características que quedan en el top 10 en ambos casos.

2. Las que tienen mayor MI (top 10): moment 2, moment 4, resistance, admittance, S, S_11, P, P_11, F_F, factor de forma.

Es por esto que como candidatos a descriptores resultan: **factor de forma (de la corriente) y desfasaje en la frecuencia fundamental (F_F)**. Según los restantes criterios definidos, tiene sentido calcularlos en ventana móvil porque son descriptores válidos de régimen y transitorio, y tiene sentido usarlos en la serie agregada. Luego, también cumplen el requerimiento deseable de que la redundancia esperada no sea alta.

Finalmente, a estos dos descriptores también se le agrega la propia señal de potencia activa agregada. Esta serie multivariada formada por tres componentes, será la utilizada en capítulos posteriores para la tarea de desagregación.

5.3.3. Conclusión

No todos los píxeles de la imagen VI son útiles, y el momento 1, junto con los momentos 3 y 5, y la THD del voltaje, son características innecesarias. Esto es esperable: píxeles que no varían no aportan información, los momentos impares dan cero frente a señales que presentan simetría respecto al signo, y el voltaje se mantiene relativamente constante. Tanto los descriptores del régimen como los del transitorio funcionan bien por separado. Hay redundancia entre algunas características (resistencia y admittance, factor de forma y factor de cresta) y combinar características mejora el desempeño. Los desempeños obtenidos son comparables con los de las publicaciones analizadas.

Las características que no fueron estudiadas en este capítulo no son despreciables. Descriptores de la imagen VI cumplirían el mismo rol que los demás descriptores de las series temporales, y tienen gran potencialidad (área, forma). La transformada de wavelet, si bien no justificada para la señal en régimen (porque es periódica), es más adecuada que la transformada de Fourier para el transitorio. Los descriptores sobre la misma también tienen gran potencialidad.

En general, el proceso de extracción de características, sin el uso de *deep learning*, se basa en encontrar descriptores escalares de las series, o encontrar descriptores vectoriales y aplicar funciones sobre ellos que resulten en descriptores escalares. En [6] se utilizan redes convolucionales para interpretar las imágenes VI. Por otro lado, varias de las características se calculan solamente usando la señal de la corriente, mientras que hay otras que utilizan ambas series como entrada.

Quizás contraintuitivo, los electrodomésticos con peor desempeño son aquellos que más consumen. Futuras líneas de investigación deberían centrarse en revertir este fenómeno.

Capítulo 5. Ingeniería de características y clasificación de electrodomésticos sobre
PLAID

Capítulo 6

Base de datos UK-DALE

6.1. Resumen

Existen diversas bases de datos públicas de NILM de libre acceso, que posibilitan el desarrollo e investigación en el área. Se decidió utilizar la base de datos UK-DALE [18] tanto para el desarrollo como para la evaluación de algoritmos de desagregación. Este capítulo presenta esta base, que es parte fundamental del presente Proyecto.

La base de datos comprende medidas de baja y de alta frecuencia, siendo 0,17 Hz y 16kHz respectivamente. La obtención de esta base de datos implicó un trabajo de más de cuatro años en el marco del doctorado de su autor.

Los datos de alta frecuencia superan los 7TB, ya que consta de varios años de medidas, por lo que descargar y procesar estos datos supuso un problema en sí mismo.

En este capítulo se describen las principales características de la base de datos y se explica en forma detallada el procesamiento de los datos de alta y baja frecuencia para utilizar los algoritmos de desagregación que se explicarán en 7.

6.2. Elección de la base de datos

La elección de esta base de datos se justifica con que recaba medidas sobre hogares del Reino Unido, donde se tiene una red eléctrica con el mismo voltaje y la misma frecuencia que en Uruguay. Sin perjuicio de ello se nota que hay otras bases de datos de libre acceso, siendo REDD [20], BLUED [10] y WHITED [15] algunos ejemplos. Esta base de datos es además la utilizada en [19], que como se verá en futuros capítulos, es la principal referencia de este trabajo en cuanto a algoritmos.

6.3. Descripción de los datos de baja frecuencia

La base de datos fue actualizada al menos una vez. Se trabajó con la versión extendida de la base de datos, publicada en 2017. La misma consta de mediciones sobre 5 hogares del Reino Unido (UK) a lo largo de distintos períodos temporales para cada casa. Para cada hogar, se tienen datos cada seis segundos de potencia activa consumida por electrodomésticos individuales, potencia aparente consumida por toda la casa, y para algunos hogares seleccionados, también se cuenta con medidas de potencia activa de toda la casa cada un segundo.

Capítulo 6. Base de datos UK-DALE

6.3.1. Potencia aparente

En esta subsección se presenta en qué consisten los datos de potencia aparente y cómo son definidos según el autor de la base de datos. Las medidas de la serie agregada de potencia provienen de medidas de corriente con una pinza como la descrita en el Apéndice B. Como se indica en el artículo de referencia [18], estos medidores no tienen forma de medir el voltaje, por lo que se obtiene la potencia aparente simplemente escalando las medidas de corriente por el valor de voltaje establecido por el proveedor de energía (230 V). Esto puede, sin embargo, muchas veces presentar errores, ya que la normativa inglesa admite tolerancias de +10 % y -6 % en los valores de voltaje [18], introduciendo errores en las medidas de potencia aparente. La potencia aparente viene dada entonces por:

$$S = I_{rms} \times V_{rms} \quad (6.1)$$

Donde I_{rms} es la lectura de la pinza de corriente y V_{rms} es una constante igual a 230V. En definitiva, las medidas de potencia aparente no son más que medidas de corriente escaladas por un factor.

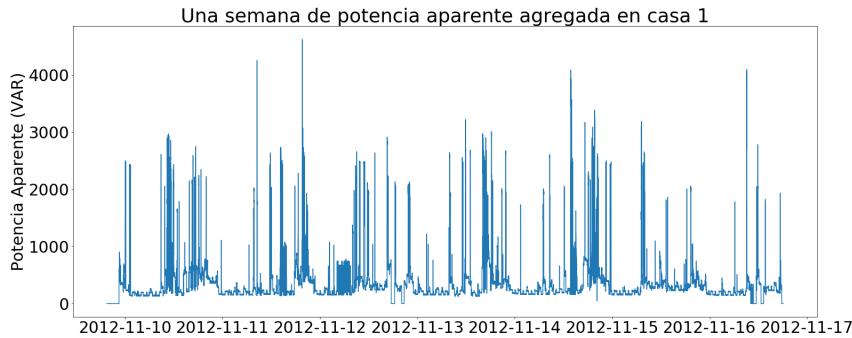


Figura 6.1: Potencia Aparente de la casa 1 durante 1 semana

6.3.2. Rango temporal de las medidas

Si bien la base de datos cuenta con medidas de cinco casas, la disponibilidad de datos para cada una de las casas y para cada tipo de dato (potencia activa, potencia aparente, etc) es muy dispar entre sí.

A continuación se presenta una tabla con el rango temporal de cada tipo de dato para cada casa en la base de datos.

Número de casa	Potencia Activa Total (1s)	Potencia Aparente Total (6s)
1	03/17/2013 al 04/26/2017	11/09/2012 al 04/26/2017
2	04/16/2013 al 10/10/2013	02/17/2013 al 10/10/2013
3	No hay datos	02/27/2013 al 04/08/2013
4	No hay datos	03/09/2013 al 10/01/2013
5	06/29/2014 al 11/13/2014	06/29/2014 al 11/13/2014

Tabla 6.1: Disponibilidad de datos agregados de baja frecuencia en UK-DALE.

6.4. Procesamiento de los datos de baja frecuencia

Por otro lado, los datos de potencia activa individuales para cada electrodoméstico abarcan rangos temporales muy similares a los de la potencia aparente total, del orden de minutos de diferencia entre unos y otros. Son estos datos las “etiquetas” o el “*ground truth*”.

6.3.3. Electrodomésticos medidos

La lista de electrodomésticos medidos varía bastante para cada casa, siendo la casa 1 la que tiene mayor cantidad de distintos electrodomésticos. Para el desarrollo de algoritmos, se escogieron los mismos electrodomésticos que en [19]. Estos electrodomésticos cumplen con dos características importantes para poder entrenar y evaluar algoritmos: tienen medidas en la mayoría de las casas, y son electrodomésticos que consumen potencias considerables.

Electrodoméstico	Casas donde hay datos
Jarra eléctrica	1 2 3 4 5
Heladera	1 2 4 5
Lavarropas	1 2 5
Microondas	1 2 5
Lavavajillas	1 2 5

Tabla 6.2: Electrodomésticos seleccionados para entrenar algoritmos.

6.4. Procesamiento de los datos de baja frecuencia

Para la desagregación, se eligió el mismo conjunto de electrodomésticos que el reportado en [19]: heladera, lavarropas, lavavajillas, jarra eléctrica y microondas. Como se verá en el Capítulo 7, para la desagregación se utilizaron redes neuronales. Como entrada a las redes, se necesitan vectores de largo fijo. Los vectores son ventanas de la serie temporal. Como salida se tiene una ventana del mismo largo, pero sobre la serie de potencia desagregada. El largo de ventana varía de acuerdo al electrodoméstico.

En los datos de baja frecuencia y para todos los aparatos, se optó por utilizar como entrada la serie de potencia aparente agregada, en lugar de la serie de potencia activa. Se decidió utilizar la potencia aparente como entrada, ya que se cuenta con medidas de todas las casas, y además el rango temporal de las medidas es mayor que el de las medidas de potencia activa, lo que implica mayor cantidad de datos.

6.4.1. Alineamiento temporal

Tanto las series de entrada a la red como las desagregadas tienen un período de 6 segundos entre cada dato. Sin embargo, existe el detalle de que las medidas de potencia agregada no están alineadas temporalmente con las de potencia desagregada (si bien se tienen medidas con la misma cadencia, no coinciden en los instantes de muestreo).

Para resolver este inconveniente se optó por correr temporalmente todas las series de cada electrodoméstico (de cada casa), de forma tal que queden perfectamente alineadas con la serie agregada de cada casa. Dado el período de muestreo de 6 segundos, el corrimiento de cada medida siempre será menor a 3 segundos, por lo que esto no supone una alteración significativa de los datos.

Capítulo 6. Base de datos UK-DALE

6.4.2. Datos faltantes

En todas las series temporales de la base de datos se tienen datos faltantes. Se entiende por datos faltantes la aparición de medidas sucesivas cuyos *timestamps* difieren más que el período de muestreo definido para cada serie. Este fenómeno aparece de forma más común en la serie de potencia activa total.

A la hora de utilizar estos datos para entrenar y evaluar algoritmos fue necesario llenar los huecos en las series temporales. Para esto se utilizó el mismo criterio que en [19]: si los huecos duran menos de 3 minutos, se completan los datos propagando hacia adelante el último dato disponible. En caso contrario, se rellena con ceros. La justificación de lo anterior presentada en [19] es que si el *gap* es pequeño se asume que el medidor perdió conectividad, si es mayor a 3 minutos entonces se asume que el medidor fue apagado junto con el electrodoméstico.

6.4.3. Activaciones y no activaciones

Si se consideran las series temporales de cada electrodoméstico, se puede ver que la mayor parte del tiempo no están consumiendo potencia. Para que el entrenamiento sea balanceado, y las redes logren aprender cuándo un electrodoméstico está encendido, se pre-procesan las series temporales para obtener vectores, de la serie agregada y desagregada, que contengan una activación (un ciclo de funcionamiento del electrodoméstico) y vectores correspondientes a ventanas en las que no hay un ciclo de trabajo completo. Como las activaciones son minoritarias en la serie temporal, se calculan todas las activaciones existentes, y luego, por cada activación, se intenta incluir también una no activación. Esto da lugar a un conjunto de entrenamiento aproximadamente balanceado entre activaciones y no activaciones.

En resumen, los datos procesados para las redes neuronales asociadas a un electrodoméstico son vectores de largo fijo que pueden contener o no una activación de cada electrodoméstico en cuestión.

6.4.4. Obtención de datos de entrada y salida

El proceso para obtener los datos de entrada y salida a las redes, en el formato adecuado, implica cortar ventanas agregadas que contengan activaciones y no activaciones. Para esto es necesario determinar, a partir de la serie desagregada, cuando está presente o no una activación. También se debe determinar el largo del vector, de forma de que las activaciones calculadas se encuentren de forma completa en los vectores de entrenamiento.

Cálculo de activaciones

A la hora de calcular las activaciones de cada electrodoméstico, se utilizó una versión corregida, ya que había un *bug*, de la función *activation-series-for-chunk* de la librería *NILMTK* [17].

Esta función toma como entrada una serie de potencia de un electrodoméstico y devuelve múltiples series de potencia, correspondientes a todas las activaciones de ese aparato. La función tiene como parámetros:

- *min-off-duration*: Ignora períodos de apagado que sean menores a este parámetro
- *min-on-duration*: Ignora activaciones de duración menor que este parámetro
- *border*: Número de muestras que se le agrega antes de cada activación (padding de las activaciones)

6.4. Procesamiento de los datos de baja frecuencia

- *on-power-threshold*: Umbral de potencia para considerar que un electrodoméstico está funcionando

En pocas palabras, esta función detecta los flancos cuando se cruza el umbral de potencia y luego filtra las activaciones según los parámetros de duraciones mínimas de apagado y encendido. Las activaciones que sobreviven a ese filtrado son las retornadas.

	<i>min_off_duration</i>	<i>min_on_duration</i>	<i>on_power_threshold</i>
Jarra eléctrica	0	12	2000
Heladera	12	60	50
Lavarropas	160	1800	20
Microondas	30	12	200
Lavavajillas	1800	1800	10

Tabla 6.3: Parámetros utilizados en la función *activation-series-for-chunk*

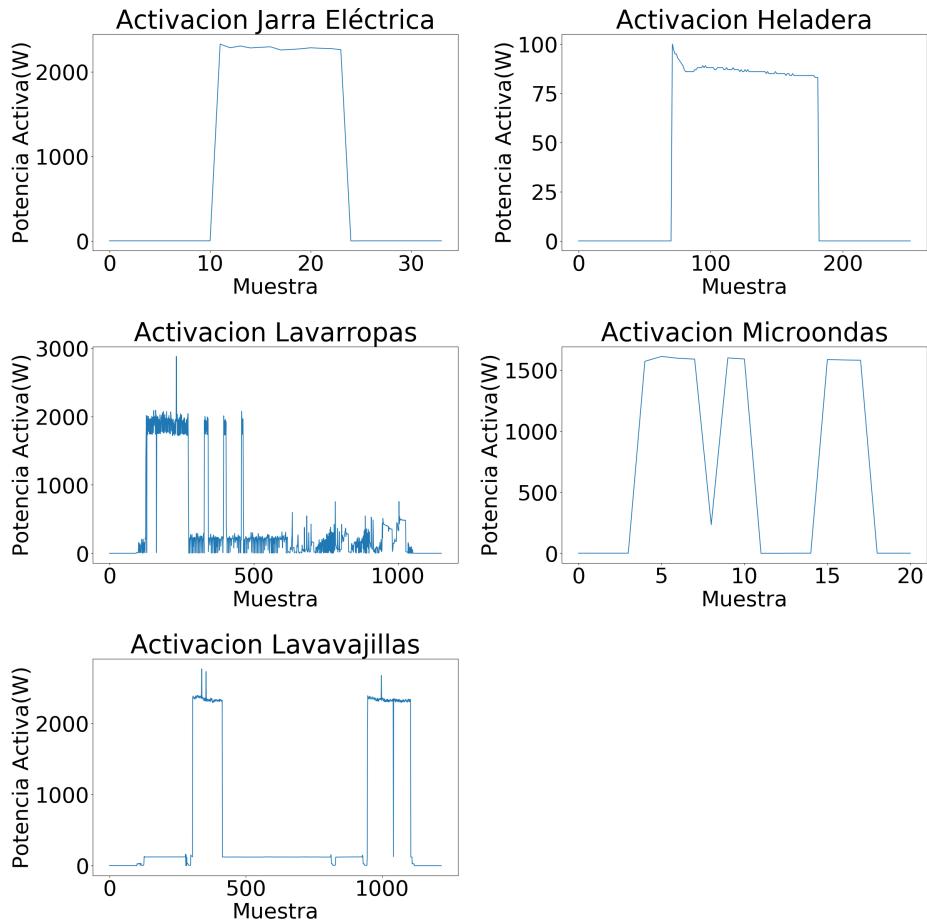


Figura 6.2: Ejemplos de activaciones (se tiene una muestra cada 6s)

Capítulo 6. Base de datos UK-DALE

Largo de ventana

Para determinar el largo de ventana a utilizar para cada electrodoméstico, inicialmente se tomaron como referencia los valores indicados en [19]. En el caso de los electrodomésticos para los cuales no se reportó el largo de ventana utilizado (microondas, lavarropas), se calcularon los histogramas de largo de activaciones (sobre el conjunto de entrenamiento) y se decidieron valores de largo de ventana que siempre incluyan completamente a todas las activaciones, a menos de *outliers*. Para los valores de referencia del *paper*, también se verificó que incluyan completamente a casi todas las activaciones. Se dejan por fuera a los *outliers* ya que si se escoge un largo de ventana muy grande se incluirían múltiples activaciones en cada ventana y se entorpece el entrenamiento.

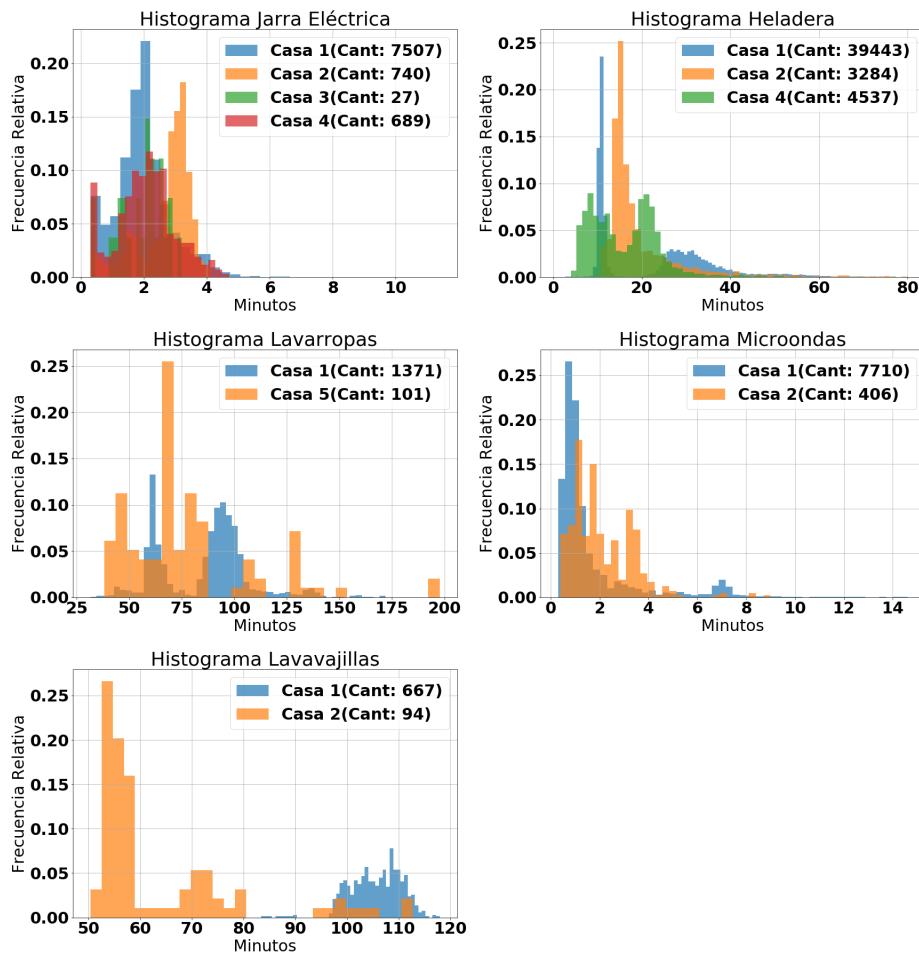


Figura 6.3: Histogramas de largo de activaciones

Sin perjuicio de lo anterior, en algunos casos puede aparecer más de una activación dentro de una ventana, o una activación parcial (no contenida completamente). Las redes se entrenaron de forma tal que para cada ventana solo reconozca la primera activación completamente contenida en la ventana. Esto se logró rellenando con ceros

6.5. Descripción de los datos de alta frecuencia

	Largo de ventana (en minutos)
Jarra eléctrica	13
Heladera	60
Lavarropas	180
Microondas	10
Lavavajillas	150

Tabla 6.4: Largos de ventana utilizados

las activaciones parciales o secundarias de cada ventana.

Una vez definidos los largos de ventana para cada electrodoméstico, los vectores de entrada y salida de las redes correspondientes a activaciones se obtienen calculando de forma aleatoria una ventana temporal de la serie agregada (univariable o multivariada) que contenga completamente a la activación.

En el trabajo de referencia se reportan solamente algunos largos de ventana. Se realizaron histogramas de duración de las activaciones de los distintos electrodomésticos, para las distintas casas. Cotejando estos histogramas de duración con los largos de ventana reportados en el trabajo de referencia se puede notar que los mismos son definidos coherentemente: el largo de ventana reportado incluye a los largos encontrados en las distintas casas. Sin embargo, para el electrodoméstico *dishwasher*, el histograma de duración muestra que las casas 1 y 2 (usadas para el entrenamiento) tienen una duración máxima de 120 minutos, mientras que la casa 5 (usada para el testing) tiene activaciones que se extienden hasta 145 minutos. La ventana definida en este caso fue de 150 minutos, lo que muestra que en la publicación, el largo de ventana es un parámetro definido *a priori*, en lugar de un parámetro que se obtiene del entrenamiento.

Cálculo de no activaciones

Para obtener los vectores de entrada y salida correspondientes a las no activaciones, se recorre todo el conjunto de activaciones y se calcula una ventana aleatoria entre activaciones consecutivas, que no incluya completamente ninguna de las dos activaciones. Puede haber casos en los que es imposible encontrar una ventana de no activación entre activaciones consecutivas ya que las mismas están demasiado cerca en el tiempo, es por esta razón que el conjunto de activaciones y no activaciones es aproximadamente balanceado y no exactamente balanceado.

6.5. Descripción de los datos de alta frecuencia

La base de datos UK-DALE, cuenta con la medición agregada a una resolución temporal muy superior a la anterior. Estas medidas son tomadas con una frecuencia de 16kHz, utilizando un medidor basado en una tarjeta de sonido [18], y consisten en medidas sin calibrar de voltaje y corriente por separado. Junto con los archivos de datos, se incluye un archivo de calibración para obtener la potencia a partir de dos archivos. Cabe destacar que el archivo de calibración de la casa 5 presenta un error: al usar las constantes de calibración en él presentes, la potencia resulta en valores que exceden los 100kW, que es demasiado consumo para cualquier combinación de electrodomésticos. Se consideró que era un error de notación y se llevaron las constantes al orden de magnitud de aquellas presentes en los demás archivos de calibración. Por otro lado, para cada versión de la base de datos hay distintos archivos de calibración, cuan-

Capítulo 6. Base de datos UK-DALE

do en verdad estos archivos son duplicados, denotando que no se consideró necesaria una re-calibración de los dispositivos entre una versión y la siguiente.

La base de datos se organiza en carpetas por versión, por casa, por año y por semana, dentro de las cuales hay archivos de formato .flac correspondientes a una hora de datos cada uno. Solamente se tienen medidas de alta frecuencia para las casas 1, 2 y 5. Estas medidas también presentan períodos con datos faltantes, archivos correspondientes a períodos mucho menores a una hora, y hasta archivos repetidos. También se encuentran algunos pocos archivos en formato .wav. En [21] se estudió el nivel de compresión de los datos disponibles en distintas bases de datos de investigación y se concluyó que esta base de datos era la que presentaba el mejor nivel de compresión. Cada archivo tiene como nombre el momento en el cual comienza la medición en formato *UNIX timestamp*. Juntando esta información con la de la frecuencia de muestreo se pueden identificar temporalmente todos los datos. Luego de estudiada la cantidad de muestras se corrobora que la frecuencia de muestreo es efectivamente 16kHz.

6.5.1. Rango temporal de las medidas

A diferencia de los datos de baja frecuencia, solo se tienen medidas de alta frecuencia para las casas 1, 2 y 5. Además el rango temporal de dichas medidas es distinto, a continuación se indica el rango temporal disponible para cada casa.

Número de casa	Potencia Activa Total (16 kHz)
1	03/15/2013 al 10/31/2016
2	04/16/2013 al 10/18/2013
3	No hay datos
4	No hay datos
5	06/29/2014 al 11/17/2014

Tabla 6.5: Disponibilidad de datos agregados de alta frecuencia en UK-DALE. Formato de fechas estadounidense.

6.6. Procesamiento de los datos de alta frecuencia

El objetivo es extraer descriptores de la serie de alta frecuencia y formar con ellos una serie multivariable de baja frecuencia (período de 6s). Esta serie de baja frecuencia se utiliza como entrada de una red neuronal como se explica en el Capítulo 7. Los descriptores son aquellos que resultaron más promisorios tras el análisis efectuado en el Capítulo 5, además de la potencia activa. El tratamiento de estos datos comienza por su descarga y termina una vez finalizado el procesamiento. El gran tamaño de la base, 7,6 TB, generó importantes desafíos tanto para su descarga como para el preprocesamiento, lo que se describe a continuación. Los datos de entrada y salida a las redes, serán ventanas de la serie multivariada que contenga activations y no activations. Estas ventanas se generan con el mismo criterio que el explicado para los datos de baja frecuencia.

6.6.1. Descriptores de alta frecuencia

Con base en las conclusiones extraídas en el Capítulo 5 para formar la serie multivariada basada en datos de alta frecuencia se escogieron: *potencia activa*, *desfasaje de la frecuencia fundamental* y *factor de forma de la corriente*. El factor de

6.6. Procesamiento de los datos de alta frecuencia

forma f se define como el ratio entre el valor RMS y el Average Rectified Value (ARV), como en la Ecuación 6.2.

6.6.2. Descarga de los datos

Web Scrapping

A la fecha, los datos están disponibles públicamente en los servidores del UK Energy Research Centre¹, ubicados en Londres, y pueden ser descargados individualmente, al *clickear*, o en tandas, mediante una aplicación que la página provee. Al ser una gran cantidad de datos, ninguna de las opciones anteriores fue viable. *Clickear* hubiera implicado demasiado tiempo, y la aplicación tenía un límite de archivos a descargar. A la extracción automática de datos de sitios web se le conoce como *web scrapping*. Hay diversos paquetes disponibles en Python que cumplen esta finalidad, entre ellos *Beautiful Soup* y *Requests*, utilizados en esta etapa. Se realizó un *script* para recorrer el directorio web de archivos e ir generando una copia local de los nombres, a ser utilizada luego. Este recorrido se realizó evitando las carpetas donde se conocía que no existían archivos de alta frecuencia. Una vez generada la copia local de nombres, se listaron los mismos según su tipo (.html, .cfg, .flac, .wav, .txt) y según la casa correspondiente. El nombre de los archivos de datos es un *UNIX timestamp* que fue convertido a fechas tradicionales y permitió mayor organización.

Automatización de la descarga

Con el fin de automatizar la descarga se optó por usar el programa por línea de comandos *wget*. Este programa cuenta con múltiples opciones, entre ellas la creación automática de directorios según la ruta, la no sobre-escribura, y la configuración de *proxy*. Además, puede tomar una lista de URLs de un archivo de texto y hacer una descarga ordenada. Dado que se utilizó la conexión a internet de la Facultad de Ingeniería, se utilizaba una ruta alternativa, reservada para fines académicos, para el tráfico, ya que se apuntaba a los servidores de un centro de investigación, como se comprobó tras un *traceroute*. Se comprobó empíricamente que la descarga de archivos mediante *wget* tenía una velocidad máxima, mayor durante la noche, que disminuía al ejecutar descargas en paralelo, en diferentes consolas. Sin embargo, la velocidad de descarga total aumentaba con el paralelismo hasta llegar a una velocidad pico de aproximadamente 10MB/s con nueve procesos. Además se utilizaron arbitrariamente los tres *proxys* disponibles: 164.73.32.10/11/12, asignando tres procesos a cada uno. Descargar en diferentes computadoras también aumentó la velocidad. Al tener estas cosas en cuenta, solo restó crear las listas de URLs que sirvieron como entrada al programa *wget*.

Creación de listas de descarga

Dado que no se contaba con un solo dispositivo capaz de almacenar todas las descargas, los datos fueron guardados en distintos discos duros externos de capacidades de 2TB o menos cada uno. Por esto, cada vez que se hacía una comprobación de los archivos guardados localmente, todos los dispositivos tenían que ser recorridos. Primero se listaron y descargaron en corto tiempo todos los archivos que no contenían medidas (archivos de calibración, readme, etc.). La creación de las listas de los archivos .flac se hacía por casa y por fecha. Para ello, se contrastaba la lista de archivos obtenida mediante *web scrapping* con los archivos almacenados localmente, filtrando además

¹[https://data.ukedc.rl.ac.uk/browse/edc/efficiency/residential/
EnergyConsumption/Domestic](https://data.ukedc.rl.ac.uk/browse/edc/efficiency/residential/EnergyConsumption/Domestic)

Capítulo 6. Base de datos UK-DALE

según el tipo. También se realizaba una comprobación de agujeros, es decir, de archivos faltantes dentro de un rango de fechas que debería estar descargado, de forma de asegurar completitud y poder bajar aquellos archivos que no se hayan descargado correctamente. Por continuidad y robustez se descargaban tandas de aproximadamente 500GB, calculadas tomando como base el peso de 200MB por .flac de una hora de duración, previa comprobación de que hubiera espacio disponible en el dispositivo de almacenamiento objetivo.

6.6.3. Procesamiento

El procesamiento implica calcular el factor de forma (Ecuación 6.2), el desfase entre la componente fundamental del voltaje y la corriente, y la potencia media en el intervalo considerado.

$$\text{factor de forma} = \frac{\sqrt{\frac{1}{T} \int_0^T i^2(t) dt}}{\frac{1}{T} \int_0^T |i(t)| dt} \quad (6.2)$$

Si bien se cuenta con la potencia aparente (y activa en algunos casos) en la base de datos original, la potencia activa se calcula nuevamente con el objetivo de posibilitar la comparación entre la base de datos y el resultado del procesamiento. Además se nota que los intervalos de tiempo no siempre coinciden (hay huecos). La serie temporal multivariable objetivo tiene un período idéntico al de la serie temporal de baja frecuencia de la base UK-DALE, es decir, 6 segundos. Por razones de eficiencia se elige calcular los descriptores no sobre todo un período $T = 6s$, sino sobre un tiempo dentro del período de 0,5s. A menos tiempo considerado en el cálculo del descriptor la varianza será mayor, bajo la asunción de un estado de régimen. Por otro lado, si el estado no es de régimen (contiene un transitorio) es probable que el cálculo sea corrompido, dando lugar a un número que no corresponde ni al estado previo al cambio ni al estado posterior. El enfoque aquí planteado es el de análisis de la información de alta frecuencia mediante descriptores del régimen. La decisión de reducir el lapso de tiempo sobre el que se calculan los descriptores a 0,5s responde al compromiso que surge de los planteos anteriores.

A la hora de procesar los datos surgen naturalmente dos enfoques. El primero de ellos implica basarse en los datos existentes e ir recorriendolos mientras se calculan los descriptores, lo cual es un enfoque “de datos hacia fechas”. El segundo es en el sentido inverso, recorrer una lista de fechas y calcular los descriptores para cada una de ellas, sería “de fechas hacia datos”. Se escoge el segundo enfoque por razones de modularidad y compatibilidad con los datos de baja frecuencia ya disponibles: dada una lista de fechas, el programa se encarga de devolver los descriptores para cada de ellas o un valor NaN si no fuera posible. En lo anterior, “fecha” es un momento con resolución de microsegundos.

Obtención de los rangos de fechas

El procesamiento se divide en dos programas, uno que se utiliza para elegir la lista de fechas, y el principal, que a partir de dicha lista calcula los descriptores. La lista de fechas se genera esencialmente consultando para una casa, cuáles son la primera y la última fecha comprendida en los archivos .flac con los que se cuenta localmente. Sobre el período de tiempo comprendido entre estas fechas, más un margen al principio y al final, se hace un muestreo con período de 6s (frecuencia de la base en baja frecuencia). A partir de esta lista de fechas, en general larga, se obtienen varias al subdividirla, con el objetivo de permitir el multiprocesamiento y establecer *checkpoints* que aporten a la robustez del algoritmo. Estas listas, los “rangos de fechas”, se guardan junto con la

6.6. Procesamiento de los datos de alta frecuencia

lista de los archivos *flac* de la casa en cuestión y con las fechas correspondientes a cada uno de esos archivos. Además, se crea una carpeta de destino para el procesamiento y se señala cuál es el archivo de calibración a ser utilizado.

Programa principal

El programa principal se diseñó en formato de pseudo-código, y luego se implementaron todas las funciones requeridas. Comienza por cargar en memoria todas las entradas previamente descritas y la ubicación de los dispositivos de almacenamiento externo. El pseudocódigo es como sigue:

```
for rango_de_fechas in rangos_de_fechas:
    inicializar tabla
    for fecha in rango_de_fechas:
        conseguir ventana de 0.5s
        if not (ventana pertenece a rango):
            cargar archivos flac cuyo rango comprende a la ventana
            calibrar archivos flac cargados
        obtener serie temporal correspondiente a ventana
        calcular descriptores
        escribir en tabla
        if (tabla esta llena):
            escribir tabla a .csv
            inicializar tabla
```

El proceso de implementar el pseudocódigo anterior tuvo que tener en cuenta varios detalles, concernientes al desempeño tanto en uso de la memoria como en velocidad de procesamiento, y a errores encontrados en la base de datos. Al hacer una comprobación del tiempo de corrida se encontró, como era de esperarse, que la gran parte del tiempo transcurría durante la carga de los flacs y el procesamiento de la serie, en ese orden. Es por esto que siempre que fuera posible, se evitaba cargar innecesariamente archivos. Por ejemplo, normalmente se tenían dos archivos en memoria al mismo momento y el segundo de ellos pasaba eventualmente a ser el primero, sin volver a cargarse en memoria. La necesidad de tener dos archivos surge de que la ventana sobre la cual se quiere hacer el cálculo no siempre está absolutamente contenida en un solo archivo. El momento de finalización de los archivos se definió teniendo en cuenta la cantidad de observaciones y su frecuencia teórica. Es por esto que fue necesario definir una tolerancia, de 1ms, para el período entre la finalización de un archivo y el comienzo del siguiente. Es muy usual que haya observaciones que se pierdan entre dos archivos consecutivos. Se asumió que no había archivos que se solapen aunque se encontraron archivos de un minuto de duración. Otras consideraciones de eficiencia surgían a la hora de cargar los archivos, ya que existen varios paquetes disponibles. No todos ellos eran útiles, sin embargo, ya que varios no eran compatibles con archivos FLAC de 24-bit como los que presenta la base. Se escogió *soundfile* debido a la comparación expuesta en el repositorio de GitHub de *torchaudio-contrib*². Por otro lado, dada la utilización de la FFT para calcular el desfasaje, se buscó una implementación rápida de la misma. La subrutina de *C* FFTW era promisoria, pero se comparó su *wrapper* de Python con *scipy.fftpack.fft*, y la última implementación resultó ser la más rápida. Se avanzó en métodos de cálculo alternativos (ya que no es necesario calcular toda la FFT para obtener el desfasaje) pero no se implementaron ya que la carga de los archivos FLAC ocupaba más del 90 % del tiempo de procesamiento. También se realizó un *profiling* de

²<https://github.com/keunwoochoi/torchaudio-contrib/issues/31#issuecomment-476183742>

Capítulo 6. Base de datos UK-DALE

la memoria utilizada buscando prevenir la saturación de la memoria RAM disponible. La memoria es administrada por el *Python memory manager* automáticamente, por lo que se usaron las herramientas incluidas en el editor *Spyder* en lugar de crear unas propias para este fin. La lista de las funciones implementadas como auxiliares al programa principal es la que sigue, ocultando todas excepto las que conciernen a los descriptores:

```
def load_flac_by_date(date, filenames, files_dates, data_locations)
def window_in_range(window, valid_range)
def date_in_range(date, valid_range)
def get_filename(date, filenames, files_dates, data_locations)
def get_series(current_flac, next_flac, window)
def get_values(series, date)
def get_duration(flac, samplerate)
def look_for_immediate_previous_date(date, files_dates)
def get_calibration_dict(full_path_to_calibration_cfg)
def calibrate_dataframe(dataframe, calibration_dict,
                       verbose=True)
def calibrate_dataframes(list_of_dataframes, calibration_dict,
                        verbose=True)
def get_window(date, width_in_seconds=6)
def load_flacs(date, filenames, files_dates, data_locations,
               tolerance=pd.to_timedelta(16/16000, unit='s'))

def fft_features(V, I, samplerate=16000, line_freq=50)
    number_of_points = len(I)
    FI = fft(I)
    FV = fft(V)

    number_of_freqs = (samplerate // 2) // line_freq
    harmonics_separation = (number_of_points // 2) // number_of_freqs
    begin_region = harmonics_separation - harmonics_separation // 2
    end_region = harmonics_separation + harmonics_separation // 2
    fundamental_index = begin_region + np.argmax(np.abs(
        FI[int(begin_region):int(end_region)]))
    F_F = np.angle(FV[fundamental_index] / FI[fundamental_index])
    P_mean = np.mean(V * I)
    return F_F, P_mean

def form_factor(current)
    ''' Computes the form factor of a time series as the ratio
    between the RMS of the series and the MAV of the series '''
    rms = np.sqrt(np.mean(current**2))
    mav = np.mean(np.abs(current))
    return rms / mav
```

6.7. Conclusión

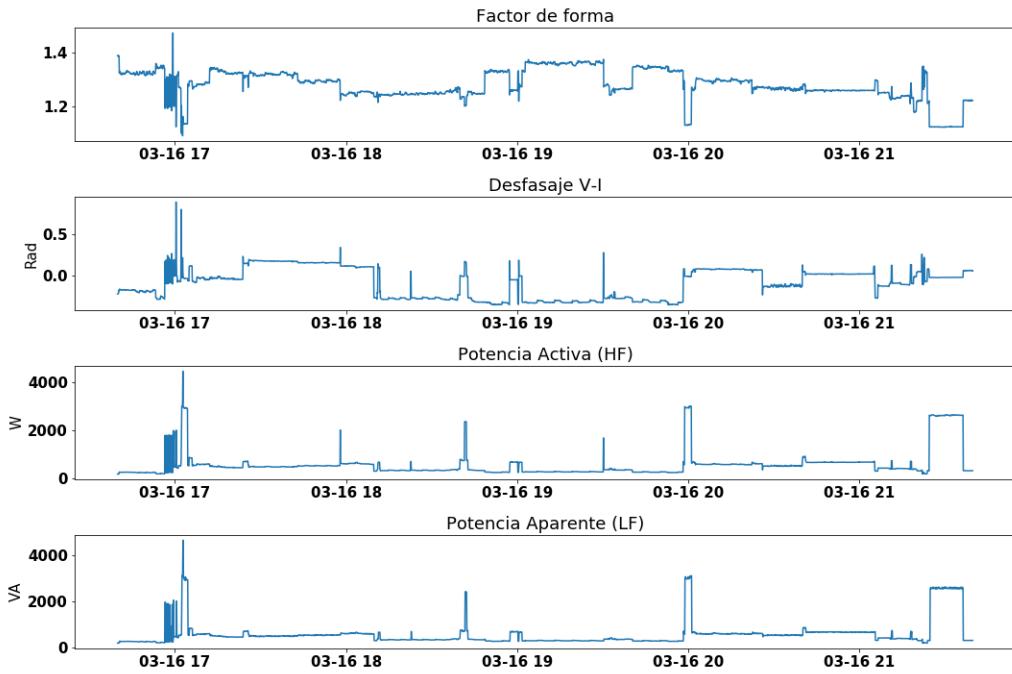


Figura 6.4: Ejemplos de los datos. De arriba hacia abajo: Factor de forma, Desfasaje V-I, Potencia Activa (HF), Potencia Aparente (LF)

6.7. Conclusión

La extracción de descriptores se realizó con éxito. Se observan los mismos en la Figura 6.4, donde también puede verse una comparación entre la potencia agregada de baja y alta frecuencia. El enfoque planteado fue de comprimir la información en series de baja frecuencia, mediante el cálculo de descriptores correspondientes a una lista de fechas predefinidas. Las restricciones de ancho de banda, infraestructura de almacenamiento, y capacidad de procesamiento fueron compensadas por los métodos expuestos en este capítulo y el tiempo utilizado. Si bien varios procesos se automatizaron, hubo una fuerte supervisión durante la ejecución de los mismos. Un ejemplo de archivo de configuración del `wget` se presenta en el Apéndice D, y el código asociado a esta parte está disponible en el repositorio, dentro de la carpeta “HF-UK”. Por otro lado, la obtención de los vectores utilizados como entrada a las redes fue ejecutada satisfactoriamente. Esto permite encontrar “activaciones” y “no-activaciones”, que como se verá en el siguiente Capítulo, formarán parte del conjunto de datos de entrenamiento.

Capítulo 6. Base de datos UK-DALE

Capítulo 7

Desagregación con redes neuronales

7.1. Resumen

El problema de desagregación a resolver consiste en estimar series de potencia individuales asociadas a cada electrodoméstico de un conjunto predefinido, a partir de ventanas de la serie temporal agregada de potencia consumida por la casa.

Este capítulo desarrolla los algoritmos utilizados con base en el trabajo de Kelly et al. [19], basados en redes neuronales artificiales. Dichas redes son funciones paramétricas, dentro de las cuales se escogen aquellas más promisorias como definitivas, como será detallado en el Capítulo 8, y se evaluarán en profundidad en el Capítulo 9 y 10. Para cada electrodoméstico a clasificar, se entrena una red neuronal dedicada.

Se expone el proceso de entrenamiento, se presentan las arquitecturas a ser utilizadas (de nombres “rectángulos” y “autoencoder”), y las modificaciones realizadas por sobre la publicación de referencia: hacer al “autoencoder” más profundo y modificar la primera capa de las redes para que acepte series multivariadas incorporando descriptores de alta frecuencia. Estas redes fueron entrenadas con los datos descritos en el Capítulo 6

7.2. Introducción a las Redes Neuronales Artificiales

Las redes neuronales artificiales (ANN) son funciones paramétricas cuyos coeficientes se ajustan utilizando datos de entrenamiento y minimizando el error respecto a un valor objetivo asociado a cada dato de entrenamiento. Es posible pensar a las ANN como una caja negra que intenta implementar una función dados ejemplos de pre-imágenes e imágenes.

Las ANN están formadas por unidades básicas llamadas neuronas. Cada neurona realiza una combinación lineal de sus entradas y da como salida el resultado de aplicarle otra función, en general no lineal, al valor obtenido por la combinación lineal. Los parámetros a ser ajustados son los pesos de cada combinación lineal que se produce en la red.

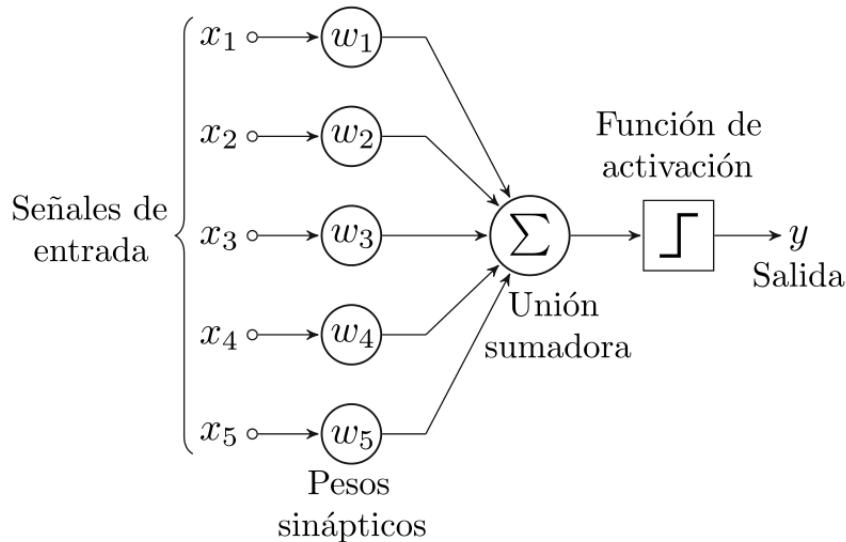


Figura 7.1: Esquema de una neurona, unidad básica que conforma una red

Apilando de distinta manera varias neuronas (poniendo como entrada de una neurona la salida de otra) se puede obtener una función continua, no lineal, con diversos grados de complejidad, que se ajuste a los datos experimentales.

Las redes neuronales pueden utilizarse tanto para clasificación como para regresión. En este trabajo se utilizará el último enfoque. A la hora de ajustar los parámetros, se debe definir una función de costo (*loss function*), y por métodos de optimización basados en el gradiente de la misma, se actualizan los parámetros en la dirección que minimice la función de costo, o direcciones similares en algoritmos más complejos. En este proyecto se utiliza como *loss function* al error cuadrático medio (MSE) entre la salida de la red y el valor objetivo.

7.2.1. Tipos de redes utilizadas

A continuación se presentan brevemente los tipos de capas o redes neuronales utilizadas para este proyecto

Redes Densas

Las redes densas son formadas por capas de neuronas apiladas, en la que cada neurona de una capa previa se conecta con todas las neuronas de la capa siguiente. Es decir, la salida de cada función de activación es una entrada de cada neurona de la siguiente capa. La disposición de los nodos se observa en la Figura 7.2.

7.2. Introducción a las Redes Neuronales Artificiales

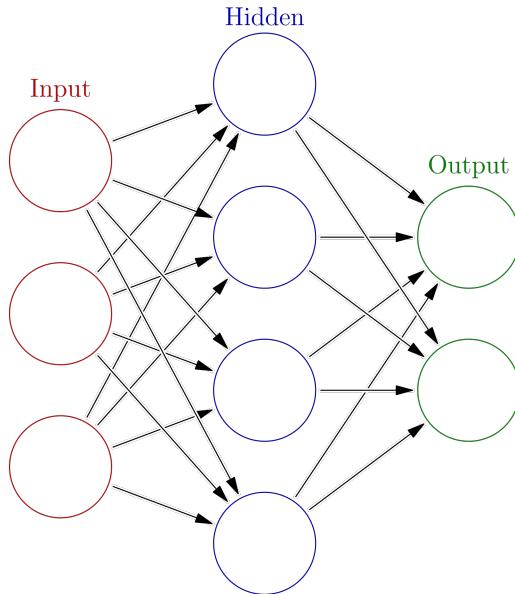


Figura 7.2: Esquema de una red neuronal densa con una capa oculta

Redes Convolucionales

Ampliamente utilizadas sobre imágenes y audio, las redes convolucionales presentan invarianza sobre el dominio de los datos. En lugar de tener conexiones densas (donde a cada punto de los datos le corresponde un peso), las neuronas de estas redes comparten sus pesos para distintos puntos de la secuencia de datos. La intuición detrás de esto es que si una neurona aprende a detectar, por ejemplo, un flanco en una señal de consumo, lo hará sin importar si dicho flanco está al inicio, en el medio o al final de la secuencia. En aplicaciones prácticas se suelen combinar capas de convolución con capas densas. Por esta razón, la propiedad de invarianza en el dominio de los datos está presente solamente en las primeras capas, convolucionales, que extraen características de bajo nivel. Por esta propiedad de invarianza en el tiempo es que todas las redes implementadas en este trabajo presentan capas de convolución al inicio.

Autoencoder

Los Autoencoder son un tipo de red neuronal formadas por dos grandes bloques: *encoder* y *decoder*. El *encoder*, consiste en una red neuronal que toma como entrada los datos de dimensión m y los lleva a una dimensión menor n , codificándolos. El *decoder* toma como entrada un vector de dimensión n y transforma este vector en datos de dimensión m , decodificándolo o reconstruyéndolo. Ambos bloques se entrena en conjunto con el objetivo de codificar y reconstruir los datos de entrada de la mejor forma posible.

La idea fundamental es lograr aprender una representación de los datos en un espacio de menor dimensión que permita recuperarlos en su espacio original. Un caso de uso típico de los autoencoders es para quitar ruido en imágenes, bajo la premisa de que las imágenes naturales (sin ruido) presentan suficiente estructura y viven en una dimensión menor que la imagen con ruido. Al pasar una imagen con ruido por

Capítulo 7. Desagregación con redes neuronales

un autoencoder, muchas veces el ruido se elimina ya que las componentes de ruido no pueden recuperarse a partir del código de menor dimensión.

El trabajar con un espacio latente de menor dimensión tiene sentido en el problema de NILM. Si la serie agregada pertenece a un subespacio de baja dimensión, en el espacio latente se codificarían los pesos sobre los vectores de alguna base del subespacio. Se espera que la serie agregada pertenezca a un subespacio de baja dimensión ya que tiene mucha estructura, originada en la suma de contribuciones de señales individuales que tienen poca variabilidad.

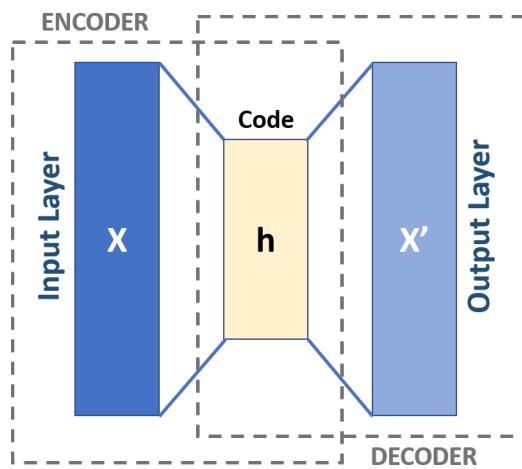


Figura 7.3: Esquema de un autoencoder

7.3. Conjuntos de datos

7.3.1. Datos sintéticos

Una de las estrategias que se explorará más adelante para evaluar si se mejora la performance del modelo base es la de utilizar *data augmentation*. Como estrategia de *data augmentation* para los datos de baja frecuencia se creó un conjunto de datos sintéticos, bajo la idea de que teniendo más datos las redes puedan generalizar mejor, ya que los datos extra aportan a la regularización.

Se comienza extrayendo las activaciones de todos los electrodomésticos de interés: jarra eléctrica, lavadora, lavavajillas, microondas y heladera. Se respetaron los mismos largos de vectores que para los datos reales. A continuación, se hará una distinción entre la creación de las muestras que contengan activaciones completas y las que no.

Supongamos que queremos crear los datos sintéticos del “electrodoméstico base”.

- Muestras con activaciones: Se comienza con una muestra, de tamaño adecuado, al electrodoméstico base, de ceros. Se escoge entre todas las activaciones una al azar. Se sortea una posición de la activación, de forma que la misma quede completamente contenida en el vector; a partir de esto se fijan los valores objetivos concordantes con este sorteo. Luego, para cada uno de los electrodomésticos distractores, se sortea con una probabilidad p (usualmente 0,4), si alguna de las activaciones correspondiente al electrodoméstico distracto se encontrará en la muestra, ya sea completamente contenida o no.

7.3. Conjuntos de datos

- Muestras sin activaciones: Se comienza con una muestra de ceros. Para cada uno de los electrodomésticos distractores que no sean el base, se sorteaa con una probabilidad p si alguna de sus activaciones aparecerán en la muestra. El vector objetivo será, tanto para la red de rectángulos como para el autoencoder, estrictamente nulo.

Se tomó el cuidado de generar una base balanceada, es decir, con la misma cantidad de datos con activaciones que sin activaciones. Al igual que es comentado en [19], este enfoque de generación de datos presenta una gran pérdida de estructura. A modo de ejemplo, la heladera tiene activaciones a una tasa aproximadamente constante, pero la base sintética no lo tiene en cuenta.

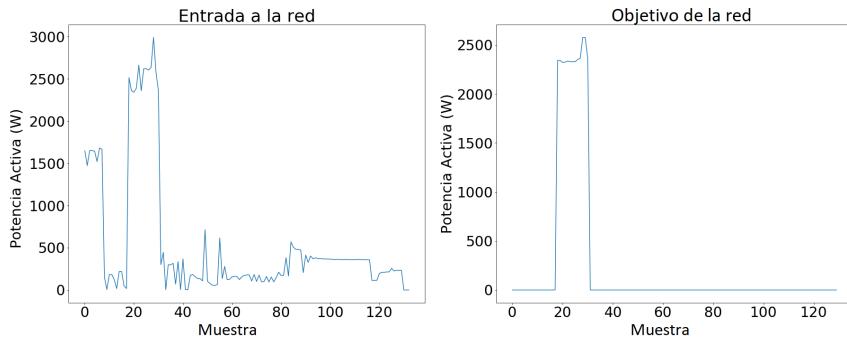


Figura 7.4: Ejemplos de los datos sintéticos para la jarra eléctrica con la red autoencoder.

7.3.2. Entrenamiento, validación y test

A la hora de entrenar un algoritmo y reportar sus resultados, es necesario tener al menos dos conjuntos de datos disjuntos: entrenamiento y test. Esto permite evaluar si el algoritmo generaliza bien al problema, es decir, si funciona para datos sobre los que no se lo entrenó, o si simplemente está “memorizando” la salida que debería entregar para cada dato de entrada. Los algoritmos se entrena utilizando los datos de entrenamiento y se reporta su *performance* sobre los datos de test. Si hay un gran deterioro entre los resultados sobre ambos conjuntos entonces es probable que exista *overfitting* y el algoritmo no generalice bien a una casa desconocida.

Si se cuenta con la cantidad de datos suficiente, es usual apartar un tercer conjunto de datos sobre el cual tampoco se entrena: el conjunto de validación. Este conjunto se utiliza para evaluar los hiperparámetros de los algoritmos (parámetros que no se entrena junto con los coeficientes del modelo). En el caso del Proyecto el conjunto de datos de validación se utilizó para definir los valores relacionados al cálculo de activaciones, no activaciones, largo de ventana e hiperparámetros de las redes. Sobre estos datos no se entrena con descenso por gradiente, sino que se utilizan para obtener estimaciones del error de generalización, y se ajustan los parámetros para minimizarlas. Una vez que se elige el mejor modelo entrenado (el que reporta mejores resultados sobre validación), se lo evalúa sobre el conjunto de test y esas métricas son las reportadas. Si no se tuviera este conjunto, resulta evidente que ajustar los hiperparámetros sobre los datos de entrenamiento se enfoca en obtener los mejores resultados posibles sobre el conjunto de entrenamiento y no en obtener un algoritmo capaz de generalizar.

En este trabajo el conjunto de entrenamiento es un conjunto de muestras que se pueden asumir generadas por cierta distribución de probabilidad P_e . De este conjunto de entrenamiento, un 20% de los vectores se usarán como conjunto de validación.

Capítulo 7. Desagregación con redes neuronales

Estos datos de validación también son generados por P_e . Estos conjuntos provienen de medidas de dos o más casas, dependiendo del electrodoméstico. Estas medidas no incluyen las últimas dos semanas del período de medición. Estas últimas dos semanas serán utilizadas para evaluar el algoritmo, ya que de ellas se extraen vectores que forman parte del primer conjunto de test. Se puede asumir que, a falta de cambios estructurales (reforma de la instalación, cambio de hábitos de consumo o reemplazo de electrodomésticos), estos datos de test también son generados por P_e . Por otro lado, se toma un segundo conjunto de test, cuyos datos provienen de otra distribución de probabilidad. Este segundo conjunto de test incluye medidas correspondientes a casas que nunca fueron vistas durante el entrenamiento de las redes. Se asume que estos datos son generados por una distribución P_{UK} , que difiere de P_e . El tercer y último conjunto de test es el que se forma con los datos pertenecientes a la base NILM-UY. Estos datos son generados por una distribución P_{UY} , que se presume difiere más de P_e de lo que lo hace P_{UK} , ya que, por ejemplo, en Uruguay se usan calefones.

	Entrenamiento	Test
Jarra eléctrica	1, 2, 3, 4	5
Heladera	1, 2, 4	5
Lavarropas	1, 5	2
Microondas	1, 2	5
Lavavajillas	1, 2	5

Tabla 7.1: Casas de entrenamiento y test en UK-DALE

7.3.3. Comentario sobre el conjunto test visto

En la subsección anterior se comentó acerca de la separación en dos conjuntos de test, uno de casas vistas, formado por las últimas dos semanas de cada casa de entrenamiento, y otro de casas no vistas durante el entrenamiento. De ahora en más se llamarán “test visto” y “test no visto”. Cuando se planeó este método no se habían explorado al detalle los datos de alta frecuencia, por lo que se suponía que tendrían una disponibilidad de datos similar a los de baja frecuencia. Lamentablemente, como se detalla en el Capítulo 6, los datos de alta frecuencia no abarcan el mismo rango temporal que los datos de baja frecuencia, por lo que el conjunto de test visto no incluye ninguna muestra de la casa 1. Esto es relevante porque la casa 1 es la mayor fuente de datos de la base UKDALE, causando dos limitantes a la evaluación: la primera es que se tiene una escasez de datos, ya que, por ejemplo, para el lavarropas solo se tienen 20 activaciones en test visto. La segunda es que cambia la distribución de los datos entre el conjunto que se entrenó y el que se evaluó. En entrenamiento, la mayoría de los ejemplos son de la casa 1, mientras que en test visto no hay ejemplos de esta casa. En el entrenamiento puede existir un sesgo hacia los electrodomésticos de la casa 1 que no se repite en el conjunto de test.

El conjunto que resulta similar al que se quisiera tener como test visto, es el conjunto de validación. Conceptualmente, las medidas provienen de la misma distribución de datos, y se puede argumentar que son tan desconocidas para el algoritmo como las activaciones pertenecientes a las últimas dos semanas. Sin embargo, como se ajustó sobre este conjunto a la hora de elegir el mejor modelo, tampoco resulta correcto considerar como test visto al conjunto de validación, ya que se tendrían resultados sobre ajustados o demasiado optimistas.

7.4. Redes de desagregación

7.3.4. Preprocesamiento de entrada a las redes

Como se indica en [19], es deseable que los datos de entrada y salida de la red estén normalizados y estandarizados. Para cumplir con esto se calcula la media de la desviación estándar de cada vector de entrada a la red σ_{entrada} y el valor de potencia máxima del target de la red \max_{target} , sobre el conjunto de datos de entrenamiento.

Luego, a cada entrada a la red, sea en entrenamiento, validación o test, se le sustrae su propia media y se la divide por σ_{entrada} . En entrenamiento, a los target se les divide por \max_{target} . De esta forma, las entradas y target de la red son valores en el intervalo $[0, 1]$. Cuando se quiere hacer una predicción, debe tomarse la salida de la red y multiplicarla por \max_{target} para llevar su valor de potencia a Watts.

7.4. Redes de desagregación

En esta sección se detallan las redes utilizadas para la desagregación de potencia eléctrica. Se tiene una red específicamente entrenada para cada uno de los cinco electrodomésticos elegidos. Se recuerda que el objetivo de estos algoritmos o redes es reconstruir la serie de potencia consumida por un electrodoméstico a partir de la serie de potencia aparente agregada de toda la casa.

En primer lugar se presentan las redes basadas en [19] y luego se comentan las variantes introducidas por los autores en busca de obtener una mejor *performance*.

7.4.1. Arquitecturas de referencia

Denoising autoencoder

Basándose en la idea de los autoencoders como algoritmos para eliminar el ruido de una señal, en [19] se propone el enfoque de considerar que la señal agregada es una señal ruidosa y la señal desagregada es una señal limpia. Bajo esta hipótesis el ruido aditivo sería la potencia consumida por todo el resto de los electrodomésticos, a priori aleatorio, ya que en distintos momentos podrían estar encendidos diferentes electrodomésticos. Esta arquitectura se presenta en la Tabla 7.5.

Para los autores esta interpretación es algo forzada: un *denoising autoencoder* se basa en algunos principios que hacen que sea eficiente en término de etiquetado de datos y que tenga sentido matemático. El primero responde al prefijo “auto”. Normalmente se considera a un autoencoder como una forma de implementar lo que hoy en día se conoce como *self supervised learning*. Esto es, no se requieren etiquetas, sino que se usa como objetivo la misma entrada o alguna parte de ella. Sin embargo, el trabajo de referencia lo utiliza como un algoritmo de aprendizaje supervisado. En cuanto a la codificación, a la hora de quitar ruido tiene sentido bajar a un espacio de dimensión menor. La dimensión del espacio latente debería ser aproximadamente la de la señal limpia. En el caso del trabajo de referencia, la dimensión del espacio latente se define igual a 128 independientemente del electrodoméstico. Es decir, no se toma en cuenta la complejidad del electrodoméstico para definir la longitud del código. Es probable que la calidad de *denoising* del algoritmo planteado en el trabajo de referencia se deba más al carácter supervisado del mismo que al pasaje por un espacio de menor dimensión. Se recomienda entonces la interpretación del mismo algoritmo como uno de aprendizaje supervisado, ya que es necesario contar con las medidas intrusivas para entrenarlo, con una capa intermedia con menor dimensión. Aún así, se seguirá utilizando la denominación “autoencoder”.

Capítulo 7. Desagregación con redes neuronales

Layer (type)	Output Shape	Param #
conv1d_2 (Conv1D)	(None, 127, 8)	40
flatten_1 (Flatten)	(None, 1016)	0
dense_3 (Dense)	(None, 1016)	1033272
dense_4 (Dense)	(None, 128)	130176
dense_5 (Dense)	(None, 1016)	131064
reshape_1 (Reshape)	(None, 127, 8)	0
zero_padding1d_1 (ZeroPadding1D)	(None, 130, 8)	0
conv1d_3 (Conv1D)	(None, 130, 1)	33

Total params: 1,294,585
Trainable params: 1,294,585
Non-trainable params: 0

Figura 7.5: Arquitectura del denoising autoencoder para la jarra eléctrica (largo de entrada 130)

Red de rectángulos

Observando que muchas veces la forma de las activaciones puede aproximarse por un rectángulo sin que esto conlleve a un error significativo en potencia, este algoritmo busca aproximar la potencia consumida por un electrodoméstico con una forma de onda rectangular.

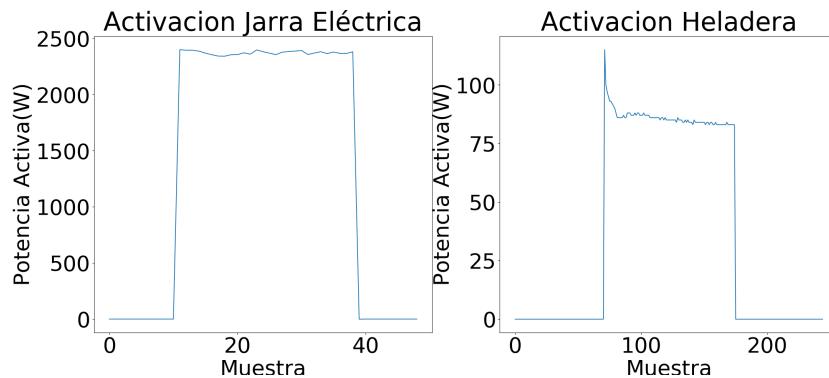


Figura 7.6: Ejemplo de activaciones ‘rectangulares’

7.4. Redes de desagregación

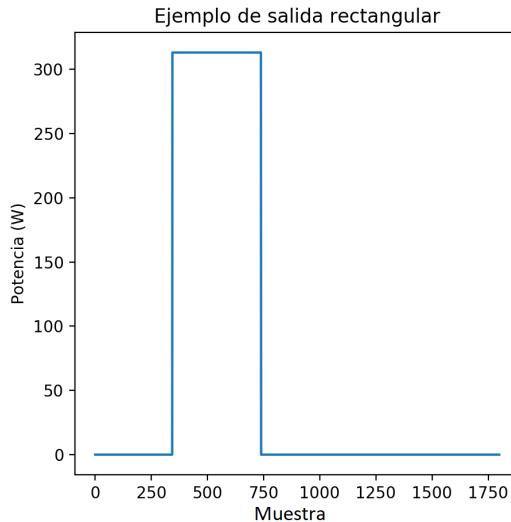


Figura 7.7: Ejemplo de salida rectangular, cuyo vector asociado es: [313, 0,19, 0,41]

Explotando la observación anterior puede reducirse el problema de desagregación de potencia a la estimación de 3 números: potencia media, inicio y fin de la activación. Conociendo esos tres valores, queda definido el rectángulo que encierra la activación. El inicio y fin de la activación se representan como *floats* que indican en qué proporción de la ventana inicia o finaliza el rectángulo. Cuando no se detecta una activación la red debe dar como salida un vector formado por tres 0, indicando que no hay ningún rectángulo. Esta red tiene un par de capas convolucionales al comienzo, y capas completamente conectadas después. Esta arquitectura se muestra en la Tabla 7.8.

7.4.2. Variantes propuestas

Incorporación de descriptores de alta frecuencia

Para aportarle más información a la entrada de las redes, se toman las series multivariadas, como fueron descritas en el Capítulo 6. Teniendo como entrada de la red una serie de potencia activa, factor de forma y desfasaje entre corriente y voltaje se busca que las redes logren predecir de mejor forma el consumo de potencia activa individual de un electrodoméstico. Recordar que esos descriptores de alta frecuencia fueron elegidos basándose en su performance en el problema de PLAID en el Capítulo 5.

Las arquitecturas utilizadas son muy similares a las presentadas en la sección anterior. La única diferencia entre las arquitecturas es que las dimensiones de la primera capa de convolución se adaptan a la dimensión de la entrada. Esto es porque se consideró que los descriptores de alta frecuencia aportan valor al facilitar la distinción de patrones de bajo nivel, pero no lo deberían cambiar los requerimientos de la arquitectura de las siguientes capas. Estas arquitecturas son las reportadas en las Tablas 7.9 y 7.10.

Big autoencoder

Si consideramos la complejidad de los modelos en función de la cantidad de parámetros, podemos observar que la red de los rectángulos es sensiblemente más compleja (y

Capítulo 7. Desagregación con redes neuronales

Layer (type)	Output Shape	Param #
conv1d_4 (Conv1D)	(None, 127, 16)	80
conv1d_5 (Conv1D)	(None, 124, 16)	1040
flatten_2 (Flatten)	(None, 1984)	0
dense_6 (Dense)	(None, 4096)	8130560
dense_7 (Dense)	(None, 3072)	12585984
dense_8 (Dense)	(None, 2048)	6293504
dense_9 (Dense)	(None, 512)	1049088
dense_10 (Dense)	(None, 3)	1539

Total params: 28,061,795

Trainable params: 28,061,795

Non-trainable params: 0

Figura 7.8: Arquitectura de la red de rectángulos para la jarra eléctrica

por ende quizás tenga mayor capacidad de aprender sobre los datos) que el autoencoder. A partir de la observación anterior y de que no resultaba del todo razonable que la dimensión del espacio latente fuera la misma para todos los electrodomésticos sin importar los distintos largos de ventana y las distintas formas de la señal, se propone una arquitectura similar pero cuya dimensión del espacio latente sea una función del largo de ventana. Además, se hace a la arquitectura más profunda. Se agrega una capa de convolución y un par de capas densamente conectadas. Estas capas intermedias tienen una cantidad de neuronas h que depende del largo de la ventana $winsize$. En particular, la capa correspondiente al código sigue la relación $h = winsize//10$. La arquitectura final se presenta en la Tabla 7.11.

7.4. Redes de desagregación

Layer (type)	Output Shape	Param #
conv1d_2 (Conv1D)	(None, 127, 16)	208
conv1d_3 (Conv1D)	(None, 124, 16)	1040
flatten_1 (Flatten)	(None, 1984)	0
dense_3 (Dense)	(None, 4096)	8130560
dense_4 (Dense)	(None, 3072)	12585984
dense_5 (Dense)	(None, 2048)	6293504
dense_6 (Dense)	(None, 512)	1049088
dense_7 (Dense)	(None, 3)	1539
Total params:	28,061,923	
Trainable params:	28,061,923	
Non-trainable params:	0	

Figura 7.9: Arquitectura de la red de rectángulos de alta frecuencia para la jarra eléctrica

```
Model: "sequential"

-----  

Layer (type)          Output Shape       Param #  

-----  

conv1d (Conv1D)       (None, 127, 8)      40  

-----  

conv1d_1 (Conv1D)     (None, 124, 8)      264  

-----  

flatten (Flatten)     (None, 992)         0  

-----  

dense (Dense)         (None, 1016)        1008888  

-----  

dense_1 (Dense)       (None, 254)         258318  

-----  

dense_2 (Dense)       (None, 13)          3315  

-----  

dense_3 (Dense)       (None, 254)         3556  

-----  

dense_4 (Dense)       (None, 1016)        259080  

-----  

reshape (Reshape)     (None, 127, 8)      0  

-----  

zero_padding1d (ZeroPadding1 (None, 130, 8)  0  

-----  

conv1d_2 (Conv1D)     (None, 130, 1)      33      71  

-----  

Total params: 1,533,494  

Trainable params: 1,533,494  

Non-trainable params: 0
```

Capítulo 7. Desagregación con redes neuronales

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 127, 8)	104
flatten (Flatten)	(None, 1016)	0
dense (Dense)	(None, 1016)	1033272
dense_1 (Dense)	(None, 128)	130176
dense_2 (Dense)	(None, 1016)	131064
reshape (Reshape)	(None, 127, 8)	0
zero_padding1d (ZeroPadding1	(None, 130, 8)	0
conv1d_1 (Conv1D)	(None, 130, 1)	33

Total params: 1,294,649
Trainable params: 1,294,649
Non-trainable params: 0

Figura 7.10: Arquitectura del autoencoder de alta frecuencia para la jarra eléctrica

7.5. Conclusión

En este capítulo se introdujeron las redes neuronales artificiales y las arquitecturas específicas a ser estudiadas. Además se expuso la división del conjunto de datos en conjuntos de entrenamiento, validación, y tres conjuntos de test. Se propusieron también dos variantes, una que permite integrar los datos de alta frecuencia, y otra que amplía la complejidad del autoencoder y que hace que la dimensión del espacio latente dependa del electrodoméstico. En el siguiente Capítulo se presenta la selección final de los modelos, que serán evaluados en los Capítulos 9 y 10.

Capítulo 8

Selección de modelos

8.1. Resumen

Este capítulo trata acerca de cómo se seleccionaron los modelos a ser utilizados en la evaluación. Se implementa una búsqueda en grilla de hiperparámetros y se realizan comparaciones para evaluar posibles mejorías. Partiendo de los modelos definidos en el capítulo anterior, se presentan ligeras modificaciones al proceso de entrenamiento. Además se evalúa cuál de todas las variantes de modelos presentadas tiene el mejor desempeño. Los modelos a ser utilizados en la evaluación son seleccionados a partir del Area Under the Curve (AUC) de la Receiver Operating Characteristic curve (ROC).

A lo largo del capítulo se presentarán algunos resultados con el fin de ilustrar el proceso seguido por los autores para la selección de modelos. No obstante, para no entorpecer la lectura, se refiere al lector al Apéndice E para consultar la totalidad de los resultados.

8.2. Métricas de desempeño

El área de aprendizaje estadístico da un marco teórico para la comprensión de los algoritmos de *machine-learning*. Esta teoría justifica la estimación del error de generalización mediante la utilización de datos que no hayan sido utilizados de forma alguna para modificar el entrenamiento. Dado que hay varios modelos posibles (uno para cada conjunto de pesos), en esta parte se explican las métricas a considerar a la hora de evaluar la capacidad de generalización o desempeño del algoritmo. Sobre estas métricas se define el concepto de “mejor” modelo.

Se reportan las métricas de clasificación usuales, True Positives (TP), True Negatives (TN), False Positives (FP), y False Negatives (FN), y métricas de desagregación que miden errores en la asignación de potencia, Mean Absolute Error (MAE) y Relative Error In Total Energy (REITE). Notar que para las métricas de clasificación no importa la potencia asignada ni en qué parte de la ventana se ha detectado el electrodoméstico, simplemente importa que la red dé como salida algún valor de potencia mayor a un umbral definido, implicando que en esa ventana de tiempo hubo consumo de potencia por el electrodoméstico. Esto no es un inconveniente ya que las ventanas a estudiar son lo suficientemente pequeñas (menos de 3hs) como para que la información de si el electrodoméstico estuvo encendido, o no, sea valiosa de todas formas. Es claro que las métricas de clasificación tiene sentido calcularlas sobre ventanas de un tamaño adecuado para encontrar ejemplos en los que el electrodoméstico este consu-

Capítulo 8. Selección de modelos

miendo y otros en los que no. Luego, las métricas se calculan sobre todo ese conjunto de ventanas.

A continuación se definen las métricas a ser reportadas, comenzando por definir algunos valores de interés:

$$n = \text{Número total de ventanas}$$

$$E = \text{Energía total consumida}$$

$$\hat{E} = \text{Energía total estimada por el algoritmo}$$

$$y_t = \text{potencia instantánea del electrodoméstico}$$

$$\hat{y}_t = \text{potencia instantánea estimada}$$

$$FPR = \frac{FP}{N} = \frac{FP}{FP + TN} = \text{Tasa de Falsos Positivos}$$

Métricas de clasificación

$$\text{recall} = \frac{TP}{TP + FN} \quad (8.1)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (8.2)$$

$$\text{accuracy} = \frac{TP + TN}{n} \quad (8.3)$$

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (8.4)$$

En caso que algún denominador se anule, el valor reportado es “nan”. Para el cálculo de estas métricas es necesario conocer TP, FN, FP y TN, para esto se deben definir criterios con el fin de establecer si tanto en el *target* como en la predicción se encuentra el electrodoméstico prendido o no. Los mismos serán explicados más adelante en el presente Capítulo.

Métricas de Regresión

$$REITE = \frac{|\hat{E} - E|}{\max(E, \hat{E})} \quad (8.5)$$

$$MAE = \frac{1}{T} \sum_{t=1}^T |\hat{y}_t - y_t| \quad (8.6)$$

Área debajo de la curva ROC

La curva ROC es simplemente el gráfico de la tasa de TP (recall) contra la tasa de FP (FPR) que se obtiene al variar el umbral de detección sobre el cual la salida de la red indica la presencia de un electrodoméstico en una ventana dada. El área abajo de la curva (AUC) es la integral de 0 a 1 de la curva ROC. Esta métrica es un escalar que representa el desempeño global, no es necesario fijar un punto de operación como en las métricas de clasificación.

Más adelante se utilizará la métrica AUC para decidir cuál es el mejor modelo, de acuerdo a las razones que siguen.

8.3. Modos de evaluación

1. En primer lugar, las métricas de “regresión” (MAE y REITE) son fáciles de interpretar, pero sin embargo sus valores no son representativos de un correcto desempeño de los algoritmos. Observando los resultados se pudo ver que a veces valores pequeños de MAE o REITE no necesariamente corresponden a modelos que se comportan bien. Por ejemplo, a veces el modelo con menor MAE da siempre la misma salida sin importar la entrada. Este último ejemplo puede explicarse para los electrodomésticos con ciclos de funcionamiento fijos (heladera por ejemplo), donde predecir siempre la misma salida resulte en un error pequeño en promedio, aunque la estrategia carece de utilidad práctica. Otro argumento por el cual no resulta tan clara la interpretación de las métricas de regresión, es que este tipo de métricas penalizan el asignar potencia en un instante distinto al que el electrodoméstico estaba consumiendo. Esto es, un desfasaje dentro de una misma ventana sumaría términos de error, incluso siendo un desfasaje en la salida de un algoritmo perfecto. No es prioritario en este Proyecto conocer el segundo exacto en el cual se enciende un electrodoméstico. Importa más tener una noción global del funcionamiento de los electrodomésticos. En búsqueda de esta métrica más “global”, aparecen como una buena opción las métricas de clasificación. Dado que los modelos fueron entrenados para minimizar un error de “regresión” (MSE), no resulta esperable que un modelo entrenado pueda obtener un buen desempeño de clasificación pero un notorio mal desempeño al aproximar la señal desagregada. Este último punto fue comprobado experimentalmente, los modelos con mejores resultados en clasificación, eran los que visualmente mejor reconstruían la señal desagregada.
2. En segundo lugar, las métricas de clasificación previamente definidas (recall, precision, accuracy, F1) se determinan para un punto de operación dado, siendo éste el umbral para el cual se considera que un electrodoméstico está encendido o no. Por lo tanto, depender de ese punto de operación resulta en métricas que no solamente dependen del algoritmo, sino que dependen también del punto de trabajo (este es un enfoque distinto a entender al algoritmo como el conjunto {modelo, umbral}). Es por esto que se decidió utilizar la AUC como medida de desempeño general de un algoritmo, a la hora de decidir cuál era el mejor modelo. Cuanto más próxima a 1 esté la AUC, significará que el modelo es capaz de detectar de mejor forma cuando un electrodoméstico está en la ventana, y debido a que el algoritmo fue entrenado minimizando una *loss* para que a la salida se obtenga la curva de consumo desagregado, es probable que también pueda reconstruir la señal desagregada de forma razonable. Una AUC cercana a 1 necesariamente implica la existencia de un punto de operación donde hay una alta tasa de TP y una baja tasa de FP.

8.3. Modos de evaluación

A la hora de evaluar los resultados obtenidos, se consideran dos enfoques. El primero trabaja sobre los vectores de entrada y salida, como fueron definidos en el Capítulo 6. En esta evaluación se puede analizar la capacidad de aprendizaje que tuvieron las redes, comparando la salida que entrega la red contra la salida esperada o *target*: un rectángulo de potencia media en el caso de la red de rectángulos, y la primera activación que aparece en la ventana en el caso de los autoencoder.

El segundo enfoque no trabaja sobre los vectores de activaciones sino que lo hace directamente sobre la serie temporal completa. Para ello, se considera la serie temporal agregada y una ventana deslizante sobre la misma con el objetivo de reconstruir la serie desagregada completa. El intervalo de la serie que cae dentro de la ventana es la

Capítulo 8. Selección de modelos

entrada a la red. Se consideran ventanas que se mueven con $STRIDE = 1$ y se estima para cada una de ellas la correspondiente salida de la red. Hay en total n ventanas, siendo n igual al largo de la serie. De esta forma los puntos son estimados tantas veces como el largo de la ventana sobre la cual se predice, por lo que el valor de cada punto será el promedio entre todas las predicciones del mismo. Por lo tanto, lo que se evalúa es el poder reconstructor de los algoritmos, en el sentido de que se quiere ver cuán parecida es la señal reconstruida de la medida desagregada. Se recuerda que los modelos fueron entrenados para detectar solamente la primera activación completa dentro de la ventana, por lo que habrá un factor de compensación a ser establecido cuando se promedia, que de no considerarse siempre se subestimaría la potencia desagregada.

Una vez reconstruida la señal, es necesario definir nuevas ventanas sobre la cual decidir si el electrodoméstico está consumiendo o no, con el objetivo de calcular las métricas de clasificación. Estas ventanas se eligen de forma tal que sean **disjuntas** y consecutivas. Luego de estimar la serie de potencia, se evalúa sobre las ventanas escogidas. Para las métricas de desagregación simplemente se calculan las diferencias de potencia entre la serie reconstruida por la red y la serie desagregada de referencia.

Observar que el “dataset” del segundo enfoque, excepto para el caso de la heladera, deja de ser balanceado porque la mayor parte del tiempo un electrodoméstico no está encendido. Este enfoque se utilizará solamente en el proceso de evaluación de los modelos y no en el de selección.

8.4. Entrenamiento

En esta sección se presentan las distintas decisiones y variantes implementadas en el entrenamiento en pos de lograr el mejor resultado posible. Dadas las arquitecturas definidas como en el Capítulo 7 y un conjunto de datos de entrenamiento fijo, lo que diferencia los modelos para un mismo electrodoméstico es el proceso de entrenamiento. El proceso de entrenamiento depende de la inicialización de los pesos de la red (en este caso uniforme Glorot [12]), del algoritmo de optimización con sus parámetros, y del número de épocas en las que se entrena, entre otras cosas.

8.4.1. Selección de la grilla

El entrenamiento de un modelo dado no siempre converge al mismo punto, debido a la inicialización aleatoria, ni hay garantías teóricas de convergencia, debido a que la superficie no es convexa. Se probaron a mano varios optimizadores y tasas de aprendizaje, además de métricas objetivo. En particular se probó y descartó utilizar la función Mean Absolute Error de Tensorflow como función de costo, por no obtener buenos resultados en ninguno de los casos. Se nota que esta métrica objetivo difiere del MAE definido arriba (8.6), ya que (i) la salida de la red se encuentra normalizada y (ii) en el caso de los rectángulos, el target es un vector tridimensional (solo una de esas dimensiones es la potencia). Por otro lado, también se probó y descartó utilizar el optimizador Adadelta, que si bien lograba converger a valores apropiados algunas veces, tenía una velocidad de convergencia mucho más lenta que los otros optimizadores testeados (Adam y Adamax). Esto sucedía para las distintas tasas de aprendizaje. La tasa de aprendizaje inicial también fue modificada, y se encontró que 0,005 era un valor excesivamente grande. Teniendo en cuenta estas consideraciones, se procedió a evaluar el modelo en una grilla definida por las 6 combinaciones posibles entre los siguientes parámetros:

- optimizer = [Adam, Adamax]
- learning rate = [0.001, 0.002, 0.0005]

8.4. Entrenamiento

Debido a la inicialización aleatoria, puede suceder (y de hecho sucede) que en una corrida un optimizador con cierta learning rate converja, y en otra corrida no. Esta variabilidad normalmente se ataca haciendo varias corridas. Se prescindió de este enfoque para mantener el tiempo de corrida limitado, y porque se asume que la grilla se encarga de reducir un poco de esa variabilidad: se espera que haya más de un modelo con desempeño razonable en la grilla. Se nota sin embargo que este efecto agrega incertidumbre a las conclusiones que puedan extraerse de una comparación entre los puntos de la grilla. Dicha comparación se realizaría de forma más justa fijando la misma semilla (*seed*) para cada punto. Se deja la puerta abierta como trabajo a futuro a la realización de una comparación más justa entre los optimizadores o puntos de la grilla.

Un ejemplo del resultado de una búsqueda en grilla para el caso base (arquitecturas de referencia, baja frecuencia) se presenta en las Tablas 8.1 y 8.2.

learning rate / optimizador	Adam	Adamax
0.0005	0.882	0.900
0.001	0.866	0.887
0.002	0.500	0.892

Tabla 8.1: AUCs de la red “rectángulos” para cada punto de la grilla entrenada con datos reales y sintéticos para el lavavajillas

learning rate / optimizador	Adam	Adamax
0.0005	0.499	0.831
0.001	0.500	0.499
0.002	0.500	0.500

Tabla 8.2: AUCs de la red “autoencoder” para cada punto de la grilla entrenada con datos reales y sintéticos para la heladera

8.4.2. Número de épocas

La cantidad de épocas utilizada en el entrenamiento es un parámetro del cual un desarrollador puede independizarse utilizando la técnica que se explica a continuación. Se entrena durante muchas épocas y se utiliza el modelo que mejor desempeño haya obtenido, sobre el conjunto de validación, de entre los modelos parciales por época. De esta manera se logra entrenar y seleccionar fácilmente los modelos a ser evaluados.

Con base en algunos resultados recientes, como [30], se decidió entrenar de esta forma y no hacer *early stopping*, ya que existen contextos en los que parar el entrenamiento en el momento que la *loss* de validación comienza a crecer no es adecuado porque de seguirse entrenando por más épocas, podría observarse un segundo descenso de la misma y obtenerse un mejor modelo.

Capítulo 8. Selección de modelos

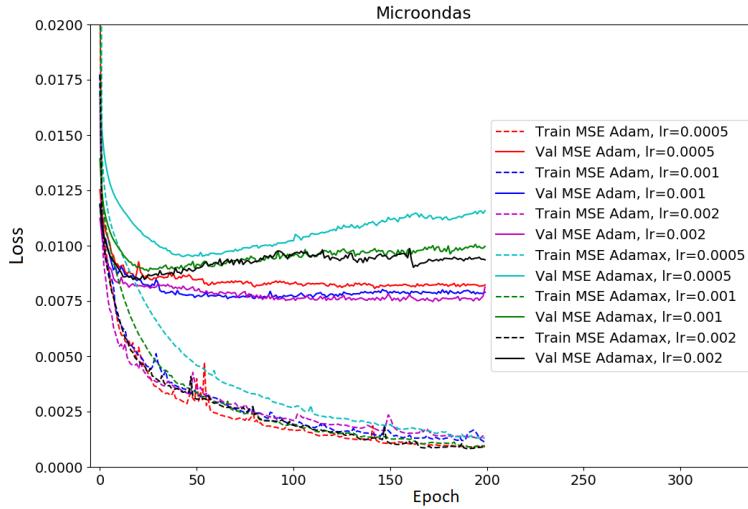


Figura 8.1: Curvas de entrenamiento. $AUC = [0,942, 0,949, 0,937, 0,500, 0,944, 0,941]$ siguiendo el mismo orden que la leyenda de la gráfica

8.4.3. Variantes que afectan el entrenamiento

Sobre las arquitecturas definidas en [19], se proponen distintas variantes en los modelos o en el proceso de entrenamiento con el fin de buscar mejores resultados. Una de las variantes en términos de arquitectura respecto a lo presentado en [19], es la de utilizar un “Big autoencoder”, mientras que las otras variantes de arquitectura consisten en modelos que admiten entradas con información de alta frecuencia, como lo son el autoencoder y rectángulos de alta frecuencia definidos en el Capítulo 7. En cuanto a los datos, para el caso de los modelos de baja frecuencia, de la misma forma que en [19], se analiza si variar el proceso de entrenamiento agregándole datos sintéticos resulta en un modelo con mejor capacidad de generalización.

Considerando los modelos originales y sus variantes, se tienen 7 modelos distintos que se listan a continuación.

	Baja Frec	Datos sintéticos	Alta frecuencia	“Big”
Rectángulos	sí	sí	sí	no
Denoising autoencoder	sí	sí	sí	sí

Tabla 8.3: Variantes y modelos utilizados

8.4.4. Mejores modelos

Cada búsqueda en grilla (ej. Tabla 8.1) devuelve, para cada optimizador, el desempeño de los pesos de la red sobre datos de validación, medido según la AUC. Para cada grilla se tiene solo un conjunto de pesos de la red que reporta la mayor AUC. Este conjunto de coeficientes de buen desempeño fue hallado por alguno de los optimizadores y se le llamará “mejor experimento”.

Se nota que la búsqueda en grilla refiere a un electrodoméstico particular y a un modelo particular (habiéndose 7 modelos por electrodoméstico). Es decir, se realizan búsquedas en grilla para electrodoméstico y para cada uno de los modelos y variaciones

8.5. Evaluación del mejor modelo sobre datos de validación

propuestas (en total se calculan electrodomésticos × modelos = $5 \times 7 = 35$ grillas). Para cada electrodoméstico, se toma el “mejor experimento” de cada una de las 7 variantes y se los compara en una tabla como la Tabla 8.4. Finalmente, se elige como “mejor modelo” a aquel “mejor experimento” que reporte mayor AUC. Notar que el “mejor modelo” de cada electrodoméstico puede ser una variante o modelo distinto para electrodomésticos distintos.

A modo de ejemplo, los resultados para el microondas pueden verse en la Tabla 8.4.

	Baja Frec	Datos sintéticos	Alta frecuencia	“Big”
Rectángulos	0.933	0.937	0.927	-
Denoising autoencoder	0.936	0.944	0.949	0.932

Tabla 8.4: AUC de los mejores experimentos para el microondas

8.5. Evaluación del mejor modelo sobre datos de validación

En base al análisis anteriormente presentado, en esta sección se presenta el mejor modelo para cada electrodoméstico y los resultados de su evaluación sobre los datos de validación.

8.5.1. Mejor modelo para cada electrodoméstico

Dependiendo de cada electrodoméstico, varió cuál arquitectura (dentro de las variantes consideradas) fue la que tuvo mayor AUC sobre los datos de validación. En la Tabla 8.5 se presenta el mejor modelo asociado a cada electrodoméstico.

Electrodoméstico	Modelo
Jarra eléctrica	Autoencoder de Alta Frecuencia
Heladera	Rectángulos de Alta Frecuencia
Lavarropas	Rectángulos de Alta Frecuencia
Microondas	Autoencoder de Alta Frecuencia
Lavavajillas	Big Autoencoder

Tabla 8.5: Mejor modelo para cada electrodoméstico

En el Apéndice E, se incluyen los valores de AUC para todos los modelos y todos los electrodomésticos que dan sustento al siguiente análisis. En base a los resultados para todos los electrodomésticos, parece oportuno destacar que los mejores modelos para todos los electrodomésticos surgen de las variantes o mejoras propuestas por los autores. Parece razonable afirmar que la incorporación de descriptores de alta frecuencia agrega información relevante para la desagregación sobre el conjunto de validación.

8.5.2. Desempeño en desagregación

La evaluación se realiza sobre las activaciones. La comparación con otros algoritmos de desagregación, Factorial Hidden Markov Model y Combinatorial Optimization se

Capítulo 8. Selección de modelos

realiza en [19]. Allí se reporta que los algoritmos basados en redes profundas son mejores que los anteriores (en términos de métricas de clasificación tanto como de regresión). El presente Proyecto implementó las dos mejores arquitecturas y presentó variantes sobre las mismas (rectángulos y autoencoder), dejando de lado el algoritmo basado en Long Short Term Memory networks (LSTM) por obtener peor desempeño. Los mejores valores de MAE reportados en [19] no superan los 28W para ningún electrodoméstico.

Electrodoméstico	Mejor modelo		Rectángulos Kelly et al.	
	MAE (W)	REITE	MAE (W)	REITE
Jarra eléctrica	88	0.040	15	-0.32
Heladera	10	0.115	22	-0.07
Lavarropas	67	0.155	28	-0.65
Microondas	86	0.241	16	-0.23
Lavavajillas	99	0.003	30	-0.53

Tabla 8.6: Resultados de desagregación en validación comparado con rectángulos de Kelly et al.

Electrodoméstico	Mejor modelo		Autoencoder Kelly et al.	
	MAE (W)	REITE	MAE (W)	REITE
Jarra eléctrica	88	0.040	16	0.02
Heladera	10	0.115	25	-0.35
Lavarropas	67	0.155	44	0.18
Microondas	86	0.241	13	0.06
Lavavajillas	99	0.003	21	-0.34

Tabla 8.7: Resultados de desagregación en validación comparado con autoencoder de Kelly et al.

En el artículo no se aclara si esta evaluación fue realizada con o sin ventanas deslizantes. La métrica REITE presentada en el mismo toma valores negativos en muchos casos, lo cual no concuerda con la fórmula planteada: la interpretación más razonable encontrada por los autores es que el signo de su valor indica si se está sobre o sub estimando la asignación de energía. Estas razones se comentan, junto con otras, en la Sección 9.6, donde se aborda la temática de la comparabilidad con la publicación de referencia.

8.5.3. Desempeño en detección o clasificación

Para la evaluación de la detección se considera la AUC obtenida para cada electrodoméstico y se evalúan también las métricas usuales de clasificación. Para esto último se utiliza como umbral a la salida de la red aquel valor que maximiza el $F1 - score$, ya que se busca tener el mejor desempeño posible teniendo en cuenta el compromiso entre el precision y el recall. En este caso, cuando el valor máximo a la salida de la red supere el valor del umbral, se considerará que la red está indicando la presencia de ese electrodoméstico. Como el umbral puede entenderse como un hiperparámetro más del algoritmo, se considera válida también la elección del umbral en la ROC en base a otro criterio distinto al propuesto, lo cual podría llevar a mejores resultados en las demás métricas de clasificación.

Como en esta sección la comparación entre la predicción y el *ground truth* se realiza contra los vectores de activaciones y no activaciones, para saber si efectivamente un

8.5. Evaluación del mejor modelo sobre datos de validación

electrodoméstico estaba consumiendo simplemente hay que ver si el target es distinto de cero. Recordar que si hay un electrodoméstico en la ventana, entonces el target corresponde a la primera activación completa.

Electrodoméstico	AUC
Jarra eléctrica	0.984
Heladera	0.901
Lavarropas	0.951
Microondas	0.949
Lavavajillas	0.997

Tabla 8.8: AUC para el mejor modelo en validación para cada electrodoméstico

Electrodoméstico	Accuracy	Precision	Recall	F1-Score	Umbral de potencia (W)
Jarra eléctrica	0.934	0.955	0.911	0.933	1283
Heladera	0.851	0.849	0.856	0.853	20
Lavarropas	0.896	0.922	0.863	0.892	280
Microondas	0.884	0.833	0.955	0.890	180
Lavavajillas	0.974	0.968	0.981	0.974	982

Tabla 8.9: Métricas de clasificación sobre validación en los mejores modelos, utilizando el umbral de potencia que optimiza el $F1$

Estas métricas corresponden a evaluaciones sobre el conjunto de validación. El testeo de los algoritmos y su desempeño final se presenta en los Capítulos 9 y 10.

Capítulo 8. Selección de modelos

Capítulo 9

Evaluación sobre UK-DALE

9.1. Resumen

En este Capítulo se presenta una primera evaluación de los modelos seleccionados en el Capítulo 8 sobre la base de datos UK-DALE [18]. Se realizan varias comparaciones. En primer lugar se comparan los resultados obtenidos con los “mejores” modelos, aquellos que fueron seleccionados por obtener los mejores valores de AUC en validación, contra la implementación propia de los modelos como fueron propuestos en [19]. En segundo lugar se comparan las métricas de evaluación obtenidas sobre los vectores de activaciones contra las métricas obtenidas mediante la evaluación por ventana deslizante como fueran definidas en el Capítulo 8 . La evaluación se realiza tanto en el período de test seleccionado dentro de las casas utilizadas para el entrenamiento, llamado Test Visto, como en las casas no utilizadas en el entrenamiento, denominado Test No Visto. Por último se comenta acerca de la comparabilidad con los resultados expuestos en [19]. Las conclusiones se resumen en la Sección 9.8.

9.2. Metodología

Las preguntas a ser respondidas, junto con el procedimiento de evaluación, serán descritas en los siguientes párrafos. Como se describe en Capítulo 7, se cuenta con dos conjuntos distintos de datos de test y dos distintas metodologías de evaluación. Los dos conjuntos de datos distintos sobre los que se evalúa son: test visto, conformado por muestras desconocidas de casas utilizadas en el entrenamiento, y test no visto, conformado por muestras de casas no utilizadas en el entrenamiento. Estos dos conjuntos de evaluación permiten estimar cuál es el desempeño de los modelos sobre una casa con la cual se entrenó en el pasado, y cuál es el desempeño de los modelos sobre una casa desconocida, respectivamente. Por otro lado, los dos métodos de evaluación consisten en: evaluación mediante activaciones, que resulta en una evaluación sobre datos con igual proporción de encendidos que de apagados, y evaluación mediante el método de ventanas deslizantes, que presenta una proporción de datos y un mecanismo de desagregación similar al del problema real. Todo lo anterior resulta en cuatro posibles formas de evaluar, que permitirán responder a las preguntas que siguen.

Este capítulo tiene un objetivo macro: reportar cómo funcionan los algoritmos seleccionados. Más específicamente, la pregunta a responder es si el enfoque planteado, basado en redes neuronales, permite brindar una solución satisfactoria al problema de NILM, evaluada de acuerdo a las métricas planteadas en Capítulo 8. Esta interro-

Capítulo 9. Evaluación sobre UK-DALE

gante puede a su vez subdividirse en otras preguntas, que individualmente aportan información relevante. Las mismas son:

- ¿Cuán bien funcionan los modelos sobre un caso de uso para el cual fueron entrenados?
- ¿Cuán bien generalizan los modelos a casas desconocidas para el mismo caso de uso?
- ¿Cómo cambia el desempeño al utilizar una metodología de evaluación que simula un caso de uso real?
- ¿Cuál sería el desempeño esperado de los modelos en un caso de uso real sobre nuevas casas del Reino Unido?

Estas preguntas pueden parecer similares, pero abarcan interrogantes fundamentalmente distintas. La primera refiere al uso ideal de un algoritmo de aprendizaje automático: datos generados por cierta distribución se dividen en train, validación y test; luego se ajustan los datos a train, cuidando que el modelo generalice bien a validación; y por último se evalúa su desempeño en test. Dado que se optimizan los modelos para que tengan un buen desempeño sobre el conjunto de datos de validación bajo la evaluación por activaciones, evaluar cuán bien funcionan en el sentido tradicional implica realizar una evaluación por activaciones sobre muestras desconocidas que surgen de las mismas casas utilizadas para entrenar.

La segunda pregunta aborda la capacidad de generalización, por lo que se utiliza un conjunto de test balanceado, pero con muestras provenientes de casas no vistas. De esta forma se analiza el error de generalización de forma aislada, sin incluir el efecto del desbalance en los datos ni simular un caso de uso real. La tercera pregunta sí cuestiona la influencia del método de evaluación, también de forma aislada, ya que no se usan casas diferentes a las del entrenamiento. Por último se realiza una evaluación final que responde al caso de uso real, con un conjunto de datos de test desbalanceado de casas desconocidas. Lo expuesto aquí se resume en la Tabla 9.1.

Método/Conjunto	Visto	No visto
Activaciones	¿Funciona para lo que se lo entrenó?	¿Cómo generaliza?
Ventanas	¿Cómo cambia el desempeño al usar proporciones reales?	¿Cómo funciona y funcionará?

Tabla 9.1: Resumen de evaluaciones realizadas

Solamente en la última evaluación se reportarán MAE, REITE y métricas de clasificación. En todas las evaluaciones se reporta la AUC, la métrica global de desempeño recomendada por los autores. La AUC es robusta al punto de operación, ya que los tiene en cuenta a todos. Por ende es, en este sentido, preferible frente a las métricas de clasificación tradicionales. Por otro lado, las métricas de regresión, si bien interpretables, no necesariamente se condicen con un comportamiento adecuado del algoritmo. Las evaluaciones a ser realizadas, presentadas anteriormente, se complementan con la comparación con los modelos base presentados en [19]. Estos modelos son rectángulos y autoencoder entrenados con datos sintéticos.

9.3. Comparación sobre test visto usando activaciones

En lo que sigue, se evaluarán los mejores modelos seleccionados en el Capítulo 8, comparándolos con los resultados obtenidos con la implementación de los modelos

9.3. Comparación sobre test visto usando activaciones

base. Los modelos base son el resultado de implementar lo reportado en [19]: las redes autoencoder y rectángulos entrenadas con datos reales y sintéticos agregados. En esta Sección y las siguientes se muestra la ROC y 3 ejemplos de funcionamiento solamente para un electrodoméstico de los cinco por razones de extensión, los resultados completos se muestran en el Apéndice E. Se presentan los resultados sobre los datos correspondientes a los últimos días de las casas sobre las que se entrenó. Este conjunto de datos es el conjunto de test más parecido al conjunto de validación, a pesar de las consideraciones sobre escasez de datos expuestas en la Sección 7.3.3.

En la Tabla 9.2 se muestran los resultados de esta evaluación. Se observa que el “mejor modelo” continúa siendo el que alcanza el mayor desempeño en tres electrodomésticos, siendo superado solamente para el caso del lavarropas y del lavavajillas. Esto indica que este modelo fue entrenado en forma exitosa. Más aún, todos los modelos tienen un desempeño similar sobre validación y test visto, lo cual implica que la capacidad de generalización de los mismos fue cuidada.

		Activaciones	
		Test Visto	Validación
Jarra Eléctrica	Mejor modelo	0.999	0.984
	Autoencoder	0.959	0.977
	Rectángulos	0.941	<u>0.982</u>
Heladera	Mejor modelo	0.941	0.901
	Autoencoder	0.500	0.500
	Rectángulos	<u>0.871</u>	0.811
Lavarropas	Mejor modelo	0.850	0.951
	Autoencoder	0.884	0.875
	Rectángulos	0.812	<u>0.900</u>
Microondas	Mejor modelo	<u>0.976</u>	0.949
	Autoencoder	0.932	<u>0.944</u>
	Rectángulos	<u>0.976</u>	0.937
Lavavajillas	Mejor modelo	0.947	0.997
	Autoencoder	0.986	0.989
	Rectángulos	0.954	<u>0.981</u>

Tabla 9.2: AUC para todos los electrodomésticos en evaluación sobre activaciones. Comparación entre mejor modelo de validación, rectángulos y autoencoder sobre test visto y validación. Las negritas señalan la mejor red entre el mejor modelo, autoencoder base y rectángulos base tanto para test visto como para validación. El subrayado indica para cada red, si se tiene mayor AUC en test visto o validación.

Gráficas

En la Figura 9.1 se muestra la ROC para un electrodoméstico y en la Figura 9.2 se grafican algunos ejemplos para observar su desempeño. La ROC se construye anotando un punto para cada umbral (el umbral es variado de forma lineal barriendo 500 puntos entre 0 y 10000 W). A medida que el umbral aumenta, es menos probable que el algoritmo señale un encendido, por lo que la tasa de falsos positivos (FPR) disminuye.

La ROC mostrada en la Figura 9.1 presenta una forma particular, satura en valores levemente superiores a $TPR = 0.9$. Esta forma se debe a que el conjunto de test visto presenta pocos ejemplos y similares entre si, por lo que existe un umbral a partir del cual se detectan la gran mayoría de electrodomésticos, este umbral corresponde al

Capítulo 9. Evaluación sobre UK-DALE

queibre de los puntos en la figura. La AUC correspondiente es de 0,947, un valor que se puede considerar alto.

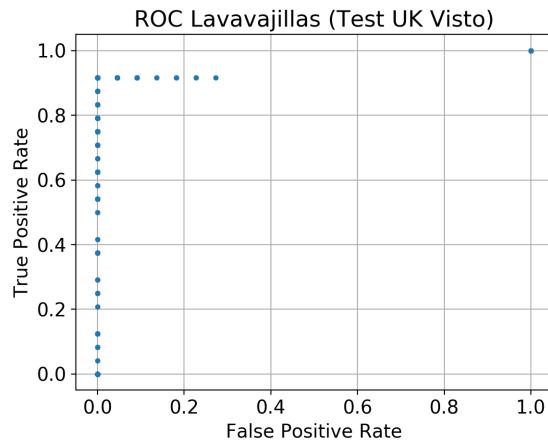


Figura 9.1: ROC Lavavajillas en test visto. $AUC = 0,9470$

En la Figura 9.2 se grafican tres ejemplos, en los cuales se observa que el auto-encoder presenta algo de ruido, y que detecta correctamente cuando se enciende el lavavajillas. Notar que el patrón de consumo en la primer y última figura es muy similar, pero las gráficas muestran dos ventanas distintas (observar el comportamiento de la curva azul antes del primer flanco).

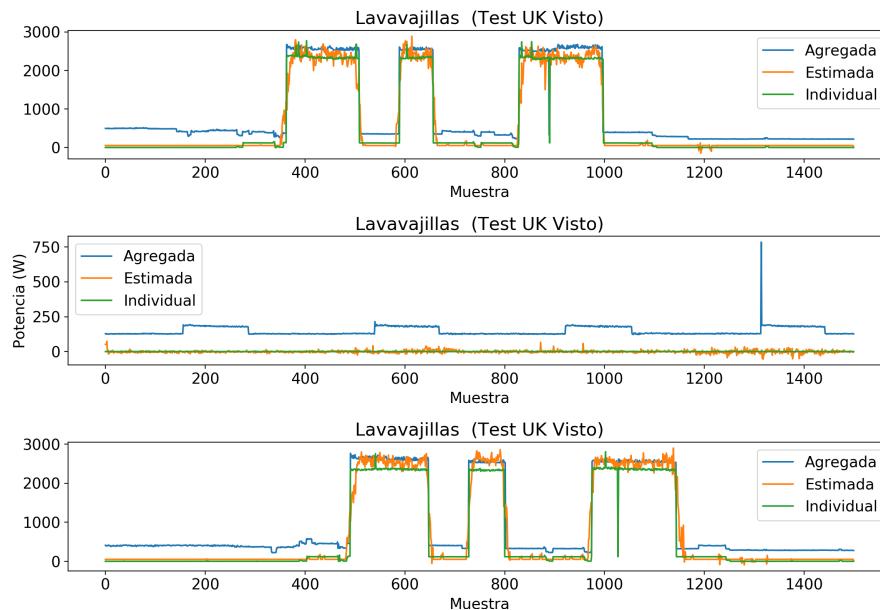


Figura 9.2: Resultados para el lavavajillas en test visto.

9.4. Comparación para test no visto usando activaciones

9.4. Comparación para test no visto usando activaciones

En esta sección se presentan los resultados sobre los datos correspondientes a las casas que no fueron utilizadas durante el entrenamiento. Este conjunto de datos continúa siendo balanceado, pero potencialmente incluye electrodomésticos distintos a los presentes durante el entrenamiento.

La Tabla 9.3 resume los resultados obtenidos al testear sobre casas distintas a las del entrenamiento. Recordar que el objetivo de esta evaluación es analizar la capacidad de generalización del modelo en forma aislada. Lo que se observa es que el mejor modelo no generaliza tan bien como cualquiera de los modelos base, con la excepción del autoencoder para la heladera. Esto indica que si bien el desempeño sobre test visto es bueno, el mismo cae al generalizar a datos que provienen de una distribución de probabilidad distinta a la que generó los datos de entrenamiento, es decir, a otras casas.

		Activaciones	
		Test Visto	Test No Visto
Jarra Eléctrica	Mejor modelo	0.999	0.983
	Autoencoder	0.959	<u>0.993</u>
	Rectángulos	0.941	0.994
Heladera	Mejor modelo	0.941	0.703
	Autoencoder	0.500	0.500
	Rectángulos	0.871	0.887
Lavarropas	Mejor modelo	<u>0.850</u>	0.795
	Autoencoder	0.884	0.908
	Rectángulos	0.812	<u>0.884</u>
Microondas	Mejor modelo	<u>0.976</u>	0.879
	Autoencoder	<u>0.932</u>	0.891
	Rectángulos	0.976	0.962
Lavavajillas	Mejor modelo	0.947	<u>0.962</u>
	Autoencoder	0.986	0.984
	Rectángulos	0.954	0.973

Tabla 9.3: AUC para todos los electrodomésticos en evaluación sobre activaciones. Comparación entre mejor modelo, rectángulos y autoencoder sobre test visto y no visto. Las negritas señalan la mejor red entre el mejor modelo, autoencoder base y rectángulos base, tanto para test visto como test no visto. El subrayado indica para cada red, si se tiene mayor AUC en test visto o test no visto.

Notar también que si bien los modelos base obtienen buenos scores de AUC, la selección del “mejor modelo” ya fue realizada, y la comparación se hace con el único objetivo de brindar una referencia. El proceso de obtención de un modelo culmina con la elección del mismo, intentando optimizar el desempeño en el conjunto de datos de validación. No debería ser esperable que el modelo generalice bien a casas no vistas cuando se lo entrenó con otro objetivo. En este sentido, sí es esperable que el desempeño se deteriore, como sucede con todos los electrodomésticos exceptuando el lavavajillas.

Gráficas

En la Figura 9.3 se muestra la ROC para el microondas y en la Figura 9.4 se grafican algunos ejemplos para observar su desempeño. El resto de los electrodomésticos se

Capítulo 9. Evaluación sobre UK-DALE

encuentran en el Apéndice E.

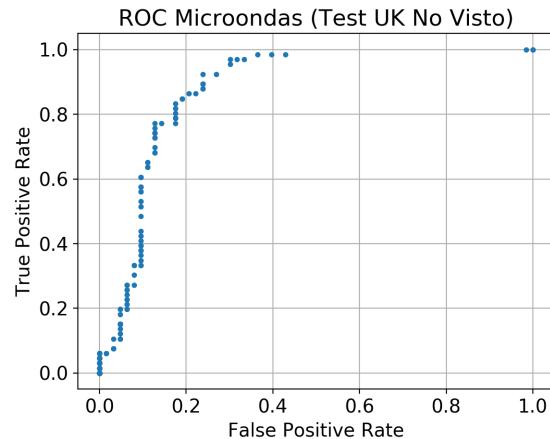


Figura 9.3: ROC Microondas en test no visto. $AUC = 0,8789$

En los ejemplos (Figura 9.4) se ve que el “mejor modelo”, el autoencoder de alta frecuencia, responde suavemente a las subidas en la curva agregada.

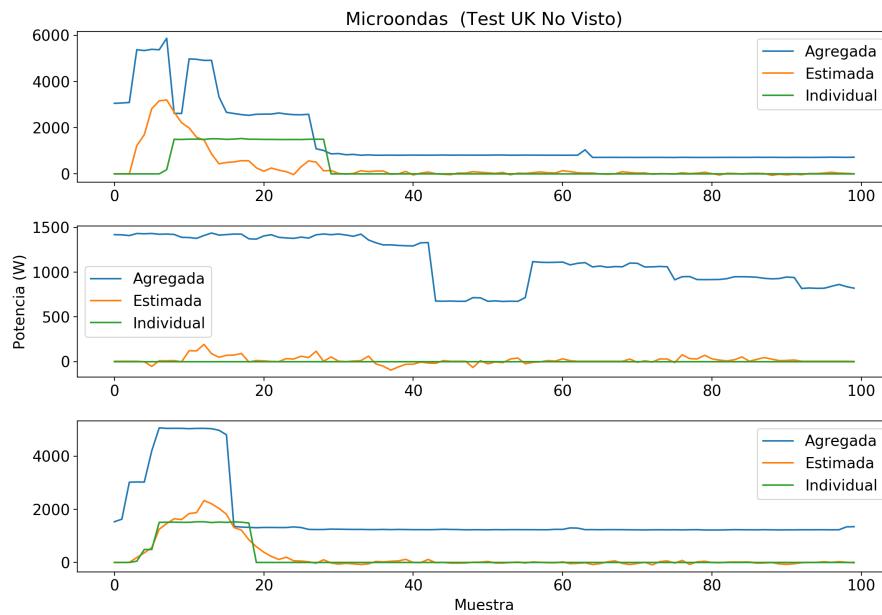


Figura 9.4: Resultados para el microondas en test no visto.

9.5. Influencia del mecanismo de evaluación

En esta Sección se modifica el mecanismo de evaluación y se procede a evaluar con ventanas deslizantes. El objetivo es ver cómo responde el mismo algoritmo o modelo,

9.5. Influencia del mecanismo de evaluación

evaluado sobre el mismo conjunto de test, a un cambio en el mecanismo de evaluación: de activaciones a ventana deslizante. La evaluación por ventanas deslizantes es la que ha de tenerse en cuenta a la hora de reportar resultados finales. Recordar que esta evaluación implica tomar toda la serie temporal que fue estimada mediante ventanas deslizantes de $STRIDE = 1$, y luego dividirla en ventanas no superpuestas sobre las cuales se evalúa. Una gran diferencia que se introduce en la evaluación mediante ventana deslizante respecto a la evaluación con activaciones es el desbalance que el primer método presenta en el conjunto datos. Esto es porque se evalúa sobre una serie agregada real, en la cual, exceptuando la heladera, los demás electrodomésticos permanecen apagados más tiempo del que están encendidos. La otra diferencia es que al estimar se realizan promedios entre algunos valores superpuestos de las ventanas, cuando sobre activaciones esto no sucede.

Para independizarse del error introducido por la generalización a otras casas, se evalúa el cambio sobre el mismo conjunto de evaluación, es decir, test visto mediante activaciones se compara con test visto mediante ventanas, y por otro lado, test no visto mediante activaciones se compara con test no visto mediante ventanas. Los resultados se resumen en la Tabla 9.4.

	Activaciones		Ventanas	
	Visto	No visto	Visto	No visto
Jarra eléctrica	0.999	0.984	1.000	0.998
Heladera	0.941	0.703	0.854	0.751
Lavarropas	0.850	0.795	0.763	0.864
Microondas	0.976	0.879	0.973	0.956
Lavavajillas	0.947	0.962	0.898	0.962

Tabla 9.4: AUC para el “mejor modelo” de cada electrodoméstico sobre UK-DALE

Los resultados demuestran que la independencia entre la generalización a casas distintas y el mecanismo de evaluación no es tal. Lo que se ve claro es que para casas conocidas, evaluar sobre activaciones reporta mejores desempeños que evaluar sobre ventanas deslizantes, pero que sucede lo inverso con el test no visto. La extensión al test no visto no solamente disminuye el desempeño sino que también hace que evaluar sobre ventanas deslizantes sea más favorable. Dicho de otro modo, la generalización completa, a casas distintas con datos desbalanceados, presenta menor deterioro que la generalización a casas distintas con datos balanceados.

Gráficas

Para el caso de datos de test visto, se utiliza como ejemplo a el microondas, su ROC se ilustra en la figura 9.5.

Capítulo 9. Evaluación sobre UK-DALE

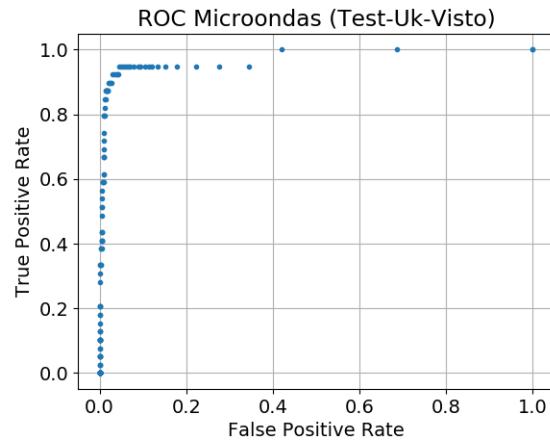


Figura 9.5: ROC Microondas Ventanas Deslizantes Test Visto. $AUC = 0,9731$

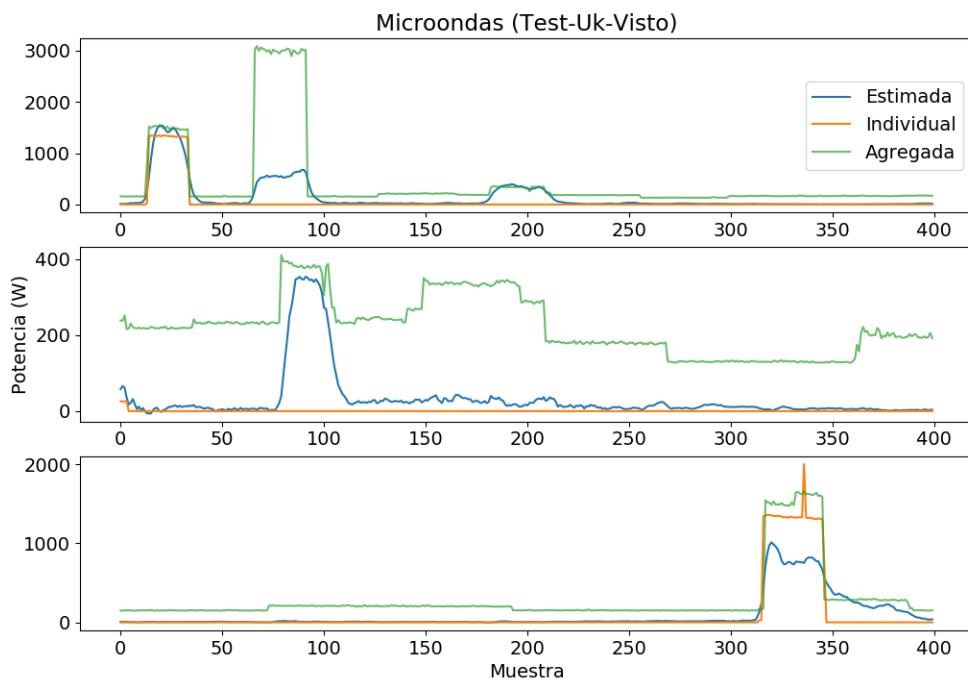


Figura 9.6: Ejemplo Microondas Ventanas Deslizantes Test Visto

9.6. Comparabilidad de implementaciones

Por último se nota que la AUC es sensible a datasets desbalanceados¹. El desempeño final de los modelos se evaluará en la Sección 9.7, incluyendo métricas de clasificación tradicionales.

9.6. Comparabilidad de implementaciones

Es de interés del Proyecto establecer comparaciones con la publicación de referencia [19]. Desgraciadamente, la comparación no es posible de forma directa, y solamente genera confusión. Algunos resultados presentados en la publicación de referencia son de difícil interpretación, como es el caso de la métrica REITE, con valores negativos. No se explica en el artículo cuál es el mecanismo de evaluación. Una comparación asumiendo que el mecanismo de evaluación es sobre activaciones y con los criterios del Proyecto se presenta en el Apéndice F. Asumir que la evaluación se realiza en ventanas deslizantes deja abierto el problema de cómo se define la estimación de un encendido o un apagado para las métricas de clasificación. Es probable que se trabaje con un umbral, pero además de ser desconocido, no se sabe sobre qué valor se aplica el umbral, máximo, media, mediana, etc.. Por último, la comparación sería injusta, ya que, debido a la extensión de la base de datos, este Proyecto utiliza una cantidad de datos que supera el doble de la utilizada en el trabajo de referencia.

9.7. Funcionamiento del sistema NILM sobre casas desconocidas del Reino Unido

Para este caso se muestra el comportamiento del modelo de la heladera en las Figuras 9.7 y 9.8. De la Tabla 9.4 se ve que la jarra eléctrica tiene un gran desempeño, mientras que la heladera presenta una AUC menor. Sin embargo, se ve que sigue las activaciones y que probablemente se desempeñe bien en clasificación.

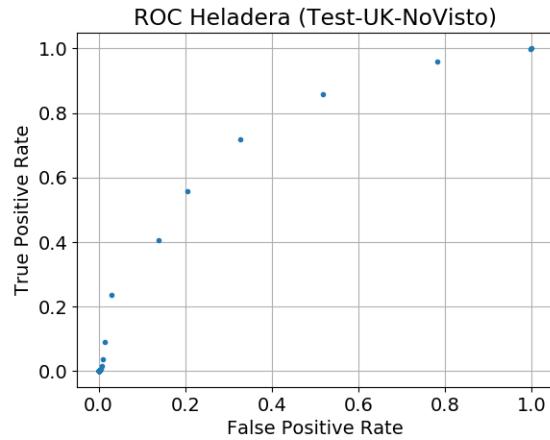


Figura 9.7: ROC Heladera Ventanas Deslizantes Test no Visto. $AUC = 0.8538$

¹Si con 1000 muestras, de las cuales 10 fueran positivas, existe algún umbral en el que se predicen correctamente estas 10 pero incorrectamente otras 90, se tendría $FPR = \frac{90}{90+990} = 0,083$ y $TPR = \frac{10}{10+0} = 1$, por lo que $AUC \geq 1 - 0,083 = 0,917$, aunque precision = 10 %

Capítulo 9. Evaluación sobre UK-DALE

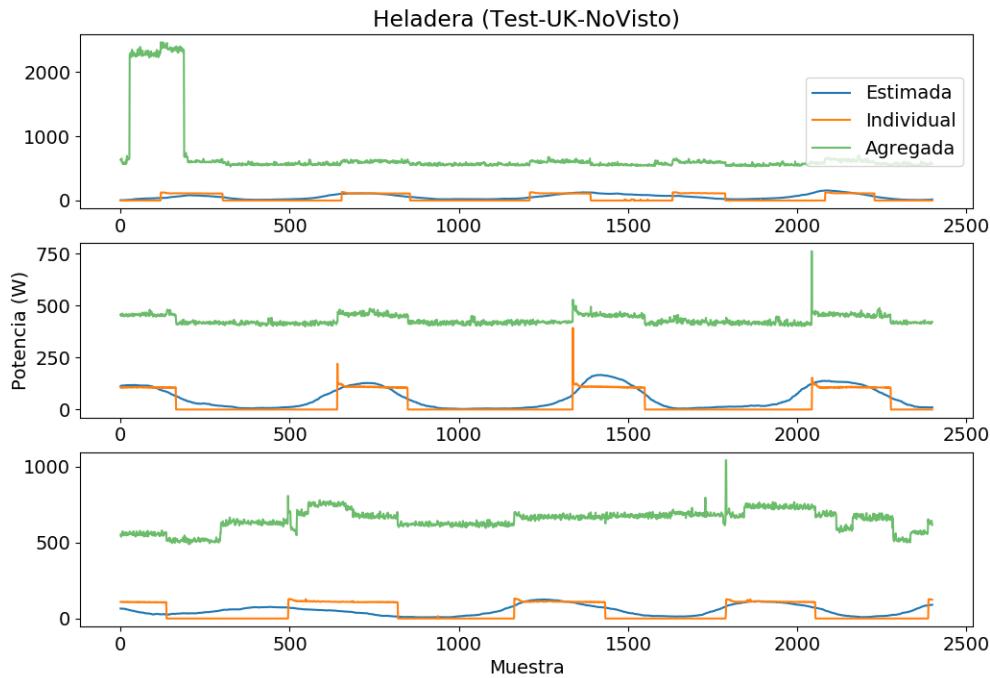


Figura 9.8: Ejemplo Heladera Ventanas Deslizantes Test no Visto

Las Tablas 9.5 y 9.6 presentan los resultados finales.

	Ventana deslizante	
Electrodoméstico	MAE (W)	REITE
Jarra eléctrica	22.48	0.609
Heladera	42.04	0.305
Lavarropas	237.98	0.962
Microondas	58.48	0.173
Lavavajillas	45.00	0.639

Tabla 9.5: Resultados de desagregación en test no visto de los mejores modelos en ventanas deslizantes.

9.8. Conclusiones

Electrodoméstico	Métrica			
	Accuracy	Precision	Recall	F1-Score
Jarra eléctrica	0.987	0.686	0.967	0.802
Heladera	0.585	0.545	0.959	0.695
Lavarropas	0.612	0.107	0.904	0.191
Microondas	0.835	0.019	0.964	0.038
Lavavajillas	0.961	0.679	0.743	0.710

Tabla 9.6: Métricas de clasificación en test no visto de los mejores modelos en ventanas deslizantes.

9.8. Conclusiones

En este Capítulo se realizaron diversas comparaciones y comentarios. Se reafirma que el “mejor modelo” alcanza el mayor desempeño para el conjunto de test visto para la mayoría de los electrodomésticos. Esto era esperable ya que este conjunto de datos es similar al de validación.

Esta superioridad no se mantiene, sin embargo, al pasar al conjunto de test no visto. En este conjunto todos los modelos presentan una capacidad de generalización similar, no hay un modelo que sea superior para todos los electrodomésticos. Estas comparaciones se realizan en el marco de evaluación sobre activaciones, resumido en la Tabla 9.3.

Otra comparación realizada es entre la evaluación por activaciones y la evaluación por ventanas deslizantes, que es cualitativamente distinta. Se ve que en este cambio de método de evaluación no se puede establecer una mejoría o deterioro del desempeño, ya que esto depende también de si los datos pertenecen a casas conocidas o no. La conclusión principal es que evaluar por activaciones da resultados más altos para el conjunto de test visto, pero da resultados peores en test no visto.

Por otro lado, para todos los electrodomésticos excepto el lavarropas, se tiene un MAE menor a 60W. Este valor corresponde al marco de evaluación más similar al real: ventanas deslizantes con datos de test no visto, presentado en la Sección 9.7. Además, en la Sección 9.6 se evalúa y se concluye que no es posible replicar los resultados experimentales del artículo [19] de Kelly debido a que no se cuenta con suficiente información.

En la Tabla 9.4 se presentan las AUCs para el “mejor modelo”. Esta tabla resume los resultados del Capítulo. La performance se deteriora, como es esperable, al pasar del conjunto de test de casas vistas al conjunto de test de casas no vistas, exceptuando el caso del lavavajillas. El lavarropas también presenta este fenómeno excepcional para el caso de la ventana deslizante. En resumen, la evaluación otorga resultados promisorios y los puntos de operación de clasificación que quedan definidos por la optimización del F1 en validación, demuestran un correcto funcionamiento sobre el caso real para 3 de 5 electrodomésticos (jarra eléctrica, heladera y lavavajillas). Para todos los electrodomésticos se tienen valores de recall altos lo que implica que rara vez no se detecta la presencia de un electrodoméstico cuando el mismo está funcionando. Para los electrodomésticos que tuvieron bajo F1, lo anterior parece indicar que casi siempre se indica que el electrodoméstico está presente aunque no sea el caso.

En el siguiente Capítulo se evaluará la capacidad de generalización de los algoritmos seleccionados sobre la base de datos NILM-UY.

Capítulo 9. Evaluación sobre UK-DALE

Capítulo 10

Evaluación sobre datos de Uruguay

10.1. Resumen

Una de las preguntas que estuvo presente durante todo el desarrollo de este trabajo fue si es factible que los modelos entrenados con la base UK-DALE [18] generalizaran a los datos de Uruguay. La respuesta a esa pregunta se incluye en este capítulo: se realizó una evaluación de los mejores modelos para cada electrodoméstico sobre los datos de NILM-UY. La evaluación fue realizada sobre la potencia desagregada utilizando el enfoque de ventana deslizante previamente explicado. Se recuerda al lector que este enfoque consiste en tomar una ventana deslizante sobre la potencia agregada y comparar la salida de la red con la potencia desagregada del electrodoméstico en cuestión.

Los resultados obtenidos son comparados con los reportados con el enfoque de ventana deslizante sobre el *dataset* UK-DALE en el capítulo previo, utilizando el mismo largo de ventana y el mismo umbral de potencia.

10.2. Procesamiento de los datos para la evaluación

El *dataset* NILM-UY está conformado por dos conjuntos de datos: señales agregadas de alta frecuencia y desagregadas. Como se detalla en el Capítulo 4, las medidas de alta frecuencia (14 kHz) consisten en archivos .flac que en un canal contienen la medida de voltaje y en el otro la de corriente. Por otro lado, las medidas individuales de cada electrodoméstico son valores de potencia activa cada 1 minuto guardadas en formato .csv.

El procesamiento de las señales agregadas fue muy similar al realizado sobre los datos de alta frecuencia de UK-DALE explicado en el Capítulo 6. A partir de las medidas de voltaje y corriente, se calcularon la potencia y los descriptores: factor de forma y desfasaje. Para que las señales fueran compatibles con los modelos entrenados, se crearon series de potencia y de los descriptores de alta frecuencia con datos cada 6 segundos. Para los modelos de alta frecuencia la entrada a las redes fue la serie multivariada, mientras que para el modelo de baja frecuencia (el del lavavajillas) la entrada era solamente la serie de potencia (una de las tres componentes de la serie multivariada).

Los datos de potencia individuales de cada electrodoméstico tuvieron que ser pasados por un bloqueador de orden 0 de forma de obtener series con datos cada 6 segundos, para que la salida de las redes sea comparable con los datos de cada electrodoméstico.

Capítulo 10. Evaluación sobre datos de Uruguay

Este procesamiento, debido a que los datos individuales de los electrodomésticos tienen una frecuencia de muestreo tan lenta, introduce desfasajes no despreciables, de hasta 1 minuto, entre la señal agregada y desagregada, que pueden perjudicar el desempeño de los algoritmos. Otra estrategia posible es realizar algún tipo de interpolación de mayor orden entre los datos.

10.3. Disponibilidad de datos

Para la Casa 1 (Casa Franco) de la base de datos se contó con todos los datos correspondientes al período citado en el Capítulo 4, sin embargo para la Casa 2 (Casa Pablo), por un comportamiento no previsto en la desconexión - reconexión del medidor, que generó datos ruidosos, no se utilizaron gran parte de los datos del final del segundo período. Al igual que fue explicado en la Sección 7.3.3, la escasez de datos a la hora de evaluar atenta contra el proceso. Además, por la naturaleza de cada uno de los hogares, pueden tenerse más o menos instancias de consumo de cada electrodoméstico.

En resumen, si bien se cuenta con una base de datos de validación, la misma dista de tener la cantidad de instancias de electrodomésticos deseable para una evaluación profunda. Sin embargo, dados los resultados que se observarán a continuación, se entiende que la base de datos permite obtener una percepción del poder de generalización de los algoritmos y los inconvenientes que los mismos pueden presentar.

10.4. Evaluación

En lo que sigue se presentarán algunos ejemplos gráficos de la salida de los algoritmos, la curva ROC y una tabla completa de métricas para cada casa. Solamente se incluye un ejemplo gráfico para cada casa para no entorpecer la lectura, se sugiere al lector interesado consultar el apéndice E para ver ejemplos de todos los electrodomésticos.

10.4.1. Casa 1 (Franco)

Para la Casa 1 se grafican ejemplos y la ROC para el lavarropas. Al visualizar las Figuras 10.1 y 10.2, se tiene una pauta de que el algoritmo no es para nada immune al cambio de electrodomésticos a la hora de evaluar. Se obtiene una ROC marginalmente mejor que si se hiciera una predicción al azar y los ejemplos obtenidos muestran que no se reconstruye bien la señal.

10.4. Evaluación

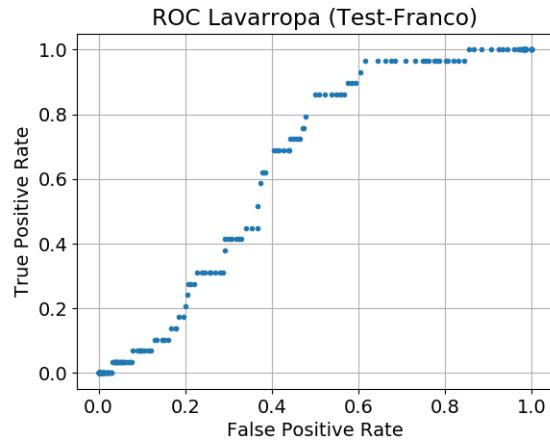


Figura 10.1: ROC Lavarropa Ventanas Deslizantes Test casa Franco. $AUC = 0,6467$

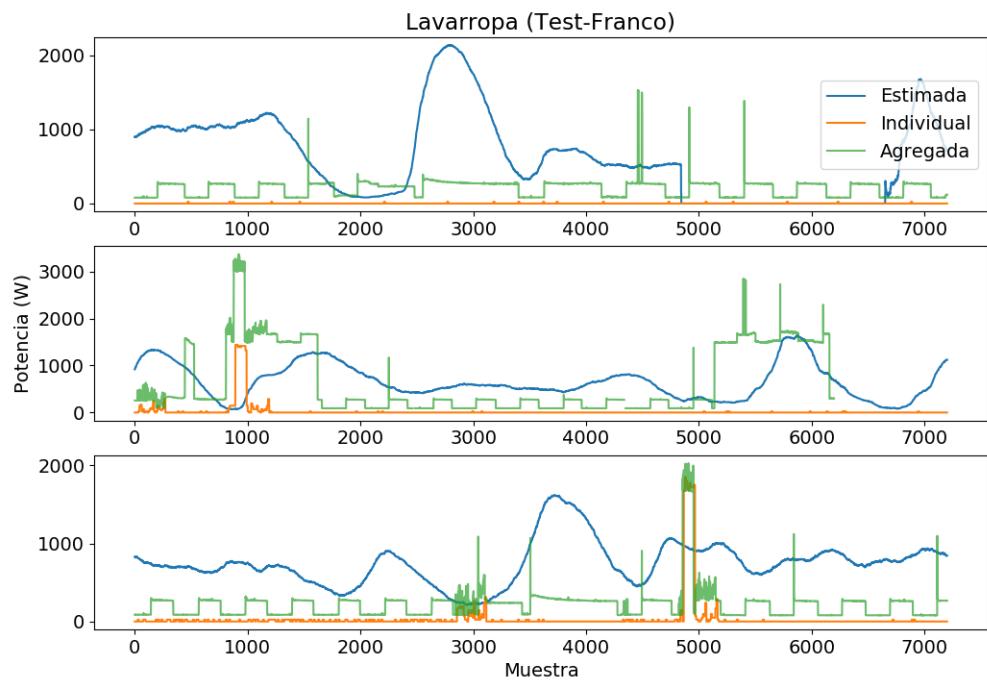


Figura 10.2: Ejemplo Lavarropa Ventanas Deslizantes Test casa Franco

Observando las métricas para todos los electrodomésticos de la casa: heladera, lavarropas y lavavajillas; lo intuido previamente se mantiene. Existe una clara degradación en los resultados.

Capítulo 10. Evaluación sobre datos de Uruguay

En el apartado de desagregación todas las métricas empeoran. Notar que si bien en el caso del microondas el MAE no aumenta tanto, dado que el REITE tiene un valor tan alto se puede considerar que el desempeño es muy bajo.

	Casa Franco (NILM-UY)		Test No Visto (UK-DALE)	
Electrodoméstico	MAE (W)	REITE	MAE (W)	REITE
Heladera	143	0.507	42	0.305
Lavarropas	793	0.995	237	0.962
Microondas	71	0.920	58	0.173

Tabla 10.1: Resultados de desagregación sobre la casa 1 de NILM-UY comparados con los resultados obtenidos en ventanas deslizantes Test No Visto UK-DALE.

Elec.	Casa Franco (NILM-UY)				Test No Visto (UK-DALE)			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
Heladera	0.759	0.781	0.959	0.861	0.585	0.545	0.959	0.695
Lavarropas	0.071	0.057	1.000	0.107	0.612	0.107	0.904	0.191
Microondas	0.506	0.018	0.650	0.036	0.835	0.019	0.964	0.038

Tabla 10.2: Resultados de clasificación sobre la casa 1 de NILM-UY comparados con los resultados obtenidos en ventanas deslizantes Test No Visto UK-DALE.

En cuanto a la clasificación, se realiza el análisis poniendo el foco sobre el F1-Score ya que es bastante representativo del desempeño general del algoritmo. Considerando esto, para la heladera se obtiene (sorpasivamente) un mejor score y para los otros dos electrodomésticos el desempeño empeora. Notar que estos resultados dependen del umbral de potencia y del largo de ventana considerado. Para todas las evaluaciones de este capítulo se mantuvieron estos valores respecto a los utilizados en UK-DALE. Queda abierta la pregunta, como trabajo futuro, de cómo se verían afectadas las métricas al considerar largos de ventana y umbrales que se ajusten al funcionamiento específico de los electrodomésticos evaluados en Uruguay, que no necesariamente tienen que ser similares a los del Reino Unido.

Observando el resto de los ejemplos gráficos (incluidos en el Apéndice E), parece ser que las redes no están generalizando. La señal reconstruida muchas veces sigue a la señal agregada independientemente de si la señal del electrodoméstico de interés está presente o no, esto puede causar que en electrodomésticos que siempre están funcionando (heladera) los valores de las métricas sean buenos, esto es debido a que se tiene una buena chance de que el electrodoméstico esté encendido en la mayoría de las ventanas sobre las que se testeó. Por otro lado, en el caso del microondas se tiene un desempeño particularmente malo.

10.4.2. Casa 2 (Pablo)

Para la Casa 2, se muestran en esta sección los resultados obtenidos para el lavavajillas.

10.4. Evaluación

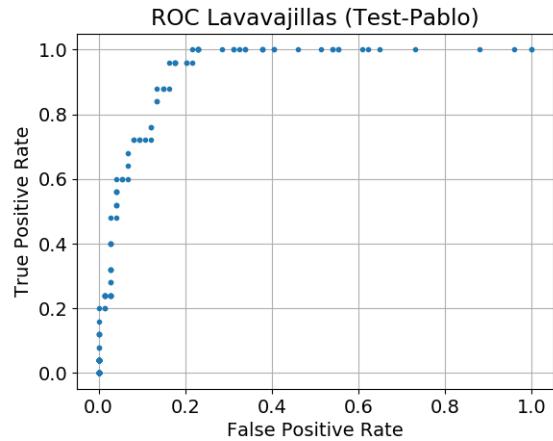


Figura 10.3: ROC Lavavajillas Ventanas Deslizantes Test casa Pablo. $AUC = 0,9376$

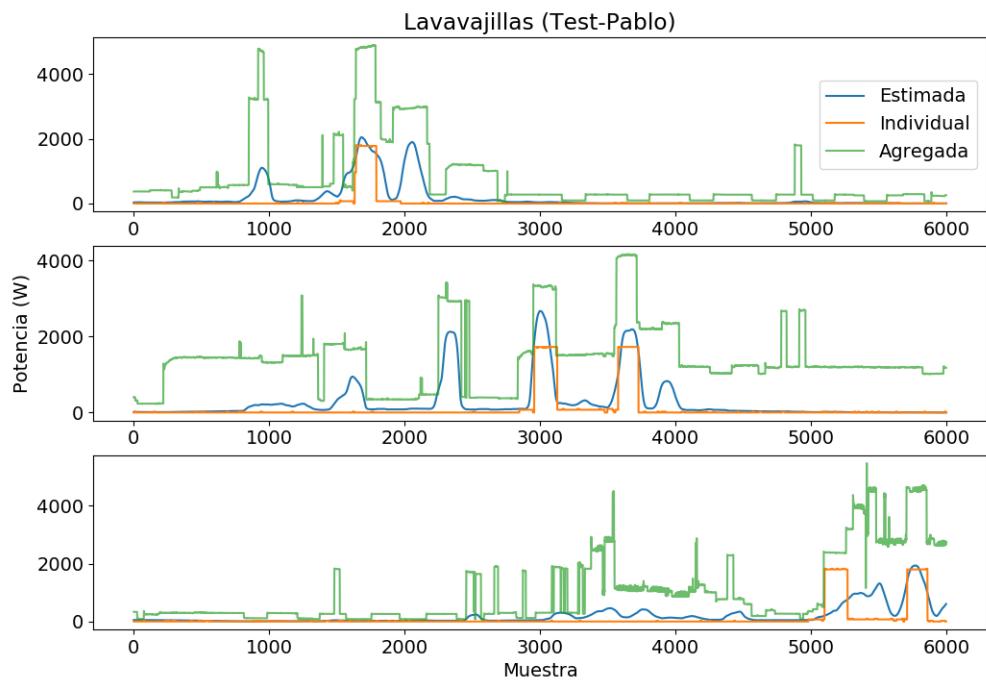


Figura 10.4: Ejemplo Lavavajillas Ventanas Deslizantes Test casa Pablo

En este electrodoméstico (que no fue evaluado en la Casa 1 por no estar), los resultados gráficos son mucho más auspiciosos. Según los ejemplos de la figura 10.4, si bien a veces se confunde, la red logra reconstruir bastante bien a la señal individual.

Capítulo 10. Evaluación sobre datos de Uruguay

A su vez, a diferencia del lavarropas de la Casa 1, para el lavavajillas de la Casa 2 la ROC obtenida muestra un buen desempeño.

Electrodoméstico	Casa Pablo (NILM-UY)		Test No Visto (UK-DALE)	
	MAE (W)	REITE	MAE (W)	REITE
Jarra eléctrica	75	0.909	22	0.609
Heladera	96	0.071	42	0.305
Lavarropas	691	0.971	238	0.962
Microondas	90	0.880	58	0.173
Lavavajillas	150	0.478	45	0.639

Tabla 10.3: Resultados de desagregación sobre la casa 2 de NILM-UY comparados con los resultados obtenidos en ventanas deslizantes Test No Visto UK-DALE.

En esta casa se tiene el conjunto completo de electrodomésticos a evaluar. En cuanto a la desagregación, el desempeño vuelve a verse impactado por el cambio de los electrodomésticos, pero en menor medida: el REITE del lavavajillas y la heladera es sorprendentemente mejor que en UK-DALE.

Elec.	Casa Pablo (NILM-UY)				Test No Visto (UK-DALE)			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
Jarra eléctrica	0.953	0.286	0.545	0.375	0.987	0.686	0.967	0.802
Heladera	0.918	0.990	0.926	0.957	0.585	0.545	0.959	0.695
Lavarropas	0.323	0.300	1.000	0.462	0.612	0.107	0.904	0.191
Microondas	0.533	0.066	0.818	0.122	0.835	0.019	0.964	0.038
Lavavajillas	0.859	0.720	0.720	0.720	0.961	0.679	0.743	0.710

Tabla 10.4: Resultados de clasificación sobre la casa 2 de NILM-UY comparados con los resultados obtenidos en ventanas deslizantes Test No Visto UK-DALE.

Considerando el F1-Score, la degradación de los resultados no es tan evidente. Solo se observa un deterioro muy marcado en la jarra eléctrica. En el resto de los electrodomésticos el desempeño es mejor incluso considerando otras métricas como la precisión. El caso de la heladera es particularmente destacable donde para todas las métricas, salvo el recall donde tiene un resultado apenas inferior, los resultados son muy superiores a los observados en UK-DALE.

	Ventanas			
	Visto UK	No visto UK	Pablo UY	Franco UY
Jarra eléctrica	1.00	0.998	0.830	-
Heladera	0.854	0.751	0.483	0.435
Lavarropas	0.763	0.864	0.808	0.647
Microondas	0.973	0.956	0.848	0.615
Lavavajillas	0.898	0.962	0.938	-

Tabla 10.5: Comparativa de AUC en ventanas deslizantes sobre las casas de NILM-UY.

10.5. Conclusiones

10.5. Conclusiones

En base a los resultados presentados en este capítulo, parece razonable concluir que los algoritmos propuestos no tienen un gran poder de generalización sobre electrodomésticos nunca vistos. Más aun, se observa que el desempeño se ve afectado cuando se consideran electrodomésticos posiblemente muy distintos, ya que se entrenó con electrodomésticos del Reino Unido, y ahora se evaluó sobre electrodomésticos comprados en Uruguay. Observando la Tabla 10.5, si bien puede verse como se degrada el desempeño al cambiar el conjunto de electrodomésticos, es sumamente destacable que para 4 de 5 electrodomésticos de la Casa Pablo, se obtenga una AUC mayor a 0,8. Esto indica que a pesar de la degradación o del bajo poder de generalización en el sentido amplio del problema, se podría encontrar umbrales que permitan a modelos entrenados con datos de unas pocas casas del Reino Unido detectar de forma razonable la presencia de electrodomésticos de Uruguay. En particular, para la heladera, el umbral obtenido en validación de UK-DALE arrojó muy buenos resultados en la Casa Pablo. Lo anterior abre la interrogante para trabajos futuros de buscar umbrales y tamaños de ventana más adecuados para los datos de Uruguay.

En cuanto a las métricas de desagregación, particularmente el MAE, se concluye que empeoran drásticamente y los algoritmos no generalizan adecuadamente en este sentido. Los resultados observados van en línea con el fenómeno observado para algunos casos de UK-DALE, al evaluar sobre electrodomésticos desconocidos, la *performance* de los algoritmos suele verse afectada.

Como cierre de este capítulo, se puede comenzar a esbozar una conclusión final: los algoritmos presentados son muy dependientes de los electrodomésticos sobre los que se entrenó y no generalizan a desagregar electrodomésticos cuyas curvas de potencia sean muy distintas a las de entrenamiento.

Capítulo 10. Evaluación sobre datos de Uruguay

Capítulo 11

Conclusiones y perspectivas

11.1. Resumen

En este Capítulo se presentan las principales conclusiones del Proyecto. Estas conclusiones abarcan las diversas aristas que conforman el Proyecto, cuyo objetivo fue realizar una solución punta a punta para el problema de desagregación del consumo de electrodomésticos. La Sección 11.3 es el lugar donde los autores plasman recomendaciones que creen pertinentes para trabajos futuros. La intención es nombrar las posibles mejoras o los enfoques distintos que podrían haberse realizado. El objetivo es aportar una visión de qué consideran los autores que es lo más promisorio para resolver el problema que motiva el Proyecto.

11.2. Conclusiones

En primer lugar, como conclusión general se logró implementar de forma satisfactoria una solución punta a punta al problema de NILM, en forma de *Minimum Viable Product* (MVP). Esta solución involucró desarrollo de hardware para el medidor, de software tanto para el medidor como para la etapa de algoritmos de desagregación y fue puesta a prueba de punta a punta durante las distintas etapas del Proyecto.

En lo que respecta al diseño del hardware del sistema de medida se destaca su robustez, suficiente como para recolectar medidas en forma desatendida por períodos de tiempo de varios meses. Además, el medidor es adecuado para el relevamiento de señales de voltaje y corriente en el marco de otros problemas o con otros objetivos. Dado que se diseñó con la idea de que la frecuencia de muestreo no sea una limitante, existe la posibilidad, previo ajuste de la configuración del software de medida, de seleccionar cualquier otra frecuencia de adquisición por debajo de los 96 kHz. También es posible ajustar la cuantización de los datos, pudiendo elegirse tomar muestras de 32, 16 u 8 bits. Más allá de la limitante en frecuencia impuesta por la tarjeta adquisidora, se advierte que las componentes de alta frecuencia pueden verse degradadas por la pinza de corriente, como se explica en el Apéndice B. Este análisis resulta un aporte en sí mismo ya que se encontró en la literatura un solo estudio de caracterización de la pinza de corriente que no hace referencia a la respuesta en frecuencia [40] y otro del mismo autor que sí lo hace de forma muy sucinta, pero que refiere a modelos anteriores de la misma [41].

A partir del relevamiento del estado del arte, se propuso una estrategia de desagregación basada en redes neuronales que actúen sobre una ventana deslizante sobre

Capítulo 11. Conclusiones y perspectivas

la serie agregada de potencia y den como salida la potencia desagregada de cada electrodoméstico. Por el funcionamiento de la ventana deslizante, a la hora de desagregar una señal se necesita al menos una cantidad de datos igual al tamaño de la ventana del electrodoméstico, por lo que este enfoque no puede ser utilizado estrictamente en tiempo real. Sin embargo, dado que los tiempos de inferencia en una ventana son bajos, aún corriendo en CPU, los algoritmos pueden utilizarse para desagregar *casi* en tiempo real, por ejemplo desagregando los últimos minutos de consumo de un hogar.

Se implementaron dos tipos de arquitecturas tomadas de un paper de referencia [19], un autoencoder y una red que estima un rectángulo que aproxima la curva de consumo de un electrodoméstico. Sobre estas arquitecturas base, se generó y evaluó una variante que incluyera entradas multidimensionales así como también una variante más profunda del autoencoder. Las entradas multidimensionales se construyeron con características de la serie agregada medida a alta frecuencia, en base a un análisis previo de búsqueda de características de alta frecuencia. Este estudio, realizado en el Capítulo 5, es también un aporte del Proyecto. En primer lugar, se realizó un estudio sobre datos de alta frecuencia individuales de un electrodoméstico de las distintas características expertas que podían extraerse y se diseñó un clasificador de electrodomésticos que alcanzó resultados similares a los del estado del arte en ese problema. En base al estudio de las características se logró determinar cuales tenían redundancia entre si, cuales lograban obtener mejores resultados, se las analizó utilizando el criterio de información mutua y se definió, según requerimientos explicados en ese capítulo, que se utilizará la fase y el factor de forma en las series multivariadas en la tarea de desagregación.

Haciendo foco en los resultados de los algoritmos en la tarea de desagregación, se concluye que los algoritmos presentados tienen una muy buena capacidad de desagregar el consumo de electrodomésticos que fueron vistos durante el entrenamiento. Este resultado coincide con lo que sucede con algunos dispositivos comerciales, como por ejemplo Smappee, que durante el primer mes de funcionamiento no realiza predicciones y luego le pide etiquetas al usuario para lograr identificar sus electrodomésticos. O también el estadounidense Sense, que indica en su sitio web que sus algoritmos logran identificar electrodomésticos una vez que han visto suficientes datos de los mismos y que la cantidad de electrodomésticos que puede identificar continuamente aumenta en base a los datos recolectados sobre todos sus usuarios. En cuanto al poder de generalización de los algoritmos presentados, según los resultados observados se concluye que no logran generalizar de forma adecuada sobre electrodomésticos que no son vistos en el entrenamiento. Si bien hay que tener en cuenta las limitantes existentes en los datos de entrenamiento, ya que solo se entrena con electrodomésticos de unas pocas casas, 4 como máximo, puede observarse claramente una degradación en los resultados cuando se evalúa sobre el conjunto de Test No Visto. Esta degradación se hace más notoria para los electrodomésticos de Uruguay de la base NILM-UY, aunque para una de las casas se obtienen valores de AUC mayores a 0,8, que se consideran más que aceptables en este contexto. Aquí podemos esbozar dos hipótesis para explicar la degradación del desempeño, en primer lugar es probable que las huellas de consumo entre electrodomésticos distintos de UK sean más similares entre sí que las de electrodomésticos de UK y de Uruguay. En segundo lugar la base NILM-UY no cuenta con un período de tiempo largo como para una evaluación profunda, similar a la realizada en UKDALE, por lo que quizás esta evaluación no es tan representativa. En resumen, se concluye que el enfoque escogido es dependiente de los electrodomésticos sobre los que se entrena, ya que depende de la forma de las activaciones de cada electrodoméstico. Los mejores resultados se obtuvieron en las métricas de REITE y AUC. Un análisis más detallado de los resultados muestra que no existe un único modelo que sea mejor para todos los electrodomésticos, en concordancia con el comentario anterior, ya que la forma de las

11.3. Perspectivas y trabajo a futuro

activaciones entre electrodomésticos es muy distinta.

En cuanto a las métricas propuestas en la literatura para evaluar desempeño, se analizaron dos métricas de regresión, MAE y REITE y las métricas de clasificación usuales además de la ROC y la AUC. En base a lo desarrollado durante el Proyecto puede concluirse que, si bien a priori las métricas de regresión son importantes para el problema, ya que miden errores en términos de potencia o energía asignada, por sí solas no dan una buena medida del desempeño del algoritmo. Se observaron casos en los que se daba una salida constante pero la MAE no daba un valor alto, por otro lado, también hubo casos en los que la MAE era mayor pero al observar las gráficas se veía que el algoritmo estaba desagregando de forma razonable la señal. El análisis de la MAE además debe hacerse teniendo en cuenta el consumo típico de cada electrodoméstico para ponerlo en perspectiva. En esta tarea, las métricas de clasificación dependen de forma directa del umbral elegido para decidir si hay o no un electrodoméstico. Para independizarse de esta elección se decidió utilizar la AUC como principal métrica de clasificación. Observando también que en los ejemplos que se tenía un valor elevado de AUC, las gráficas mostraban que el algoritmo estaba reconstruyendo bien la señal desagregada, se concluye que la AUC es una métrica adecuada de clasificación para NILM.

Sobre la utilidad de las características de alta frecuencia y la modificación a los modelos de referencia puede concluirse que estas variantes permitieron un mejor ajuste a los datos, es decir, un mejor desempeño en validación. Sin embargo, al evaluar los modelos en los distintos conjuntos de test, si bien muchas veces los mejores resultados, tanto en clasificación como en desagregación, eran arrojados por estos “mejores modelos”, también son varios los casos en los que los modelos de referencia eran los que obtenían mejores resultados.

Además de la evaluación de los modelos implementados se realizó una comparación con los resultados de referencia [19]. Esta comparación se realizó por completitud, pero se entiende que para poder hacerla de forma justa habría que acceder al protocolo completo del experimento realizado por los autores. En este Proyecto se ha tratado de salvar esta restricción para que trabajos futuros puedan tomar los resultados aquí expuestos y realizar comparaciones adecuadas. Esto se hizo explicitando totalmente el protocolo de entrenamiento y evaluación, generando un marco que puede ser tomado como referencia por futuros trabajos. Los modelos fueron implementados en el reciente TensorFlow 2, un ecosistema estable al mediano plazo.

11.3. Perspectivas y trabajo a futuro

Sobre los trabajos a futuro que puedan continuar en esta línea, se considera que podría realizarse un análisis más exhaustivo de los experimentos y del entrenamiento de las redes neuronales, considerar otras bases de datos públicas como WHITED [15], REDD [20] y BLUED [10] para tener más datos de entrenamiento y evaluación. También se podría incorporar características externas, como ser la hora del día, el mes o la temperatura exterior para brindarle más información a los modelos. Si se quisieran obtener resultados que superen el estado del arte actual se podría considerar un problema más limitado en alcance, focalizándose en algo menos general, como puede ser un tipo de electrodoméstico en específico o un enfoque que no requiera desarrollar un modelo para cada electrodoméstico, ya que esto multiplica los problemas y divide los éxitos a la hora de llevar a cabo el proyecto. En el anexo G se realiza un análisis de distintos enfoques: supervisados, no supervisados, semi-supervisados, de fusión de clasificadores que podrían seguirse a futuro.

Dado los buenos resultados en clasificación sobre electrodomésticos conocidos, una de las aplicaciones que UTE podría extraer de este Proyecto es un detector de calefones

Capítulo 11. Conclusiones y perspectivas

a partir de las medidas de sus *smart meters*. Entrenando una red de rectángulos como la del modelo base o un autoencoder base o su variante profunda, que tenga como entrada medidas 15 minutales, que corresponden al promedio de 900 muestras, y que tenga como *targets* las medidas de los medidores *smart plug* que tiene UTE ya instalados. Si bien en el Proyecto no se ha testeado el desempeño de algoritmos con esta tasa de muestreo, sí se comprobó que el entrenamiento basado en los medidores de calefones permite utilizar aprendizaje automático y que las redes tienen potencial a la hora de aprender funciones arbitrarias en el marco de este problema.

Sobre los trabajos a futuro que puedan continuar en esta línea, se considera que podría realizarse un análisis más exhaustivo de los experimentos y del entrenamiento de las redes neuronales, considerar otras bases de datos públicas como WHITED [15], REDD [20] y BLUED [10] para tener más datos de entrenamiento y evaluación. También se podría incorporar características externas, como ser la hora del día, el mes o la temperatura exterior para brindarle más información a los modelos. Si se quisieran obtener resultados que superen el estado del arte actual se podría considerar un problema más limitado en alcance, focalizándose en algo menos general, como puede ser un tipo de electrodoméstico en específico o un enfoque que no requiera desarrollar un modelo para cada electrodoméstico, ya que esto multiplica los problemas y divide los éxitos a la hora de llevar a cabo el proyecto. En el anexo G se realiza un análisis de distintos enfoques: supervisados, no supervisados, semi-supervisados, de fusión de clasificadores que podrían seguirse a futuro.

11.4. Conclusiones de la asignatura

En lo que se refiere a los objetivos de la asignatura Proyecto de Fin de Carrera, consideramos que se han cumplido con los mismos. Se abordó un trabajo integrador, que fue transversal a distintas ramas de la ingeniería eléctrica como lo son la electrónica, telecomunicaciones, procesamiento de señales, potencia y aprendizaje automático. Implicó un fortalecimiento de habilidades de diseño hardware, desarrollo de software, desarrollo y evaluación de modelos. Se creció mucho en lo referido al manejo de proyectos en general, entendiendo que un buen manejo de plazos, requerimientos y motivación es crucial, así como lo valioso de conocer los pasos a seguir y resultados a esperar. También se fortalecieron habilidades blandas como lo son el trabajo en equipo, la interacción con especialistas y la documentación.

Apéndice A

Circuitos de referencia

En este Apéndice se presentará el trabajo del Ing. Juan Pablo Maggi y del Phd. Jack Kelly en el diseño de circuitos de medida de voltaje y corriente de la red eléctrica.

Círculo del Ing. Juan Pablo Maggi

En UTE ya había experiencia acumulada en el diseño e implementación de sistemas de medición. En particular, el Ing. Maggi desarrolló una solución que se presenta en la figura A para tomar medidas de la red. Este bloque acondicionador forma parte de un medidor inteligente de su autoría.

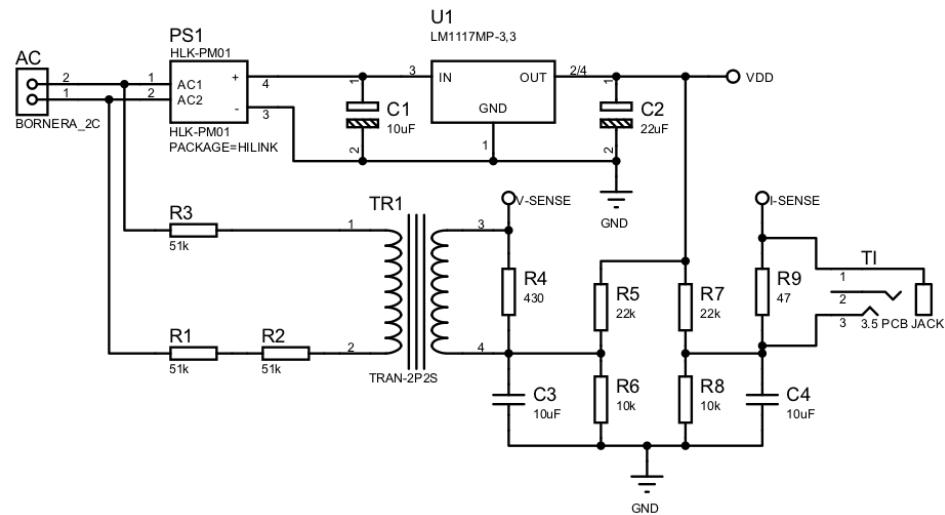


Figura A.1: Bloque acondicionador de señal del WiFi Meter, diseñado por el Ing. Maggi.

Los puntos de lectura del ADC son V-SENSE (para el voltaje) e I-SENSE (para la corriente). Las entradas son: para la corriente, la misma pinza utilizada por el grupo que se conecta en el Jack de audio, y para el voltaje, la bornera AC.

Esta solución consiste en una rama que genera un offset, sobre el cual se agregan las señales a medir (proporcionales a la corriente y al voltaje). Esta primera rama se implementa mediante un conversor AC/DC que alimenta un regulador de tensión,

Apéndice A. Circuitos de referencia

usado para fijar el voltaje de referencia a través de un divisor resistivo. La necesidad de generar un offset y agregárselo a las señales a medir, fue impuesta en este caso por el ADC elegido, cuyo rango de entrada solo admite señales positivas.

A este voltaje de referencia se le suma, por un lado, la salida de un transformador de tensión que sirve de aislación galvánica. La relación entre la entrada y la salida del bloque del circuito correspondiente al transformador viene dada por un divisor resistivo. De esta forma, se obtiene la lectura del voltaje de la red. Por otro lado, al voltaje de referencia también se le suma el voltaje generado en una resistencia debido a la corriente de la red (escalada por la pinza de corriente), obteniéndose así la lectura de corriente de la red.

Para la adición del offset a las señales a medir, se utilizan condensadores de desacople, y el circuito también incluye un pequeño filtro para reducir el ruido de la señal de offset.

Circuito de Jack Kelly

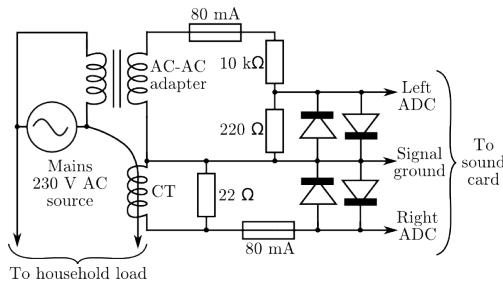


Figura A.2: Bloque acondicionador de señal del sistema adquisidor de Jack Kelly

El circuito de Jack Kelly, presentado en la figura A y usado para recabar la base de datos UK-DALE [18], consiste en un transformador de corriente, uno de voltaje, resistencias, fusibles y diodos Zener. En particular, adapta la entrada a una salida de 0,7 V.

Para medir el voltaje de forma segura, utilizó un transformador AC/AC. La salida pico de circuito abierto de 11V. Basarse en los estudios de Open Energy Monitor le permite afirmar que el transformador es lineal en esa zona. Este voltaje se reduce a 0,7 V a través de un divisor resistivo. Por otro lado, para medir la corriente, utiliza la pinza de corriente SCT-013-000. Esta pinza se conecta en paralelo a una resistencia que permite $0,89V_{peak}$ a la salida. Para proteger a la tarjeta de sonido a sobrecargas, se incluye en ambos canales fusibles de 80mA y un par de diodos tradicionales con 1,3V de *forward voltage*.

Apéndice B

Pinza de corriente

B.1. Resumen

Las pinzas de corriente (*current clamps*) son instrumentos de medida que entregan una corriente o voltaje proporcional a la corriente que pasa por un cable. Siendo un elemento importante en la creación de un medidor de corriente no intrusivo, en este capítulo se presentan algunas de sus características. En particular, se reportan los experimentos realizados sobre la pinza de corriente YHDC SCT013-000 CT con el objetivo de relevar su respuesta en frecuencia. Para estas pruebas se utiliza la pinza, un banco de carga, un generador de ondas y un osciloscopio. Los resultados muestran que, experimentando bajo el supuesto de que el banco de resistencias utilizado tiene una respuesta en frecuencia plana, la pinza tiene un comportamiento atenuante en frecuencias altas. Además, se extraen otras conclusiones al considerar que la resistencia presentada por el banco de cargas depende de la frecuencia. Una discusión acerca del método, las hipótesis y su alcance es presentada junto con las conclusiones.

B.2. Introducción

En el contexto de los objetivos del proyecto, se hace necesario seleccionar una pinza de corriente que permita adquirir la señal con una respuesta adecuada y que, además, sea de bajo costo.

Existen distintas soluciones tecnológicas para realizar la medida. Una pinza de corriente puede estar implementada como un transformador de corriente, es decir, mediante un cable enrollado en un núcleo magnético en forma de anillo. Por ese anillo pasa el cable que conduce la corriente a medir. Otras tecnologías incluyen el medidor aleta de hierro (*iron vane*), que no tiene un ancho de banda mucho más grande que 100Hz, las pinzas que funcionan por efecto Hall, que relacionan linealmente el voltaje y el campo magnético generado por la corriente, y las sondas Rogowski, que otorgan un voltaje proporcional a la derivada de la corriente, y son de alto costo.

De la revisión del estado del arte, surge que tanto en los trabajos académicos de referencia [11, 18], como en el diseño de dispositivos comerciales se utilizan transformadores de corriente como pinzas de corriente.

Teniendo en cuenta las referencias, disponibilidad en el mercado y costos, se seleccionó la pinza de corriente SCT-013-000. Esta pinza fue validada por el proyecto Open Energy Monitor [40].

Apéndice B. Pinza de corriente



Figura B.1: Pinza de corriente SCT 013

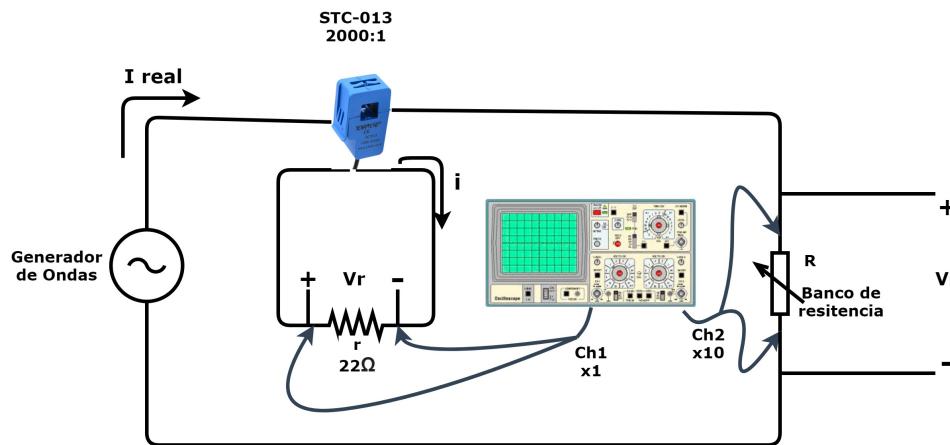


Figura B.2: Diagrama del experimento realizado

Más allá de estas validaciones previas, se consideró relevante realizar un conjunto de experimentos para relevar las características técnicas previo a su uso.

B.3. Metodología

El objetivo de los experimentos es caracterizar el ancho de banda de la pinza de corriente y comprobar que no es menor que 7kHz. Con el fin de hacer pasar por la pinza una corriente de magnitud similar a la consumida por una residencia, se utilizó el generador de ondas APS-1102 GW Instek, y un banco de carga resistivo utilizado usualmente en aplicaciones de potencia a 50Hz. Los instrumentos de medida utilizados fueron el osciloscopio de dos canales Siglent SDS1102CM+ y un multímetro Fluke 115.

El experimento se dispone de la siguiente manera: El generador de ondas se conecta al banco de resistencias. El cable que sale del borne positivo del generador pasa por la pinza de corriente dando $N = 6$ vueltas. Un canal del osciloscopio mide con puntas x10, previamente compensadas, los bornes del banco de resistencias. El otro canal mide el voltaje generado por la pinza de corriente en los bornes de una resistencia de medida de 22Ω . La disposición se muestra en la Figura B.2.

La corriente real (la que se considera medida con fiabilidad), se calcula como

$$I_{real} = \frac{V_R}{R}, \quad (B.1)$$

B.4. Medidas

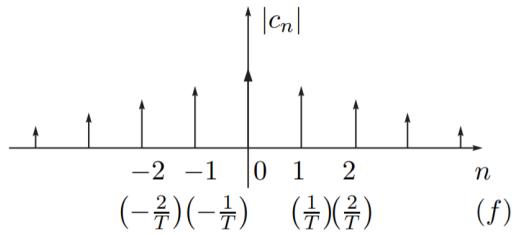


Figura B.3: Espectro de una onda cuadrada

siendo V_R y R el voltaje en bornes y la resistencia del banco de carga resistivo, respectivamente. La corriente estimada se obtiene a partir del voltaje de la resistencia de medida r , que se encuentra en el secundario de la pinza, teniendo en cuenta la relación de transformación (2000 : 1) y la cantidad de vueltas N . Esto es,

$$I_m = \frac{V_r}{r} \frac{2000}{1} \frac{1}{N}. \quad (\text{B.2})$$

Las resistencias R y r se miden con el multímetro.

Se generan ondas sinusoidales, de distintas frecuencias, que generan, aproximadamente, una corriente eficaz de 2A. El generador que se tiene produce sinusoides sin distorsión apreciable a simple vista hasta los 500Hz, por lo que el rango de frecuencias evaluado va de 10Hz a 500Hz. Un ejemplo de una señal sinusoidal se puede observar en la Figura B.4. Dado que con las señales anteriores no se caracteriza el comportamiento de la pinza a frecuencias mayores a 500Hz, se generan ondas cuadradas (de media nula), cuyo espectro tiene, teóricamente, soporte infinito. Los coeficientes de Fourier de la onda cuadrada cumplen

$$c_n = \begin{cases} 0, & n \text{ par} \\ -\frac{jI_p}{n\pi}, & n \text{ impar} \end{cases} \quad (\text{B.3})$$

siendo I_p el valor de pico de la onda.

El espectro de una onda cuadrada ideal se presenta en la Figura B.3. Estas ondas cuadradas se generan de distinta frecuencia, con un tope impuesto por el generador de 550Hz, con el objetivo de que los armónicos se extiendan sobre el espectro. Un ejemplo de onda cuadrada generada por la fuente se muestra en la Figura B.4. Por otro lado, los voltajes medidos en bornes de la resistencia de medida r para los dos tipos de onda se observan en la Figura B.5.

B.4. Medidas

Las medidas tomadas surgen de la extracción de las muestras del osciloscopio. Son 20480 puntos para cada medición. En la tabla B.1 están los valores leídos en la pantalla del generador para cada medida recabada.

Reportar ganancia implica estimar, y comparar, la magnitud de dos señales. La comparación entre dos sinusoides, con el objetivo de reportar ganancia, se realiza en frecuencia. Esto es, se aplica la transformada de Fourier a ambas señales, y se reporta el ratio entre el máximo de ambos espectros (correspondiente a la frecuencia fundamental). Se escoge este camino porque la estimación de la magnitud de la componente fundamental es más robusta y simple que la estimación de otra magnitud, como la amplitud, a partir de la serie temporal. Estimar la amplitud a partir de la serie temporal

Apéndice B. Pinza de corriente

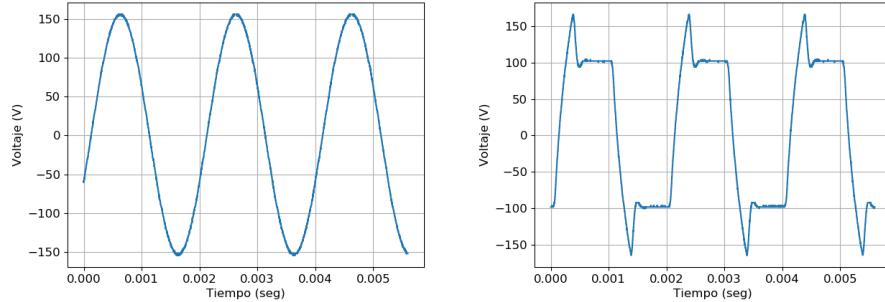


Figura B.4: Voltaje en bornes del banco de carga. A la izquierda, voltaje en bornes del banco cuando el generador se configura para producir una señal sinusoidal. A la derecha, voltaje en bornes del banco cuando el generador se configura para producir una onda cuadrada.

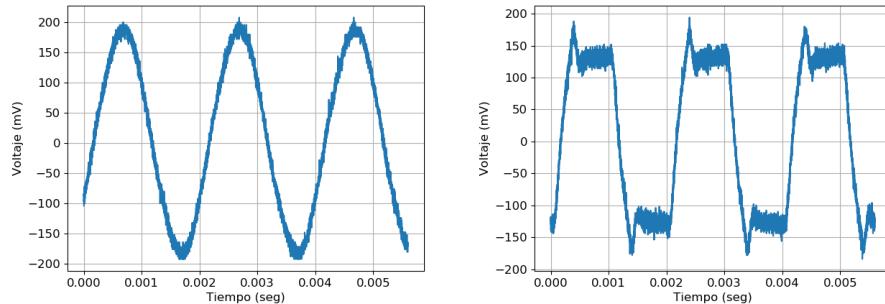


Figura B.5: Voltaje en bornes de r . A la izquierda, voltaje en bornes de r cuando el generador se configura para producir una señal sinusoidal. A la derecha, voltaje en bornes de r cuando el generador se configura para producir una onda cuadrada.

implicaría usar estimadores con gran varianza (como el máximo), recurrir a optimizar sobre diversos parámetros (como amplitud, frecuencia y fase), o a usar métricas de la magnitud como la potencia.

Por otro lado, se cuantifica la ganancia para los primeros siete armónicos para cada onda cuadrada generada. Para evitar el etiquetado manual de los armónicos se encuentra, para cada intervalo definido entre dos armónicos pares consecutivos, un máximo. Este máximo corresponde, en teoría, a un armónico impar. La Figura B.6 muestra la ganancia en función de la frecuencia, resultado de la superposición de todas las ondas cuadradas, junto con la ganancia calculada mediante señales sinusoidales. En esta figura se observa que hay una diferencia (2%) entre la ganancia calculada a partir de las ondas sinusoidales y la ganancia calculada a partir de las ondas cuadradas.

B.5. Impedancia del banco de carga

Las medidas y resultados reportados anteriormente son válidos dentro de la hipótesis de que la resistencia del banco de carga es constante, es decir, que tiene una res-

B.5. Impedancia del banco de carga

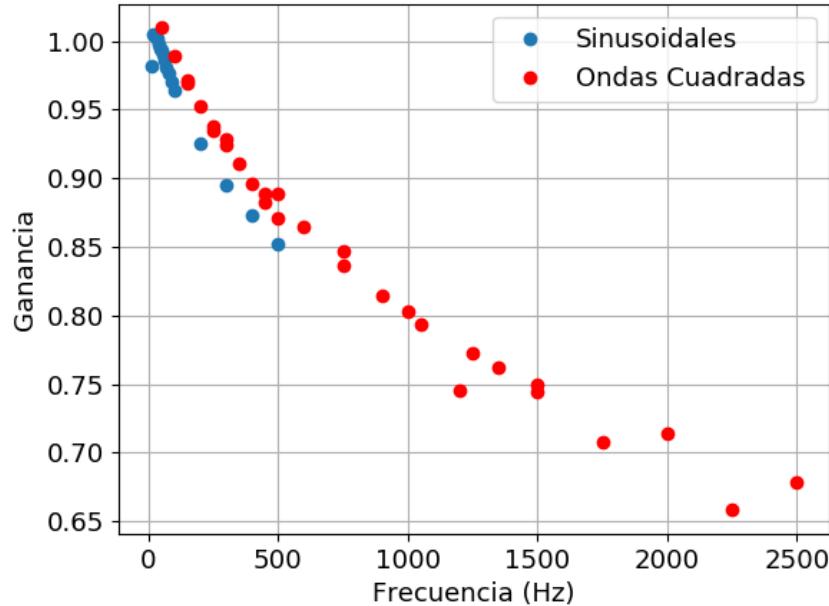


Figura B.6: Ganancia en función de la frecuencia - R constante

puesta en frecuencia plana. Las resistencias que conforman el banco consisten de un bobinado de espaciado considerable alrededor de un núcleo de cerámica. En la tabla B.1 se observan el voltaje y la corriente reportados por el generador de ondas, V_F e I_F , datos que no fueron tomados en cuenta hasta el momento. Se observa que hay una variación entre el cociente entre estas cantidades. Esta observación da lugar al cuestionamiento de esta hipótesis.

Se realizó entonces un experimento sobre el banco de resistencias para ampliar la comprensión del problema planteado. Se dispone un divisor resistivo (Figura B.8), poniendo en serie al banco de cargas y a una resistencia 1000 veces más grande. Se procede a relevar un diagrama de Bode, conectando el instrumento Analog Discovery 2 como generador de onda (hacia el divisor resistivo) y como osciloscopio (midiendo los bornes del banco V_R) a la vez. La respuesta obtenida es la de la (Figura B.7). Nótese que la señal generada es de amplitud 5V, por lo que la corriente pasando por el banco es de algunos μA , órdenes de magnitud menor que la utilizada en el experimento principal. Esta diferencia hace imposible la extrapolación. Solamente es posible afirmar que es plausible que la respuesta en frecuencia del banco de resistencias no sea plana.

Para explorar este asunto más profundamente se plantea modelar el banco de resistencias como una impedancia formada por la serie de una R y una L : $|Z(w)| = \sqrt{R^2 + (wL)^2}$. Los datos usados en este caso fueron $V_F(w)$ e $I_F(w)$, siendo su cociente la referencia para cada frecuencia w . Reescribiendo el modelo se obtienen las ecuaciones dadas por $R^2 + w^2 L^2 = (V_F(w)/I_F(w))^2 \quad \forall w$, que permiten encontrar mediante el método de mínimos cuadrados las cantidades R^2 y L^2 . Los valores obtenidos fueron $R = 46,7\Omega$ y $L = 9,07\text{mH}$, siendo R del orden del valor medido con el multímetro ($R = 36\Omega$). Además, simulando este último experimento en LTSpice con los valores

Apéndice B. Pinza de corriente

Frecuencia	Sinusoide		Onda cuadrada	
	V_F	I_F	V_F	I_F
10	89.9	2	-	-
20	91.0	2	-	-
30	91.9	2	-	-
35	92.3	2	-	-
40	92.7	2	-	-
45	92.9	2	-	-
50	93.3	2	100	2.13
55	93.6	2	-	-
60	93.8	2	-	-
70	94.3	2	-	-
80	94.8	2	-	-
90	95.3	2	-	-
100	95.8	2	100.1	2.08
150	-	-	100.1	2.04
200	99.8	2	100.2	2
250	-	-	100.2	1.98
300	102.4	2	100.2	1.95
350	-	-	100.3	1.93
400	104.7	2	100.3	1.91
450	-	-	100.3	1.89
500	107.3	2	100.4	1.88
550	-	-	100.4	1.86

Tabla B.1: Lecturas del generador de señal

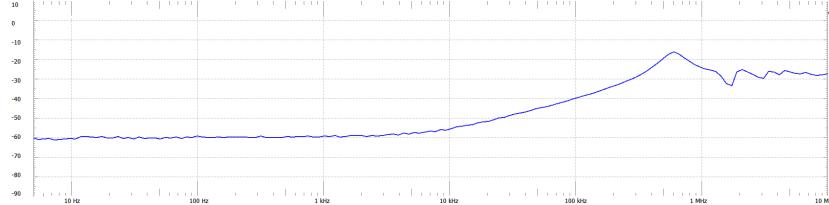


Figura B.7: Respuesta en frecuencia del banco

estimados se obtiene una respuesta en frecuencia que presenta un polo a menos de una octava del polo del diagrama de Bode experimental.

La comparación entre $V_F(w)/I_F(w)$ y la $Z(w)$ generada utilizando los estimados de R y de L se muestra en la Figura B.9. En esta figura se observa que la forma de las curvas es cualitativamente distinta, ya que las concavidades difieren. Esto es, en última instancia, lo que indica que el modelo supuesto no es correcto.

En la Figura B.10 se presenta la respuesta en frecuencia hasta 500Hz recabada con las señales sinusoidales, bajo los distintos modelos de $Z(w)$. Dejando de lado la hipótesis de un banco de carga con respuesta en frecuencia plana y tomando como hipótesis que los valores de $V_F(w)$ y de I_F son confiables a los efectos de determinar $Z(w)$, se obtiene una atenuación menor.

Para analizar el resultado de las ondas cuadradas bajo la hipótesis del banco con

B.6. Conclusiones

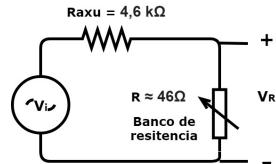


Figura B.8: Divisor resistivo para relevar diagrama de Bode del banco de resistencia

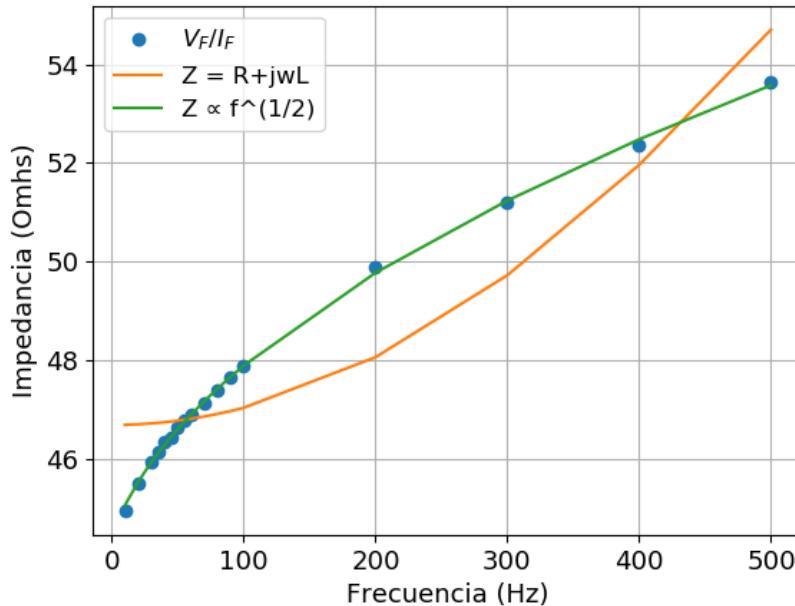


Figura B.9: Impedancia $V_F(w)/I_F = Z(w)$, el ajuste polinómico y el modelo RL

carga variable, es necesario asignar valores de $Z(w)$ para aquellas frecuencias en las que no hay valores de $V_F(w)$. Para esto es inevitable ajustar un modelo a los datos y extrapolar. Un modelo polinómico de grado dos en Z , $w = aZ^2 + bZ + c$, es simple y se ajusta a los datos como se muestra en la Figura B.9. Este modelo es, sin embargo, arbitrario, y determinar su validez y sentido físico requiere un análisis que no se presenta en este trabajo. En la Figura B.11, abajo, se presentan los resultados que surgen de asumirlo válido. Se observa una dispersión considerable más allá de los 3kHz. Esto es causa de la pequeña magnitud de las componentes de alta frecuencia.

B.6. Conclusiones

En este capítulo se reportaron resultados de un experimento realizado sobre un ejemplar de pinza de corriente SCT013-000. Las conclusiones a extraer de las medidas recabadas están sujetas al conjunto de hipótesis considerado: se puede asumir un banco de carga puramente resistivo ($Z(w) = R = \text{cte}$), o no ($Z(w) = V_F(w)/I_F$).

Apéndice B. Pinza de corriente

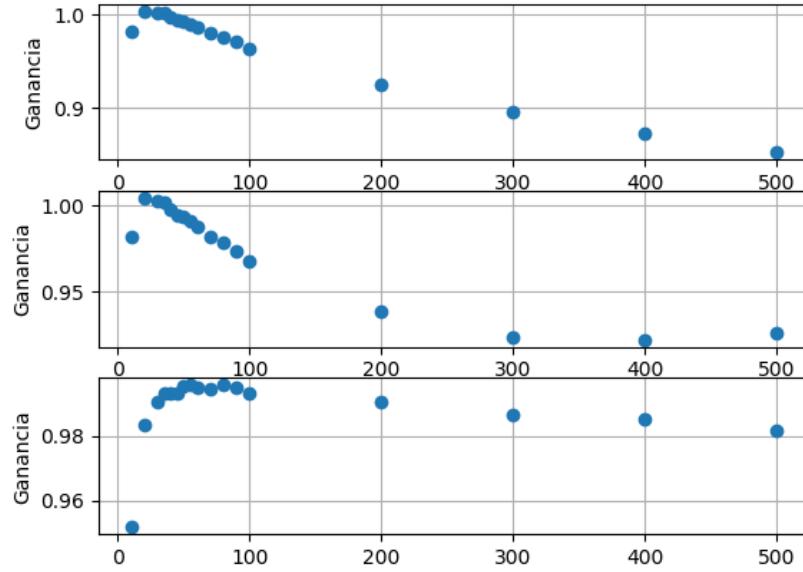


Figura B.10: Ganancia recabada mediante señales sinusoidales. Arriba: $Z(w) = R$, Centro: $Z(w) = \sqrt{R^2 + (wL)^2}$, Abajo: $Z(w) \propto 1/w^2$

Partiendo de la base de que el banco de carga es puramente resistivo se concluye que la pinza de corriente atenúa considerablemente componentes en frecuencia más allá de los 50Hz. Por otro lado, se discute y se muestra que esta suposición es, al menos, dudosa.

Partiendo de que los valores de $V_F(w)$ e I_F son confiables a los efectos de determinar $Z(w)$, se observa que la diferencia cualitativa entre un modelo lineal (con R y L) y el ratio $V_F(w)/I_F$ es insalvable, ya que las concavidades de las curvas son distintas. Esto implica que el banco de carga no es correctamente modelado por una resistencia y una inductancia. Más aún, teniendo en cuenta que la concavidad de la curva $Z(w)$ no cambia al añadir un componente capacitivo, se concluye que el banco de carga no es correctamente modelado por un modelo lineal (Z como combinación de R, L, y C).

Sin embargo, un modelo polinómico de orden 2 puede usarse para ajustar los datos. Bajo este modelo y asumiendo su validez en un amplio rango de frecuencias, se observa que la pinza de corriente atenúa la señal menos de un 5 % en frecuencias menores a 2,7kHz en el rango de interés. Esto es acorde a lo que se lee en [40], que reporta un error del 3 % hasta 3kHz. Más allá de esa frecuencia, la dispersión aumenta y es probable que la asunción de validez del modelo polinómico sea inadecuada.

Sin restringirse a las hipótesis presentadas, se observa que, en el análisis en frecuencia de las ondas cuadradas, y a pesar de la atenuación, la pinza de corriente permite distinguir del ruido componentes de hasta al menos 7kHz, como se muestra en la Figura B.12. Si bien esto depende, en general, de la magnitud de la señal a la entrada, se concluye que para una onda cuadrada y una corriente eficaz similar a la consumida en residencias, la pinza de corriente no anula completamente las componentes de altas frecuencias.

B.6. Conclusiones

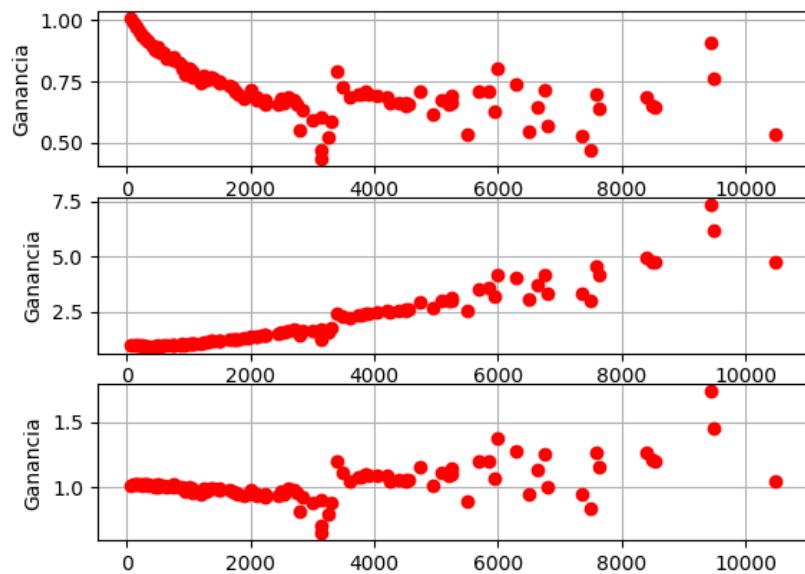


Figura B.11: Ganancia recabada mediante ondas cuadradas. Arriba: $Z(w) = R$, Centro: $Z(w) = \sqrt{R^2 + (wL)^2}$, Abajo: $Z(w) \propto 1/w^2$

Apéndice B. Pinza de corriente

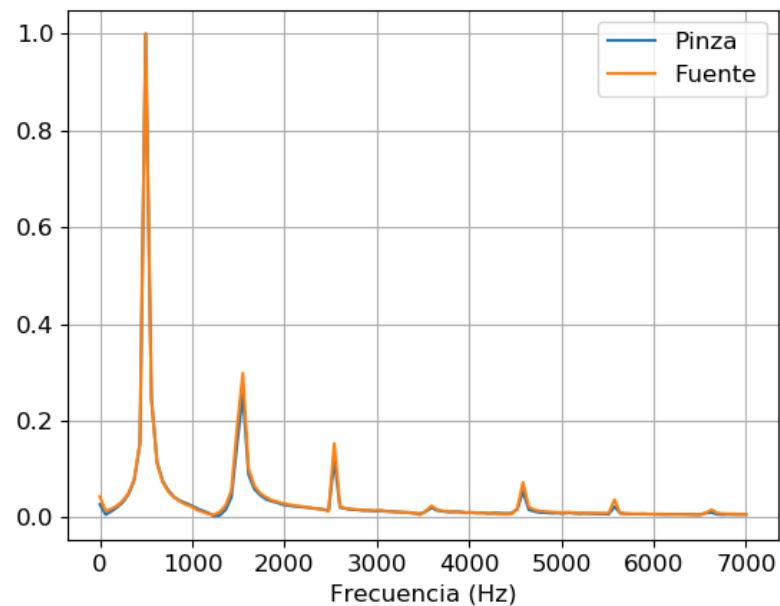


Figura B.12: FFT de la señal leída en bornes de la resistencia de la pinza y en bornes del banco de carga para una onda cuadrada de 500Hz

Apéndice C

Hoja de datos

Split core current transformer

YHDC®

Model: SCT-013

Rated input current: 5A/100A

Characteristics: Openning size: 13mm*13mm,
Non-linearity $\pm 3\%$ (10%—120% of rated input current)
1m leading wire, standard $\Phi 3.5$ three core plug output.
Current output type and voltage output type (voltage output
type built-in sampling resistor)

Purpose: Used for current measurement, monitor and protection for AC motor,
lighting equipment, air compressor etc

Core material: ferrite

Mechanical strength: the number of switching is not less than 1000 times(test at 25°C)

Safty index: Dielectric strength(between shell and output)1000V AC/1min

Fire resistance property: In accordance with UL94-V0

Work temperature: -25°C ~+70°C



Outline size diagram: (in mm)

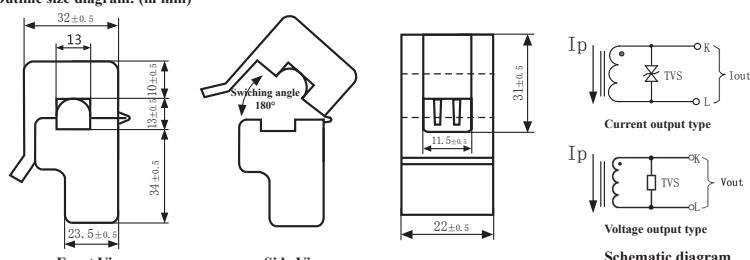


Table of technical parameter:

Model	SCT-013-000	SCT-013-005	SCT-013-010	SCT-013-015	SCT-013-020
Input current	0-100A	0-5A	0-10A	0-15A	0-20A
Output type	0-50mA	0-1V	0-1V	0-1V	0-1V
Model	SCT-013-025	SCT-013-030	SCT-013-050	SCT-013-060	SCT-013-000V
Input current	0-25A	0-30A	0-50A	0-60A	0-100A
Output type	0-1V	0-1V	0-1V	0-1V	0-1V

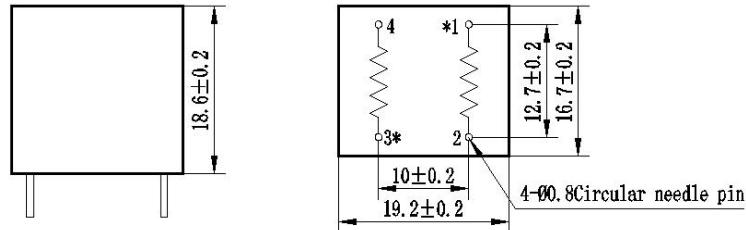
※ Output type: voltage output type built-in sampling resistor, current output type built-in protective diode.

ZMPT101B

Current-type Voltage Transformer

Small size, high accuracy, good consistency, for voltage and power measurement

Structural parameters:



Remarks: primary input: 1、2 pins secondary output: 3、4pins
Or
primary input: 3、4 pins secondary output::1、2pins
“*” Same polarity

Front view

Bottom view

The main technical parameters:

Model	ZMPT101B
Rated input current	2mA
Rated output current	2mA
turns ratio	1000:1000
phase angle error	≤20' (input 2mA, sampling resistor 100Ω)
operating range	0~1000V 0~10mA (sampling resistor 100Ω)
linearity	≤0.2% (20%dot~120%dot)
Permissible error	-0.3%≤ f ≤+0.2% (input 2mA, sampling resistor 100Ω)
isolation voltage	4000V
application	voltage and power measurement
Encapsulation	Epoxy
installation	PCB mounting (Pin Length>3mm)
Operating temperature	-40°C~+60°C
Case Material	ABS (Note: ABS CASE is NOT available for wave-soldering)

Tel: 86-25-52601870

E-mail: zm@zeming-e.com

Apéndice C. Hojas de datos

Qingxian Zeming Langxi Electronic

www.micro-transformer.com

Direction for use:

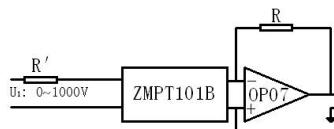


Figure I

1. The typical usage of the product is for the active output (Figure I). R' is a limiting resistor, R is a sampling resistor.
2. The product can be directly through the resistance sampling , easy to use (Figure II).

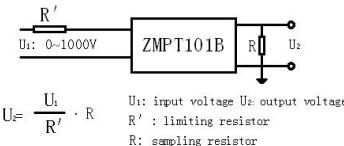


Figure II

$U_2 = \frac{U_1}{R'} \cdot R$

U_1 : input voltage U_2 : output voltage
 R' : limiting resistor
 R : sampling resistor

Tel: 86-25-52601870

E-mail: zm@zeming-e.com

Apéndice D

Archivo de Configuración Wget

Archivo de Configuración Wget

```
# La url base el resto son relativas a ella
base = https://data.ukedc.rl.ac.uk
# Continuar descargando las descargas parciales
continue = on
# Forzar la creación de directorios
dirstruct = on
# De donde leo los links a entrar
input = list/h1_datos_0
# No sobreescribir ni bajar dos veces lo mismo
#no-clobber = on

# Donde estoy parado
dir_prefix=/media/franchesoni/SATback-02/datos_franco/HF-UK

verbose = on

# You can lower or raise the default number of retries when
# downloading a file default is 20.
tries = 5

# Lowering the maximum depth of the recursive retrieval is handy to
# prevent newbies from going too "deep" when they unwittingly start
# the recursive retrieval. The default is 5.
reclevel = 10

# By default Wget uses "passive FTP" transfer where the client
# initiates the data connection to the server rather than the other
# way around. That is required on systems behind NAT where the client
# computer cannot be easily reached from the Internet. However, some
# firewalls software explicitly supports active FTP and in fact has
# problems supporting passive transfer. If you are in such
# environment, use "passive_ftp = off" to revert to active FTP.
passive_ftp = off

# Set this to on to use timestamping by default:
#timestamping = on

# You can set the default proxies for Wget to use for http, https, and ftp.
# They will override the value in the environment.
```

Apéndice D. Archivo de Configuración Wget

```
https_proxy=164.73.32.10:3128/
http_proxy=164.73.32.10:3128/
ftp_proxy=164.73.32.10:3128/

# To have Wget follow FTP links from HTML files by default, set this
# to on:
follow_ftp = on
```

Apéndice E

Resultados de los algoritmos

E.1. Resumen

En este apéndice se incluyen de forma más exhaustiva que en el resto de la documentación, tablas, gráficas y resultados tanto del entrenamiento como de la validación y test de los algoritmos de desagregación desarrollados en 7.

E.2. Resultados de validación sobre UK-DALE

E.2.1. AUC de validación para todos los modelos y experimentos propuestos

En la tabla E.1 se incluyen los resultados de AUC de validación para **todos** los modelos entrenados.

E.3. Resultados de test sobre activaciones en UK-DALE

E.3.1. Casas vistas durante el entrenamiento

Se presentan los resultados sobre los datos correspondientes a los últimos días de datos de las casas sobre las que se entrenó.

Gráficas

A continuación se muestra la ROC para cada electrodoméstico y se grafican algunos ejemplos para observar su desempeño.

Apéndice E. Resultados de los algoritmos

Elec	Model	Data	Run	Adam .0005	Adam 0.001	Adam 0.002	AMax 0.0005	AMax 0.001	AMax 0.002	Best
dish	rectangulos	data	lf_0	0.9863	0.9914	0.9692	0.9875	0.9881	0.977	0.9914
dish	rectangulos	data	lf.syn_0	0.9806	0.9794	0.9775	0.9777	0.9753	0.9723	0.9806
dish	rectangulos	data_lf	hf_0	0.9778	0.9739	0.9708	0.9893	0.9825	0.9491	0.9893
dish	autoencoder	data	ae_lf_run	0.9763	0.5	0.5	0.9846	0.9793	0.5	0.9846
dish	autoencoder	data	ae_big_lf_run	0.9752	0.5	0.5	0.9811	0.5	0.9968	0.9968
dish	autoencoder	data	ae_lf_syn_run	0.9752	0.9799	0.5	0.9886	0.9737	0.5	0.9886
dish	autoencoder	data_hf	ae_hf_run	0.977	0.9753	0.5	0.9412	0.9661	0.5	0.977
fridge	rectangulos	data	lf_0	0.8001	0.5	0.5	0.8102	0.7953	0.8139	0.8139
fridge	rectangulos	data	lf.syn_0	0.811	0.5	0.5	0.8037	0.809	0.8028	0.811
fridge	rectangulos	data_hf	hf_0	0.8689	0.6374	0.5	0.9012	0.868	0.8797	0.9012
fridge	autoencoder	data	ae_lf_run	0.5	0.5	0.5	0.5	0.5	0.5	0.5
fridge	autoencoder	data	ae_big_lf_run	0.5	0.5	0.5	0.5	0.5	0.5	0.5
fridge	autoencoder	data	ae_lf_syn_run	0.5	0.5	0.5	0.5	0.5	0.5	0.5001
fridge	autoencoder	data_hf	ae_hf_run	0.5	0.5	0.5	0.8312	0.4998	0.5	0.8312
kettle	rectangulos	data	lf_0	0.9748	0.9767	0.5003	0.9707	0.9741	0.9686	0.9767
kettle	rectangulos	data	lf.syn_0	0.9695	0.9757	0.5	0.9783	0.9759	0.9815	0.9815
kettle	rectangulos	data_lf	hf_0	0.9652	0.9714	0.9529	0.9642	0.9673	0.9555	0.9714
kettle	autoencoder	data	ae_lf_run	0.9778	0.9769	0.9792	0.9791	0.9772	0.9781	0.9792
kettle	autoencoder	data	ae_big_lf_run	0.9754	0.9726	0.9646	0.9779	0.9795	0.971	0.9795
kettle	autoencoder	data	ae_lf_syn_run	0.9772	0.9768	0.9745	0.9768	0.9765	0.9755	0.9772
kettle	autoencoder	data_hf	ae_hf_run	0.9763	0.984	0.979	0.9762	0.9759	0.9794	0.984
microwave	rectangulos	data	lf_0	0.8999	0.9205	0.5	0.9277	0.9331	0.9231	0.9331
microwave	rectangulos	data	lf.syn_0	0.9125	0.9329	0.5221	0.9357	0.9368	0.912	0.9368
microwave	rectangulos	data_hf	hf_0	0.9203	0.919	0.5007	0.9267	0.9247	0.9122	0.9267
microwave	autoencoder	data	ae_lf_run	0.934	0.9315	0.9256	0.9358	0.9339	0.9263	0.9358
microwave	autoencoder	data	ae_big_lf_run	0.9126	0.9226	0.5	0.9321	0.932	0.9233	0.9321
microwave	autoencoder	data	ae_lf_syn_run	0.9316	0.9337	0.934	0.9439	0.9318	0.9314	0.9439
microwave	autoencoder	data_hf	ae_hf_run	0.942	0.9492	0.9374	0.9442	0.9407	0.9492	
washing	rectangulos	data	lf_0	0.883	0.9002	0.8817	0.8893	0.8582	0.9037	0.9037
washing	rectangulos	data	lf.syn_0	0.8823	0.8658	0.5	0.9004	0.8867	0.8919	0.9004
washing	rectangulos	data_hf	hf_0	0.9252	0.9176	0.8893	0.9509	0.9419	0.9362	0.9509
washing	autoencoder	data	ae_lf_run	0.8774	0.8709	0.5	0.8774	0.5	0.5	0.8774
washing	autoencoder	data	ae_big_lf_run	0.5	0.5	0.5	0.8496	0.5	0.5	0.8496
washing	autoencoder	data	ae_lf_syn_run	0.8633	0.8746	0.5	0.8682	0.4983	0.5	0.8746
washing	autoencoder	data_hf	ae_hf_run	0.8761	0.5	0.49	0.494	0.5	0.5	0.8761

Tabla E.1: AUC de validación para todos los modelos entrenados

E.3. Resultados de test sobre activaciones en UK-DALE

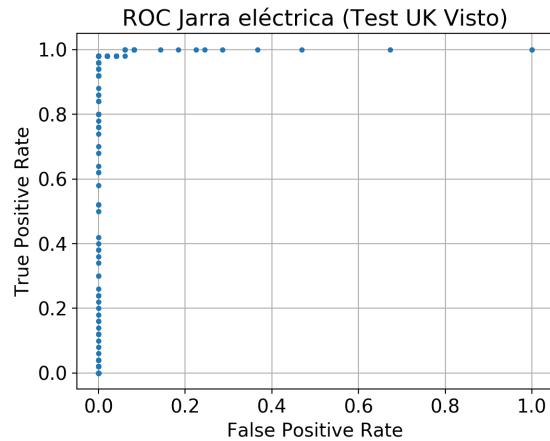


Figura E.1: ROC Jarra eléctrica en test visto. $AUC = 0,9988$

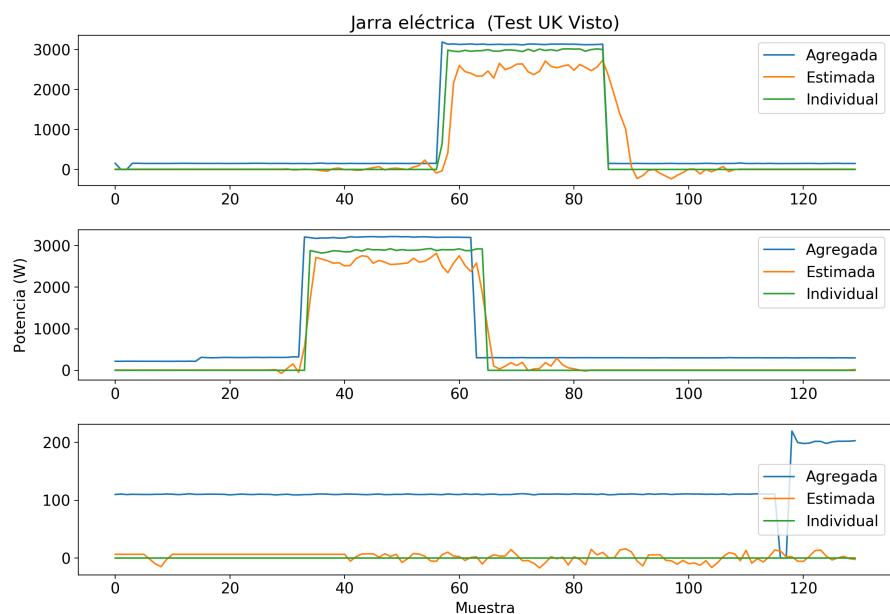


Figura E.2: Resultados para la jarra eléctrica en test visto.

Apéndice E. Resultados de los algoritmos

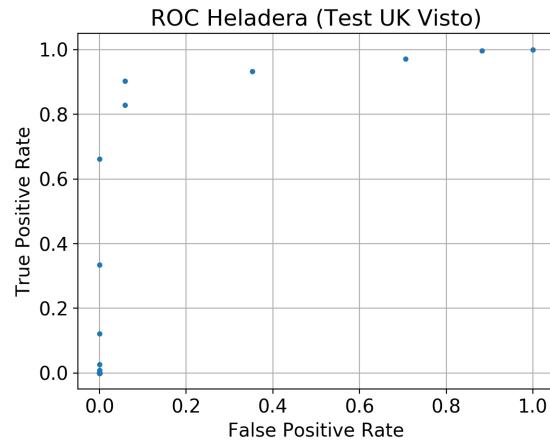


Figura E.3: ROC Heladera en test visto. $AUC = 0,9411$

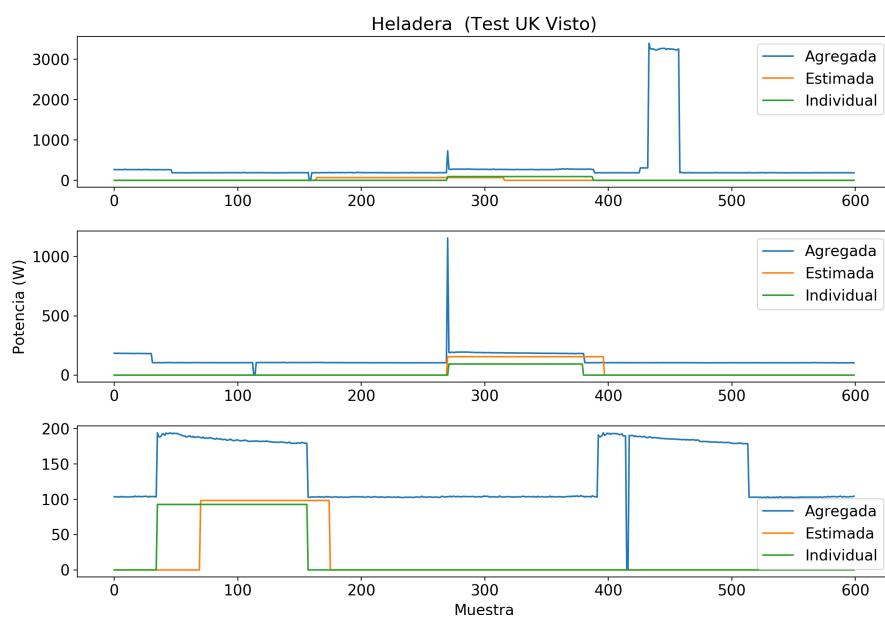


Figura E.4: Resultados para la heladera en test visto.

E.3. Resultados de test sobre activaciones en UK-DALE

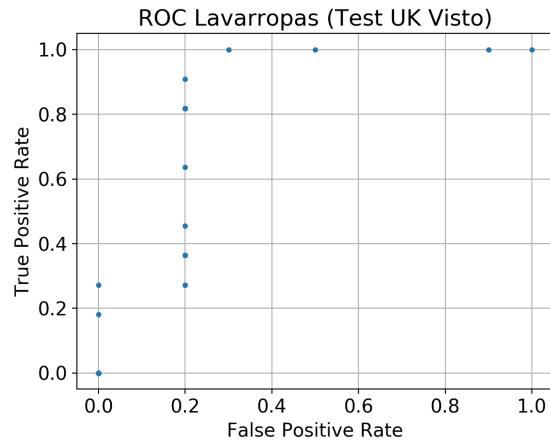


Figura E.5: ROC Lavarropas en test visto. $AUC = 0,85$

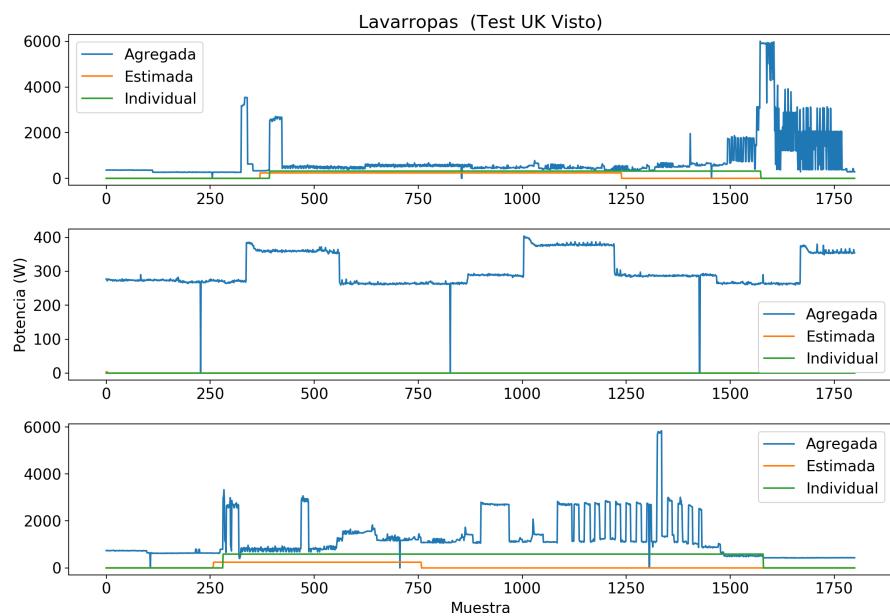


Figura E.6: Resultados para el lavarropas en test visto.

Apéndice E. Resultados de los algoritmos

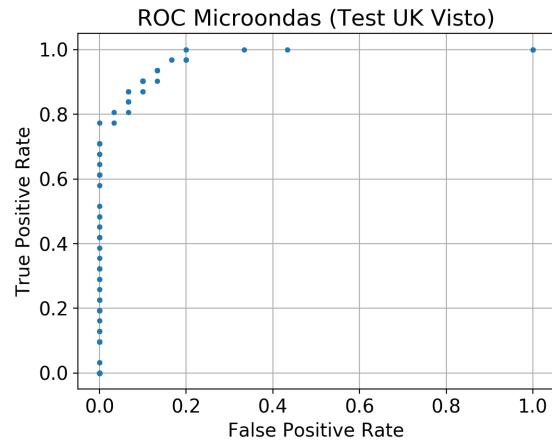


Figura E.7: ROC Microondas en test visto. $AUC = 0,9758$

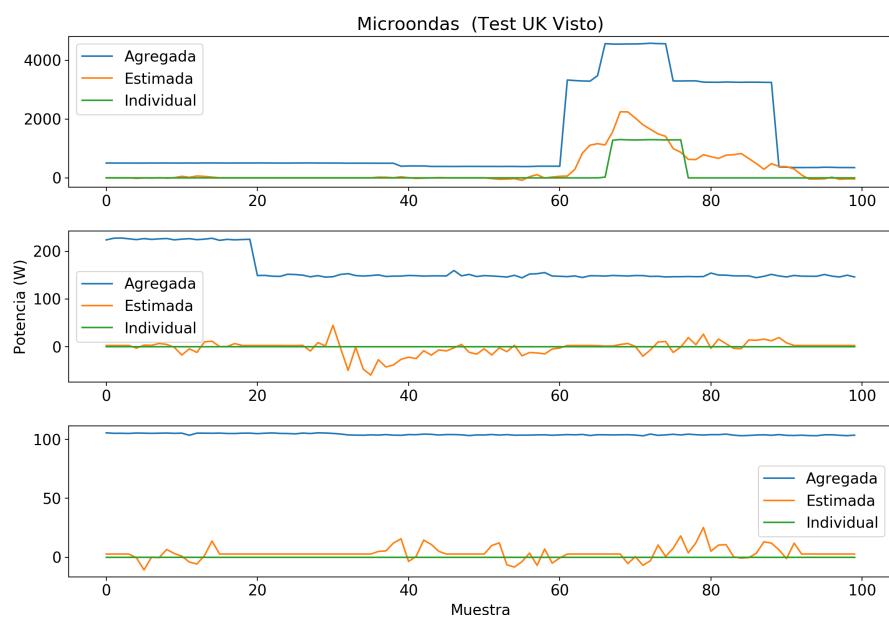


Figura E.8: Resultados para el microondas en test visto.

E.3. Resultados de test sobre activaciones en UK-DALE

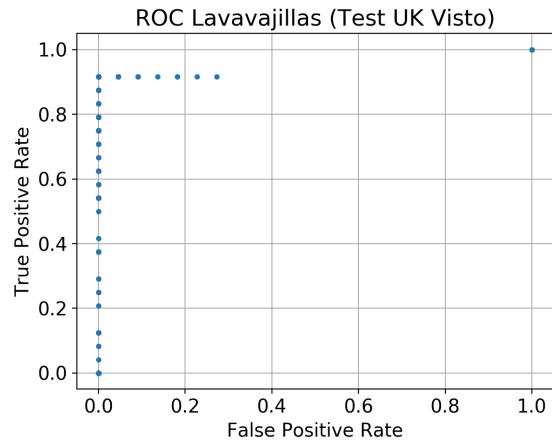


Figura E.9: ROC Lavavajillas en test visto. $AUC = 0,9470$

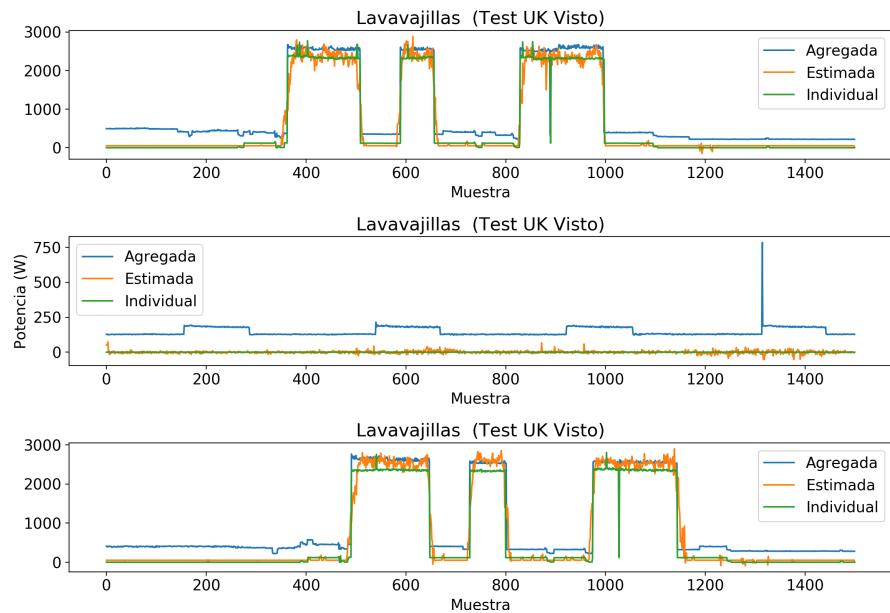


Figura E.10: Resultados para el lavavajillas en test visto.

E.3.2. Casas no vistas

Gráficas

A continuación se muestra la ROC para cada electrodoméstico y se grafican algunos ejemplos para observar su desempeño.

Apéndice E. Resultados de los algoritmos

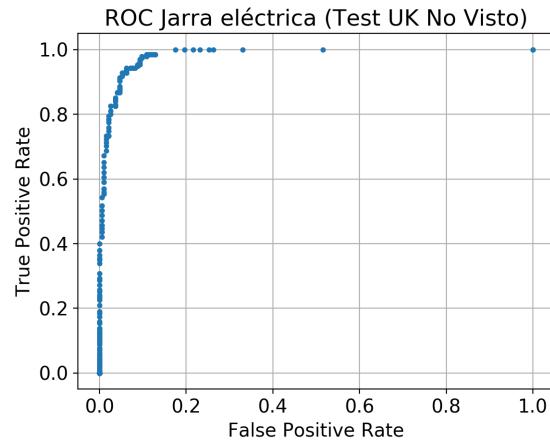


Figura E.11: ROC Jarra eléctrica en test no visto. $AUC = 0,9835$

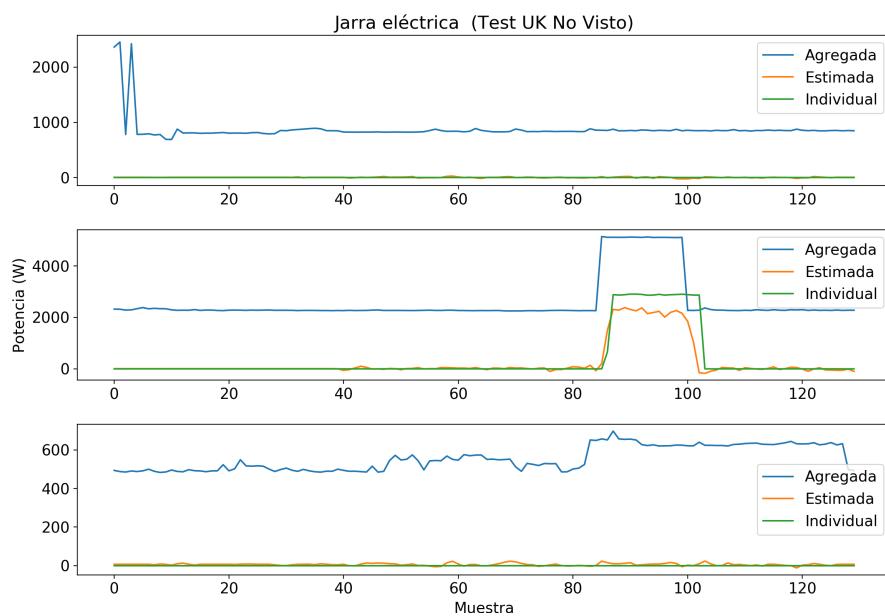


Figura E.12: Resultados para la jarra eléctrica en test no visto.

E.3. Resultados de test sobre activaciones en UK-DALE

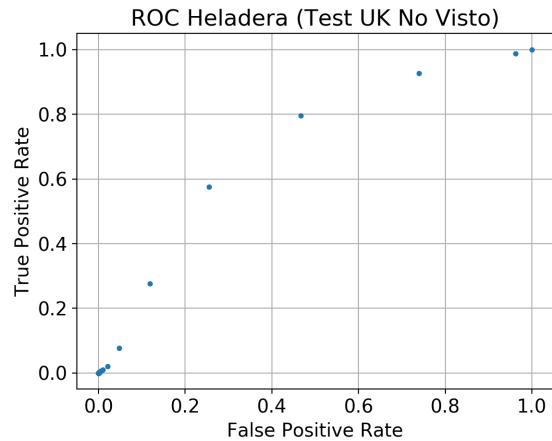


Figura E.13: ROC Heladera en test no visto. $AUC = 0,7030$

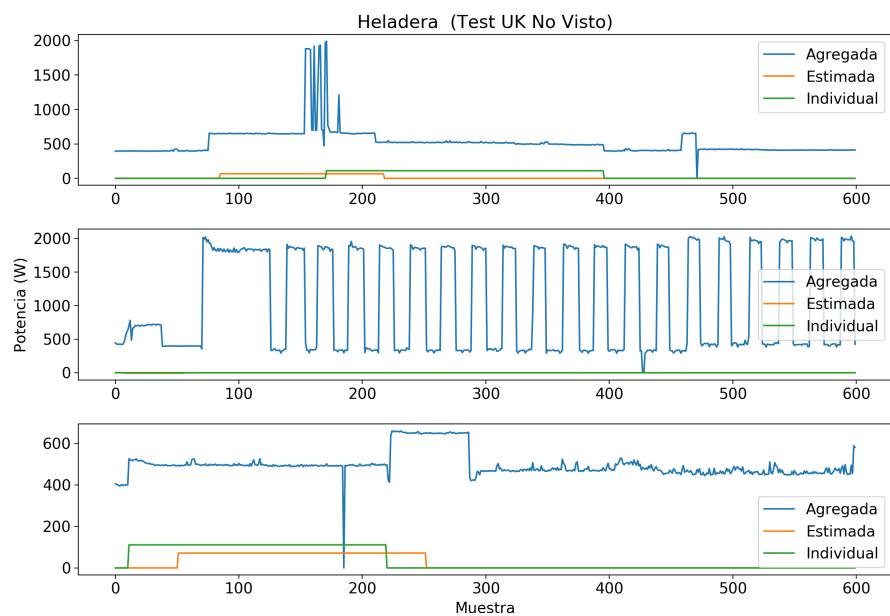


Figura E.14: Resultados para la heladera en test no visto.

Apéndice E. Resultados de los algoritmos

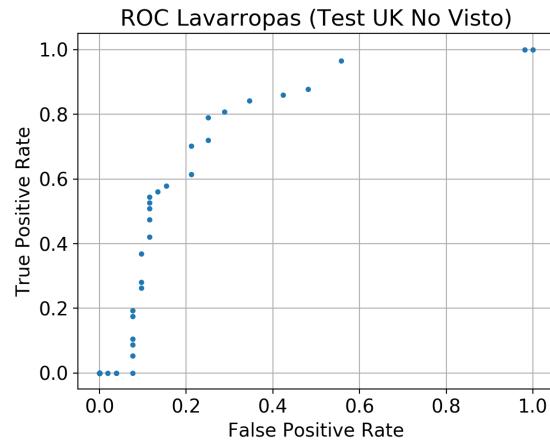


Figura E.15: ROC Lavarropas en test no visto. $AUC = 0,7949$

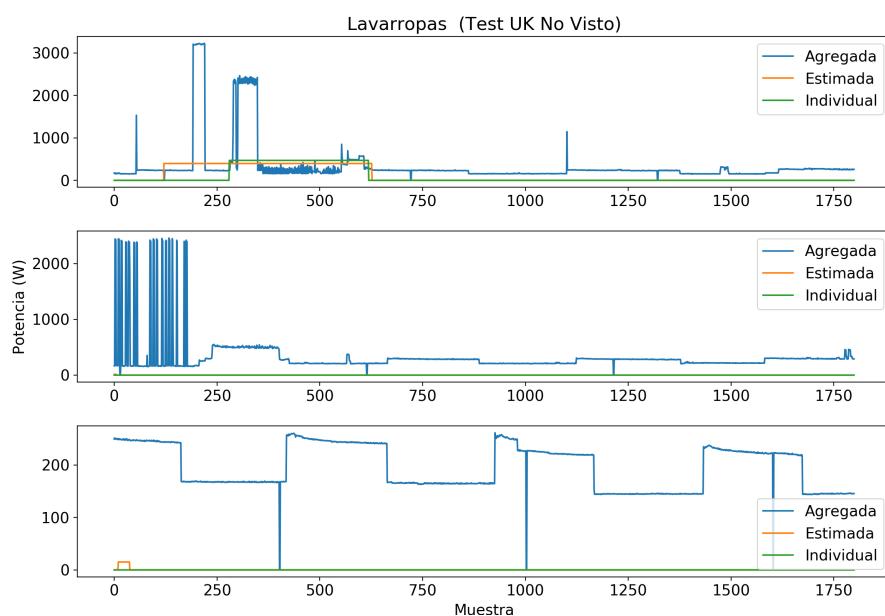


Figura E.16: Resultados para el lavarropas en test no visto.

E.3. Resultados de test sobre activaciones en UK-DALE

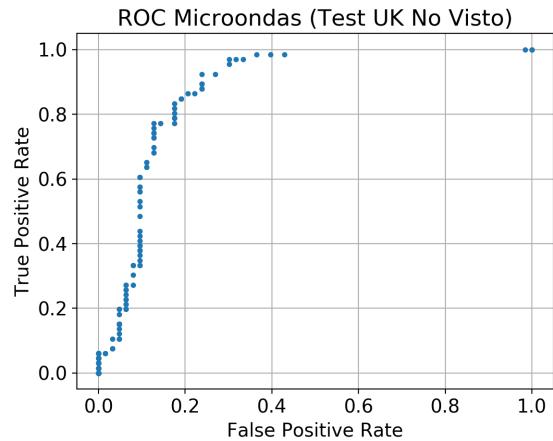


Figura E.17: ROC Microondas en test no visto. $AUC = 0,8789$

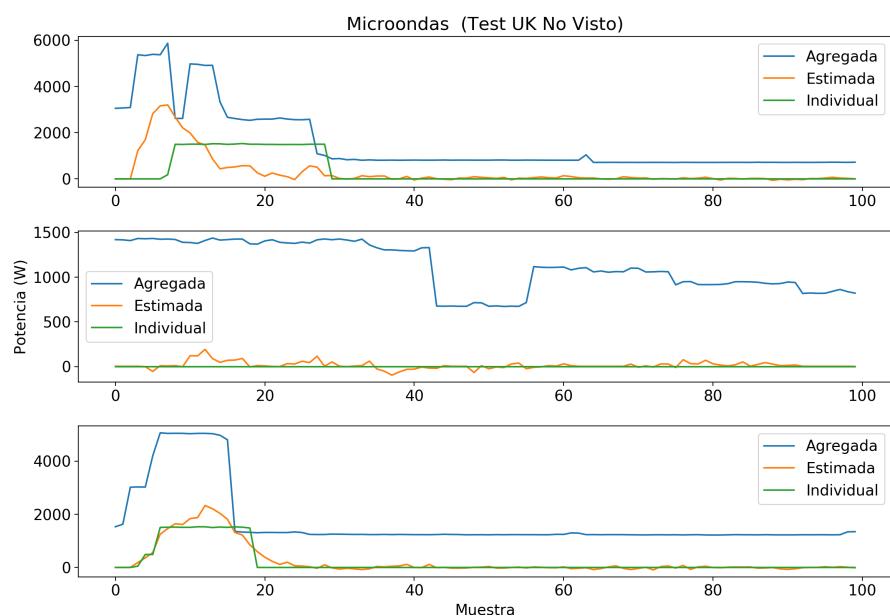


Figura E.18: Resultados para el microondas en test no visto.

Apéndice E. Resultados de los algoritmos

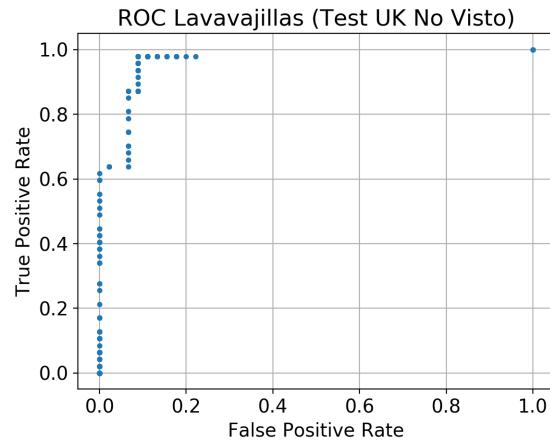


Figura E.19: ROC Lavavajillas en test no visto. $AUC = 0.9617$

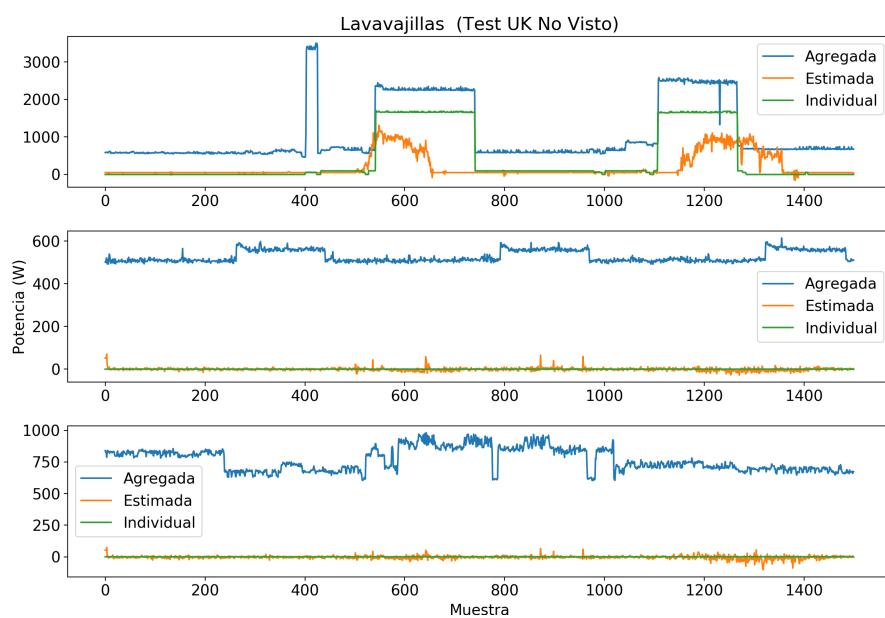


Figura E.20: Resultados para el lavavajillas en test no visto.

E.4. Resultados de test sobre ventanas deslizantes en UK-DALE

E.4. Resultados de test sobre ventanas deslizantes en UK-DALE

E.4.1. Casas vistas durante el entrenamiento

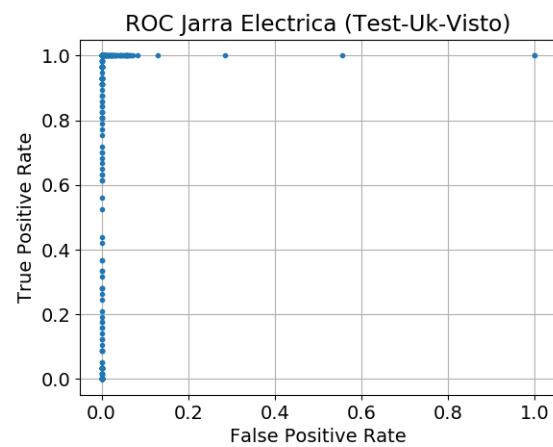


Figura E.21: ROC Jarra Electrica Ventanas Deslizantes Test Visto. $AUC = 0,99995$

Apéndice E. Resultados de los algoritmos

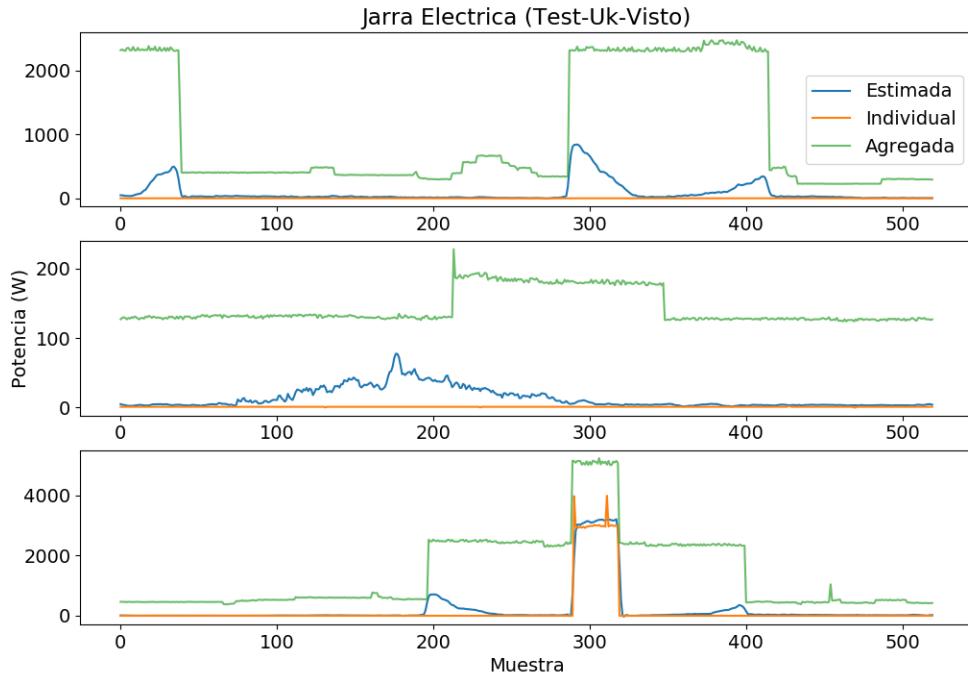


Figura E.22: Ejemplo Jarra Electrica Ventanas Deslizantes Test Visto. $AUC = 0,99995$

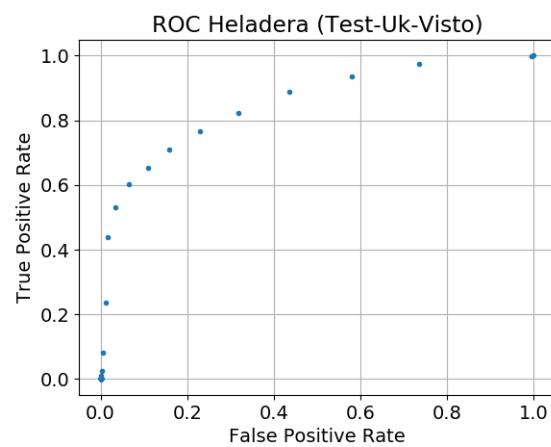


Figura E.23: ROC Heladera Ventanas Deslizantes Test Visto

E.4. Resultados de test sobre ventanas deslizantes en UK-DALE

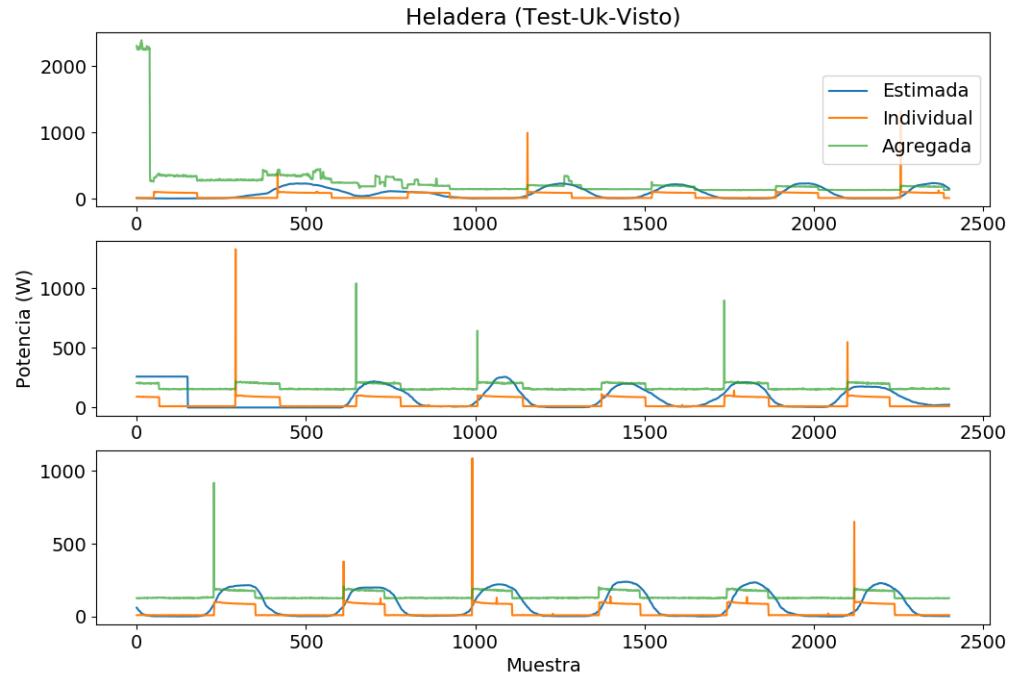


Figura E.24: Ejemplo Heladera Ventanas Deslizantes Test Visto

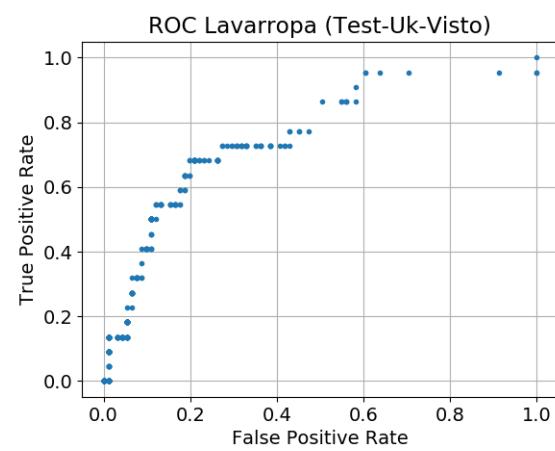


Figura E.25: ROC Lavarropas Ventanas Deslizantes Test Visto. $AUC = 0,7635$

Apéndice E. Resultados de los algoritmos

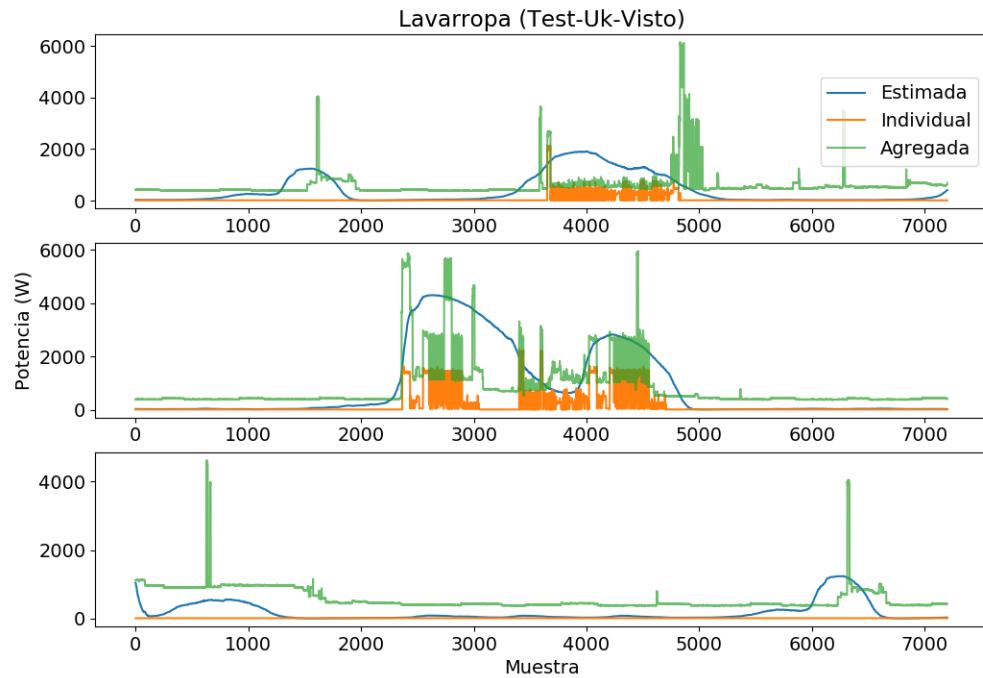


Figura E.26: Ejemplo Lavarropa Ventanas Deslizantes Test Visto

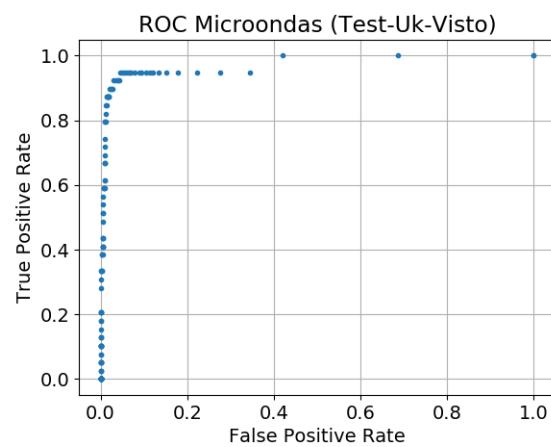


Figura E.27: ROC Microondas Ventanas Deslizantes Test Visto. $AUC = 0,9731$

E.4. Resultados de test sobre ventanas deslizantes en UK-DALE

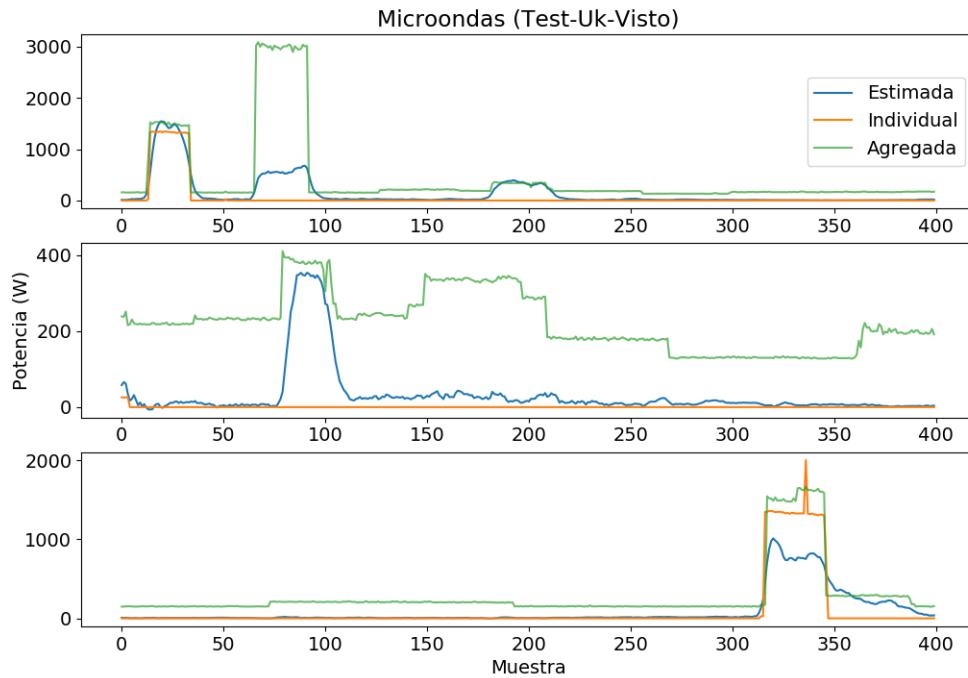


Figura E.28: Ejemplo Microondas Ventanas Deslizantes Test Visto

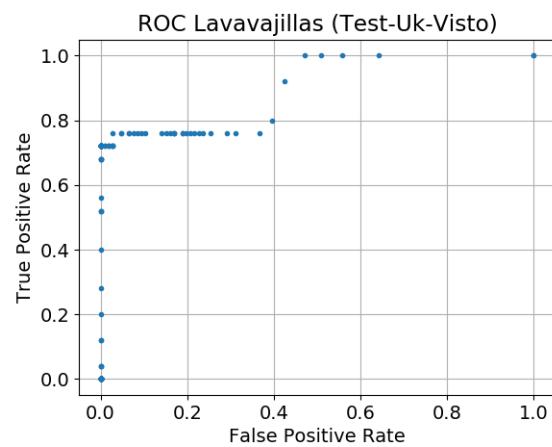


Figura E.29: ROC Lavavajillas Ventanas Deslizantes Test Visto. $AUC = 0,8985$

Apéndice E. Resultados de los algoritmos

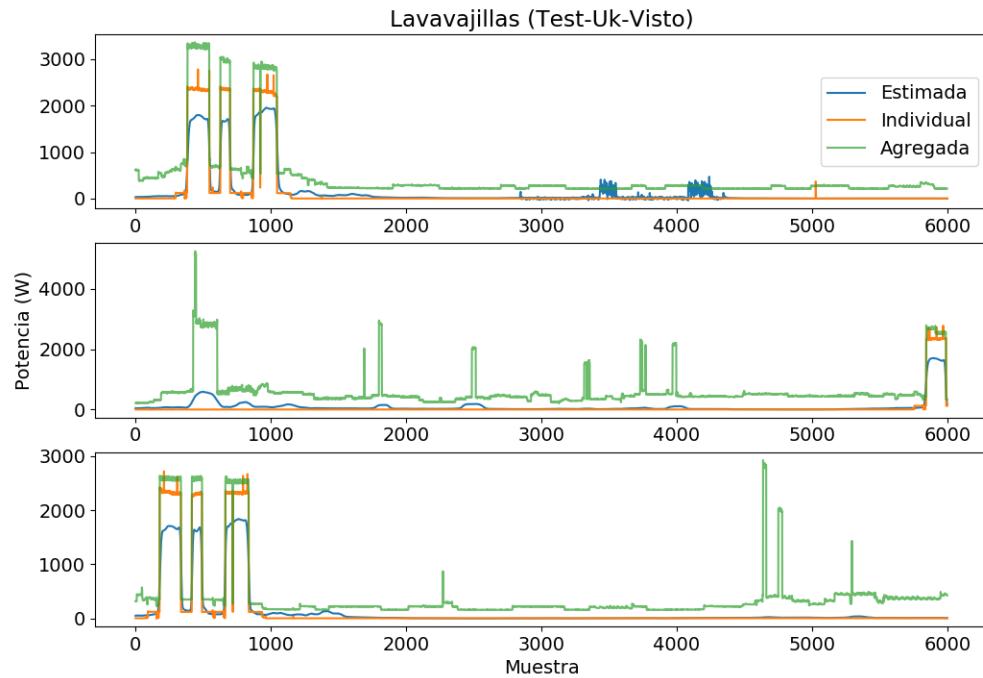


Figura E.30: Ejemplo Lavavajillas Ventanas Deslizantes Test Visto

E.4.2. Casas no vistas durante el entrenamiento

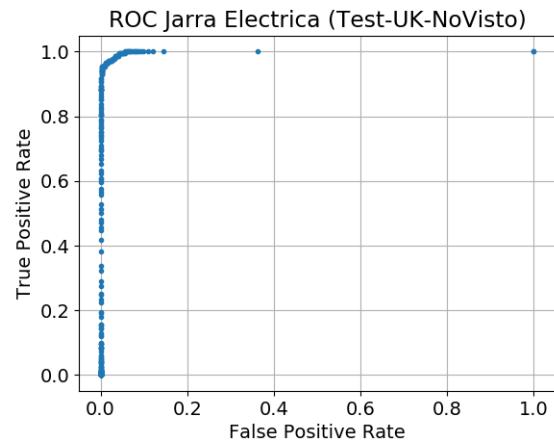


Figura E.31: ROC Jarra Electrica Ventanas Deslizantes Test no Visto. $AUC = 0,9982$

E.4. Resultados de test sobre ventanas deslizantes en UK-DALE

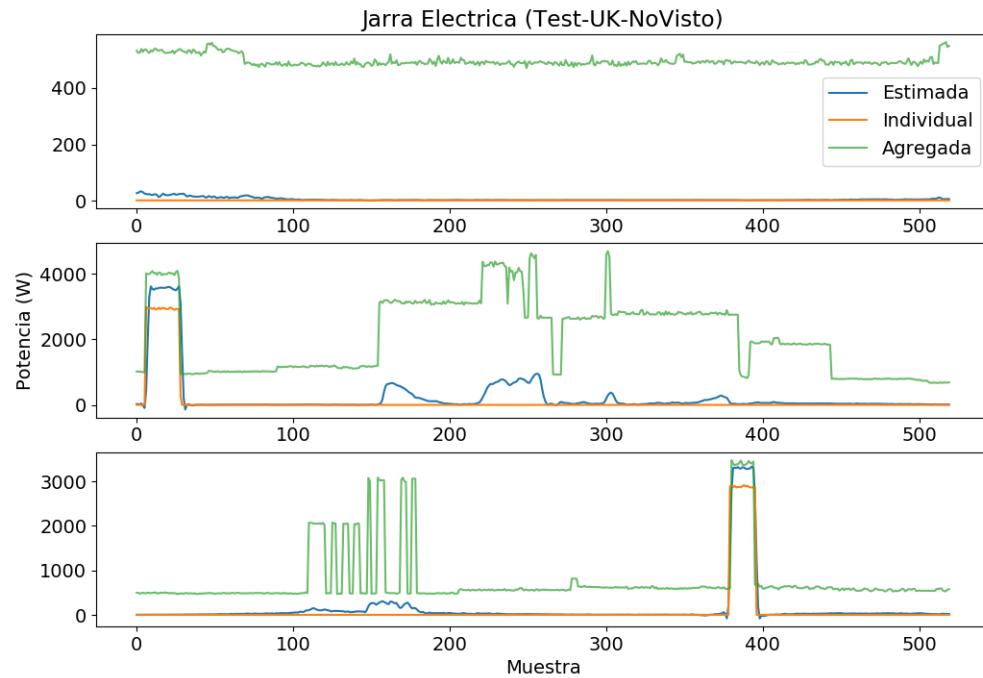


Figura E.32: Ejemplo Jarra Electrica Ventanas Deslizantes Test no Visto

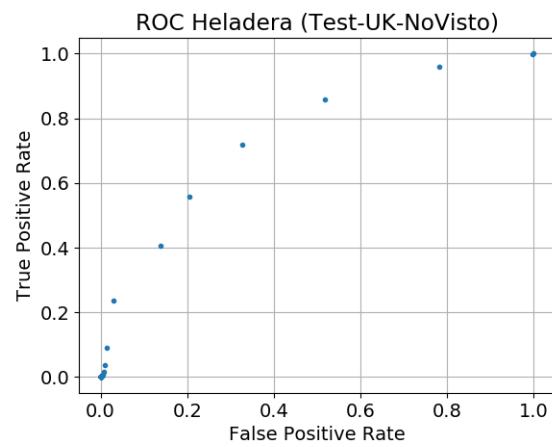


Figura E.33: ROC Heladera Ventanas Deslizantes Test no Visto. $AUC = 0,8538$

Apéndice E. Resultados de los algoritmos

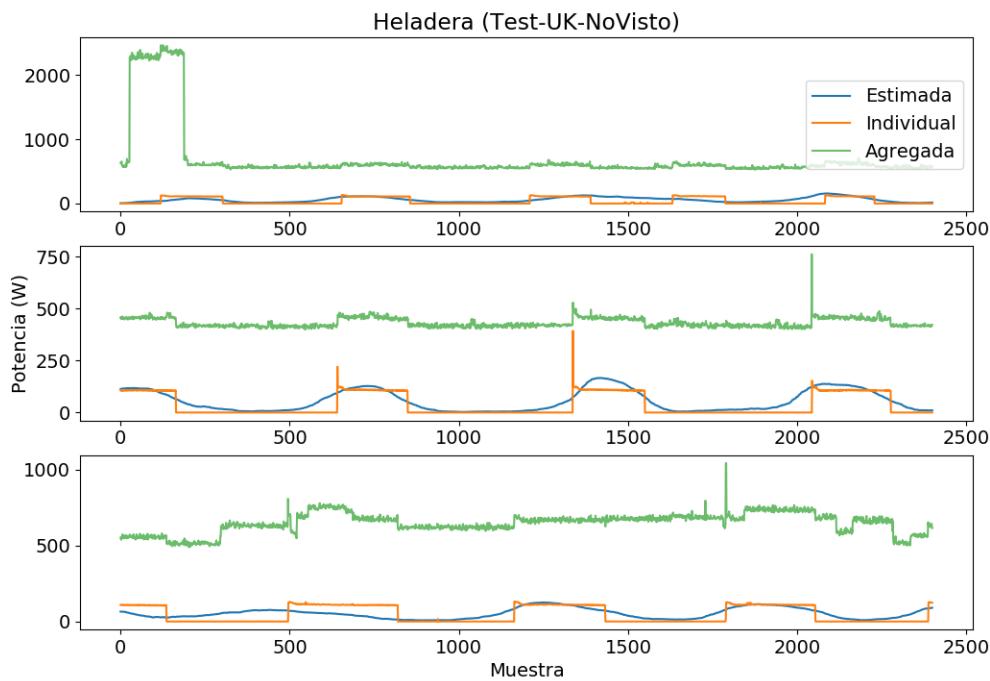


Figura E.34: Ejemplo Heladera Ventanas Deslizantes Test no Visto

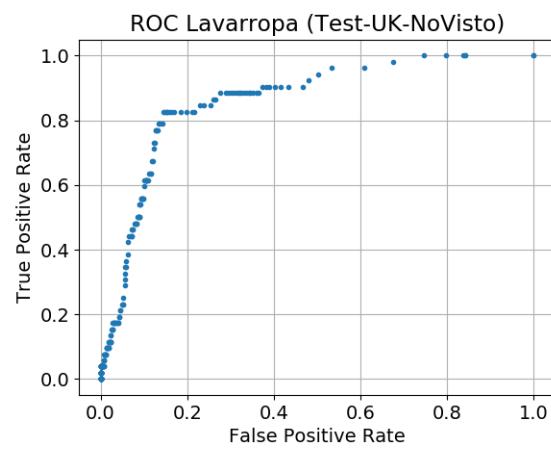


Figura E.35: ROC Lavarropas Ventanas Deslizantes Test no Visto. $AUC = 0,7635$

E.4. Resultados de test sobre ventanas deslizantes en UK-DALE

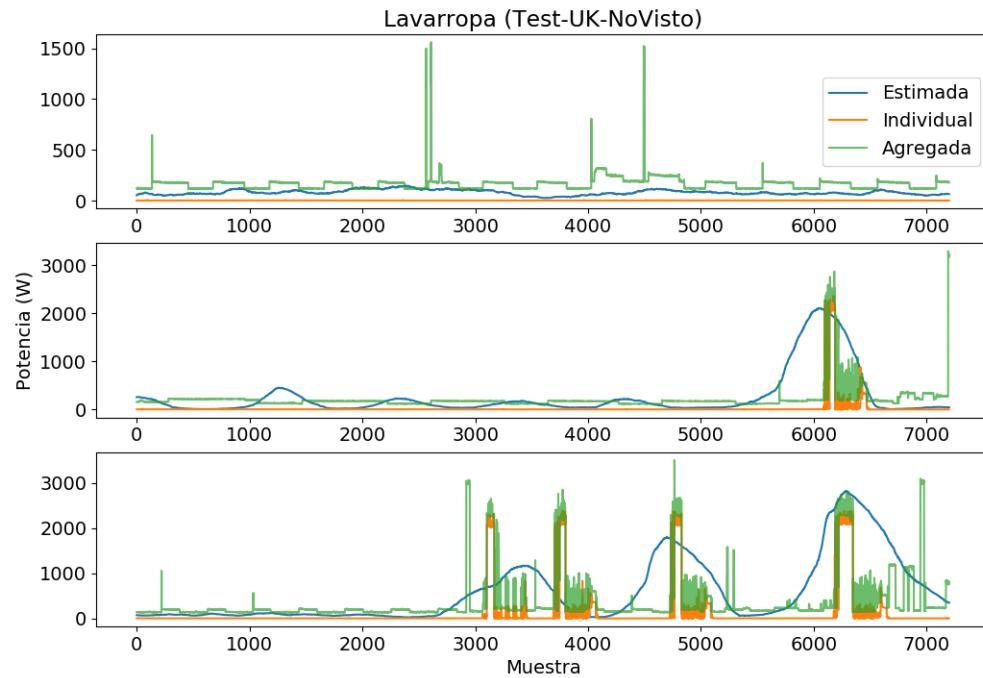


Figura E.36: Ejemplo Lavarropa Ventanas Deslizantes Test no Visto

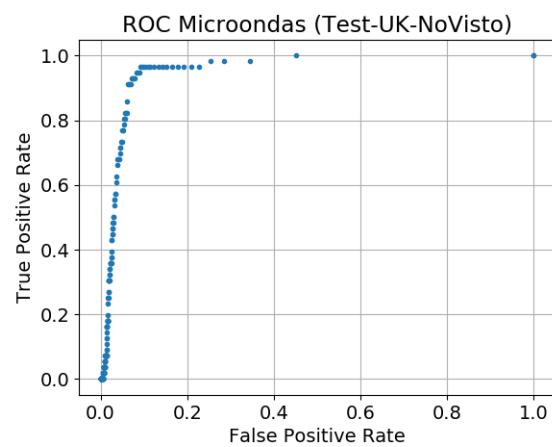


Figura E.37: ROC Microondas Ventanas Deslizantes Test no Visto. $AUC = 0,9731$

Apéndice E. Resultados de los algoritmos

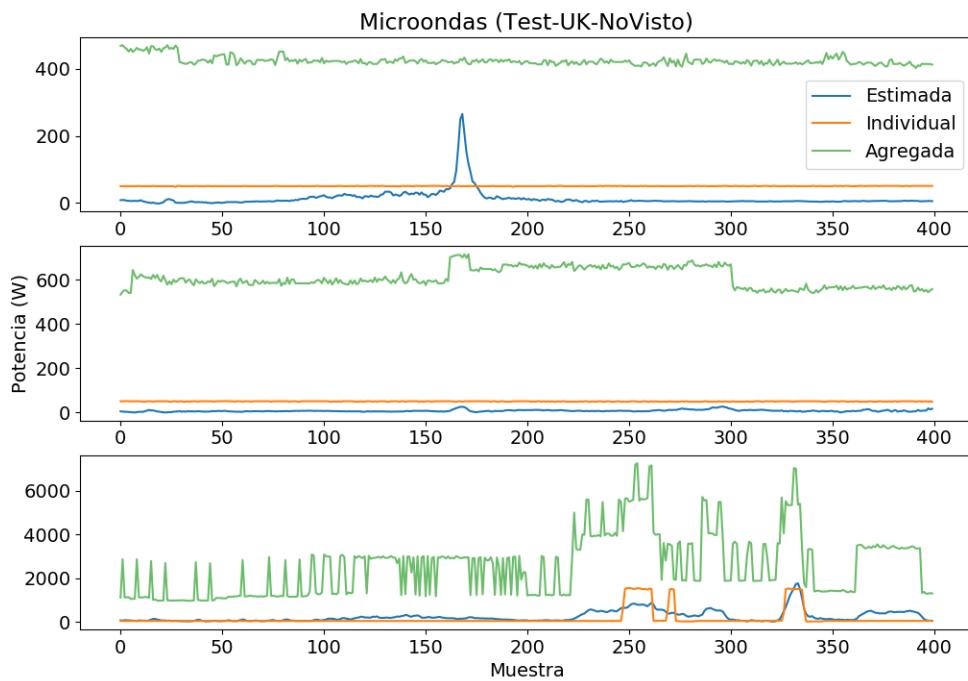


Figura E.38: Ejemplo Microondas Ventanas Deslizantes Test no Visto

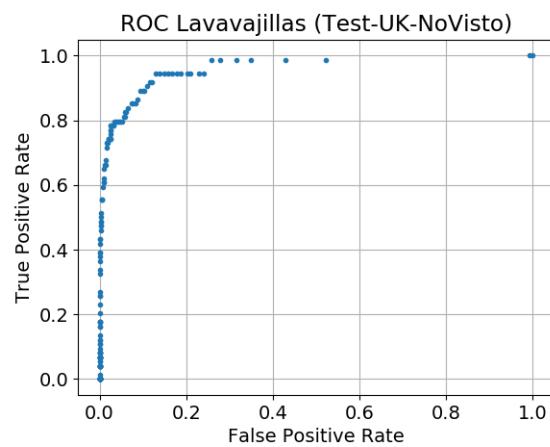


Figura E.39: ROC Lavavajillas Ventanas Deslizantes Test no Visto. $AUC = 0,8985$

E.5. Resultado de test en ventanas deslizantes sobre NILM-UY

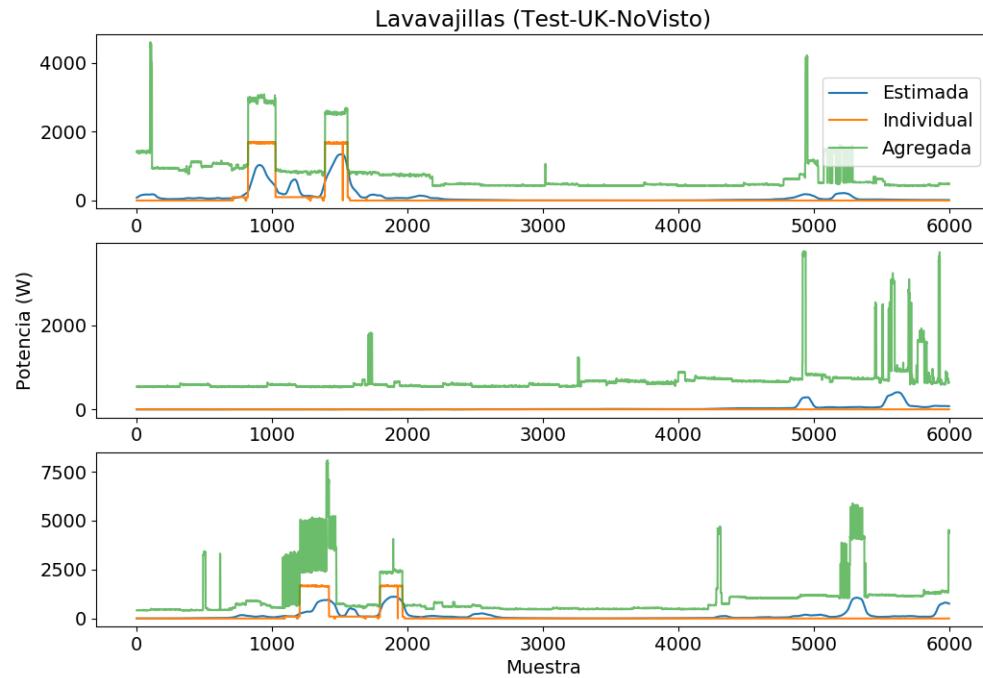


Figura E.40: Ejemplo Lavavajillas Ventanas Deslizantes Test no Visto

E.5. Resultado de test en ventanas deslizantes sobre NILM-UY

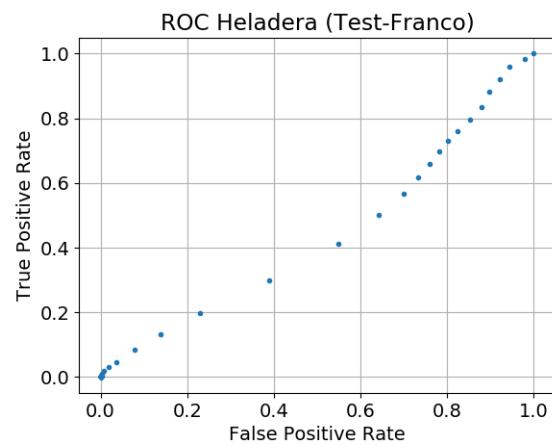


Figura E.41: ROC Heladera Ventanas Deslizantes Test casa Franco. $AUC = 0,4351$

Apéndice E. Resultados de los algoritmos

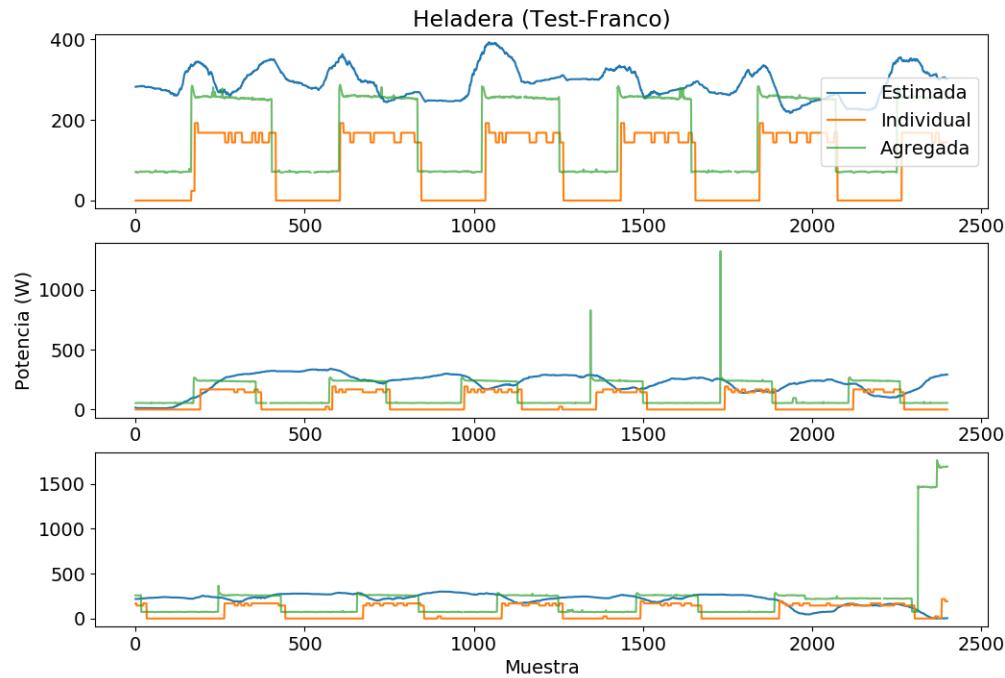


Figura E.42: Ejemplo Heladera Ventanas Deslizantes Test casa Franco

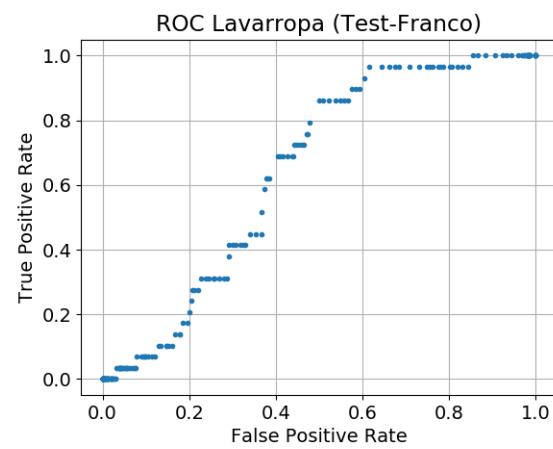


Figura E.43: ROC Lavarropas Ventanas Deslizantes Test casa Franco. $AUC = 0,6467$

E.5. Resultado de test en ventanas deslizantes sobre NILM-UY

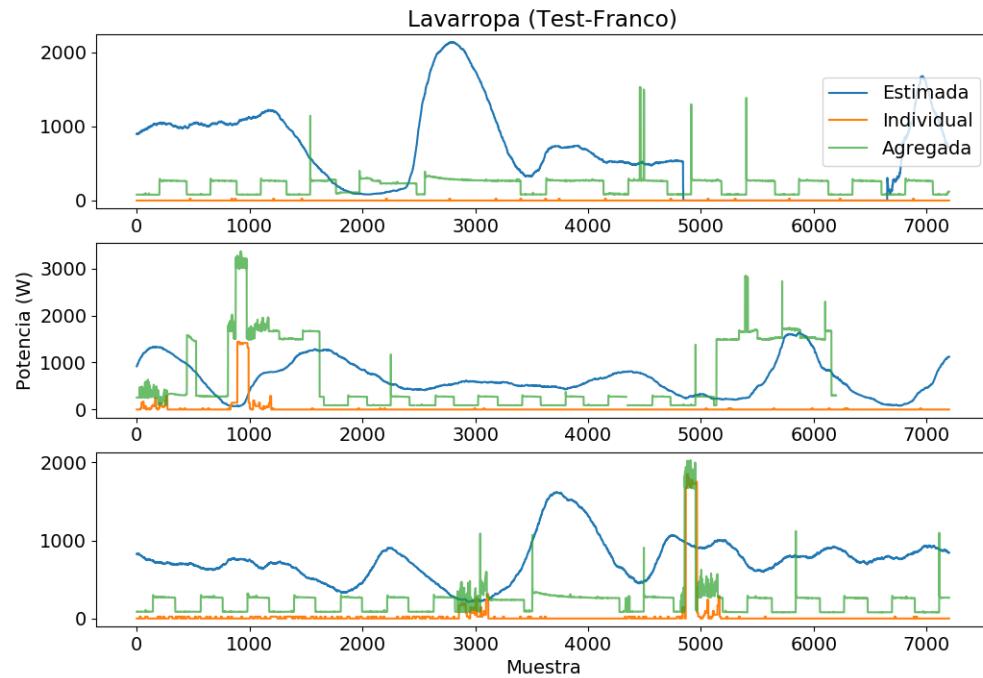


Figura E.44: Ejemplo Lavarropa Ventanas Deslizantes Test casa Franco

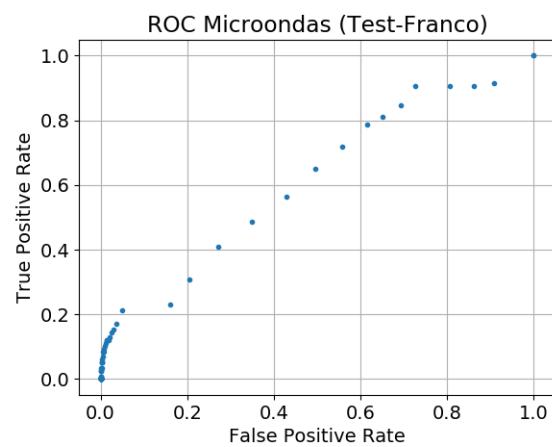


Figura E.45: ROC Microondas Ventanas Deslizantes Test casa Franco. $AUC = 0,6154$

Apéndice E. Resultados de los algoritmos

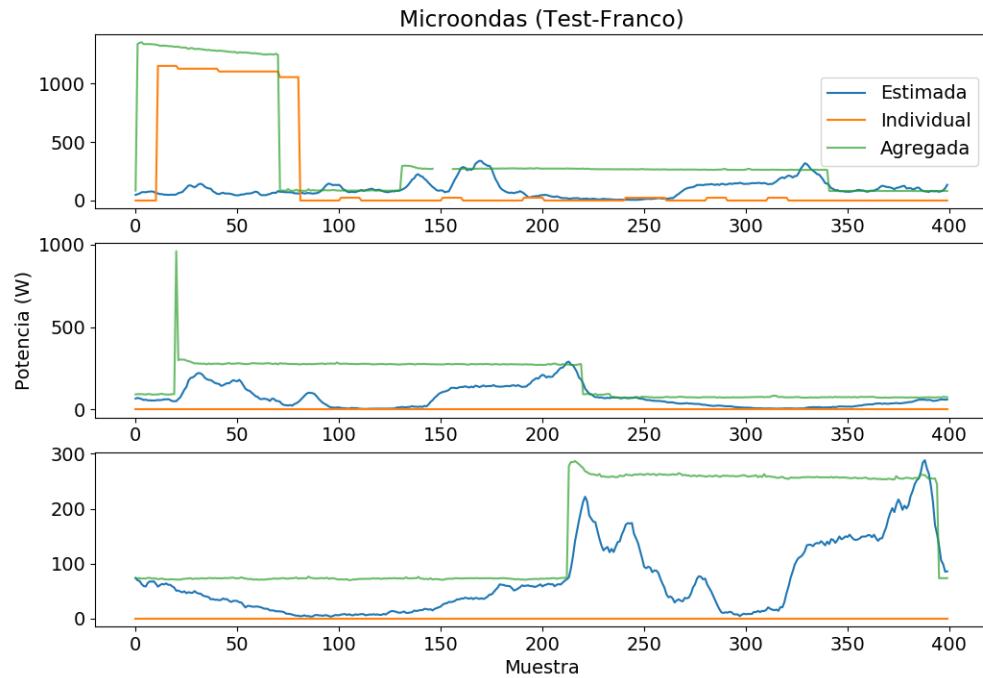


Figura E.46: Ejemplo Microondas Ventanas Deslizantes Test casa Franco

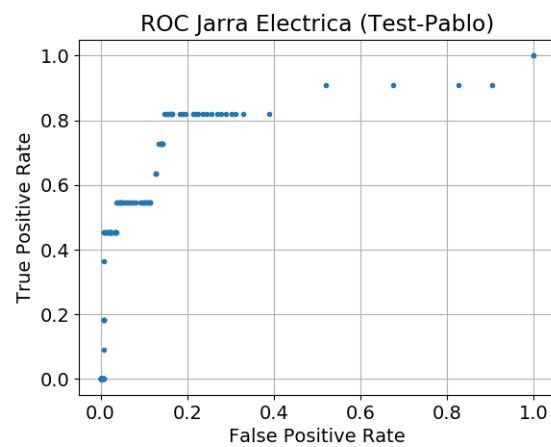


Figura E.47: ROC Jarra Electrica Ventanas Deslizantes Test casa Pablo. $AUC = 0,8295$

E.5. Resultado de test en ventanas deslizantes sobre NILM-UY

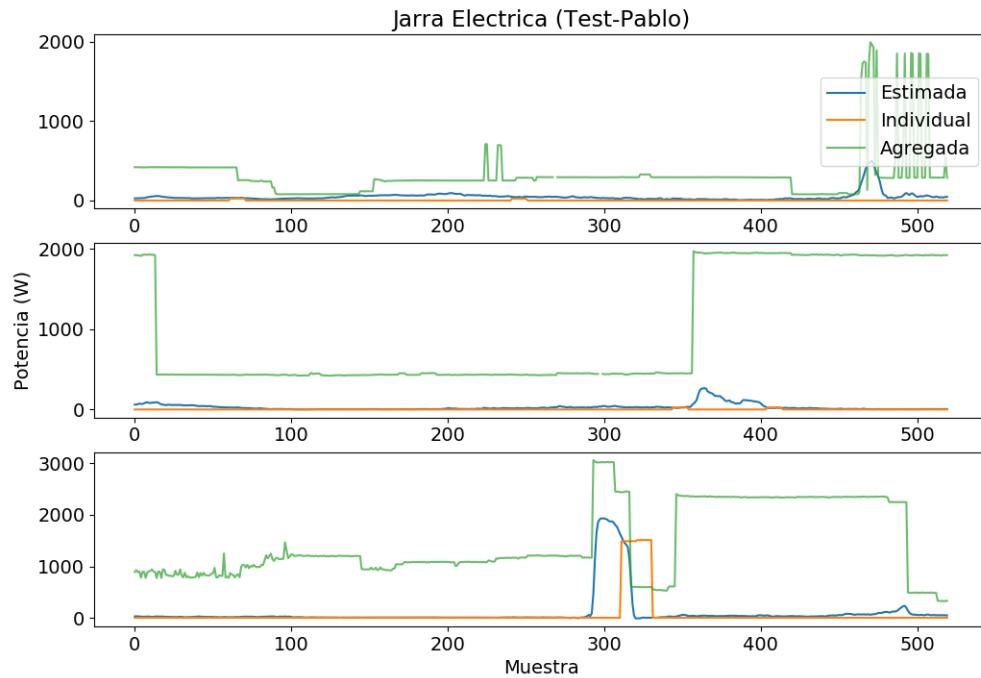


Figura E.48: Ejemplo Jarra Electrica Ventanas Deslizantes Test casa Pablo

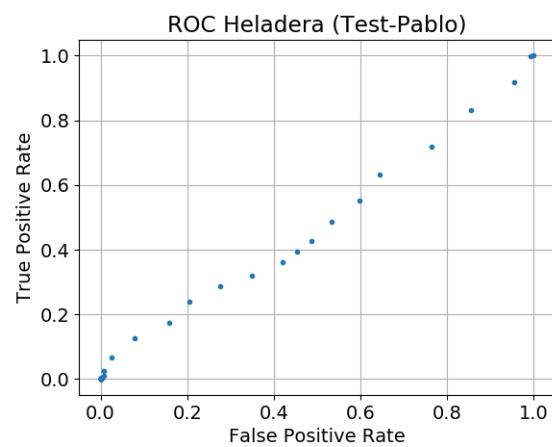


Figura E.49: ROC Heladera Ventanas Deslizantes Test casa Pablo. $AUC = 0,4833$

Apéndice E. Resultados de los algoritmos

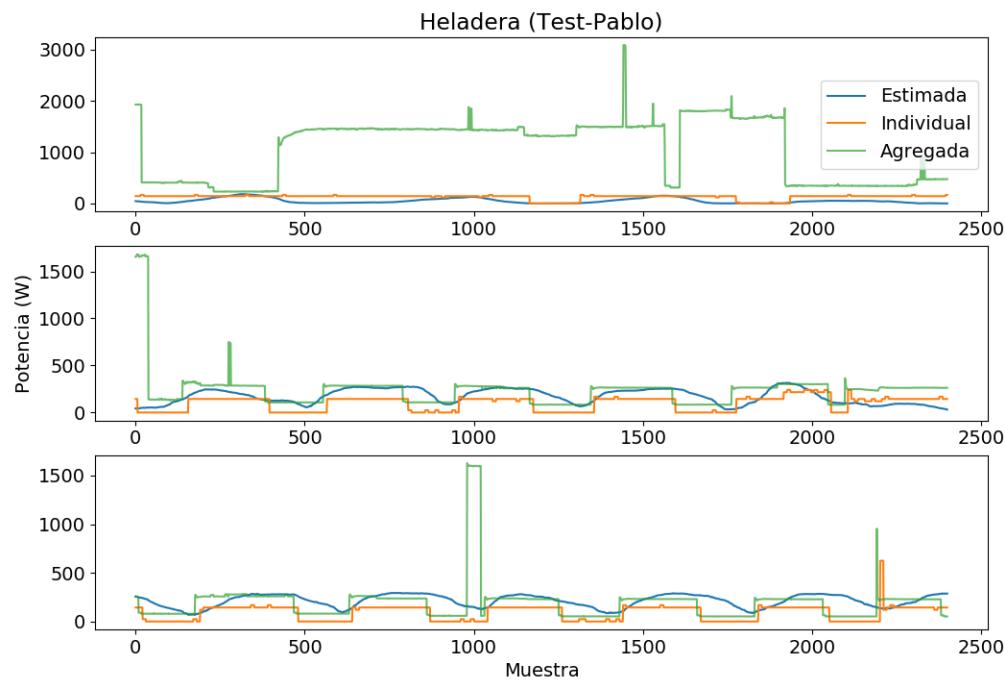


Figura E.50: Ejemplo Heladera Ventanas Deslizantes Test casa Pablo

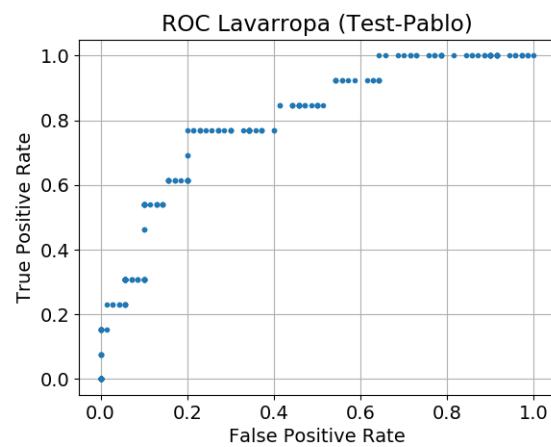


Figura E.51: ROC Lavarropas Ventanas Deslizantes Test casa Pablo. $AUC = 0,8077$

E.5. Resultado de test en ventanas deslizantes sobre NILM-UY

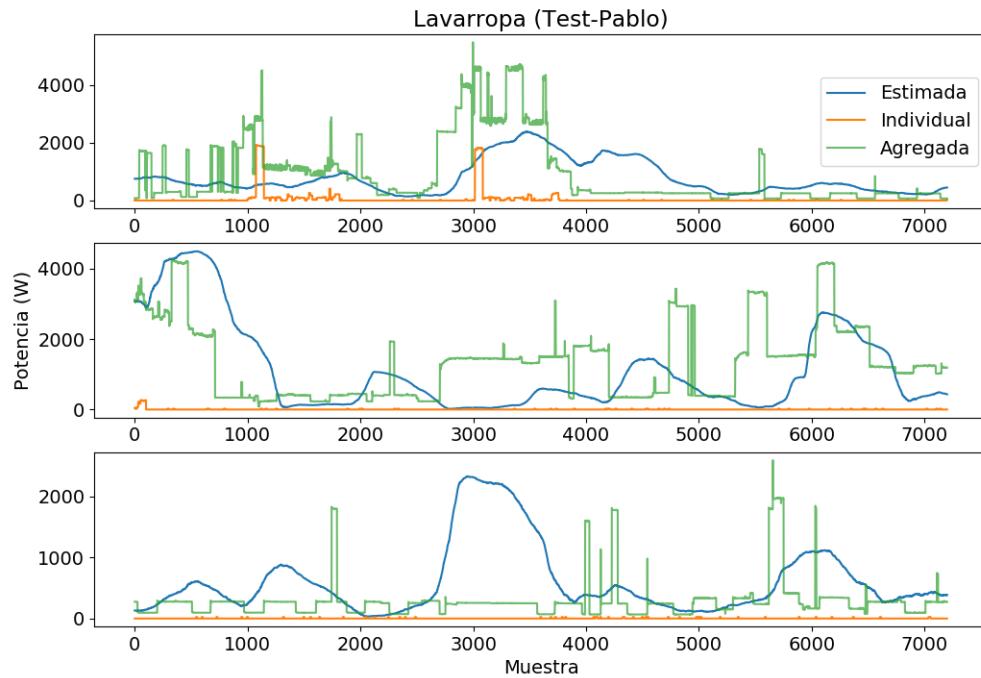


Figura E.52: Ejemplo Lavarropa Ventanas Deslizantes Test casa Pablo

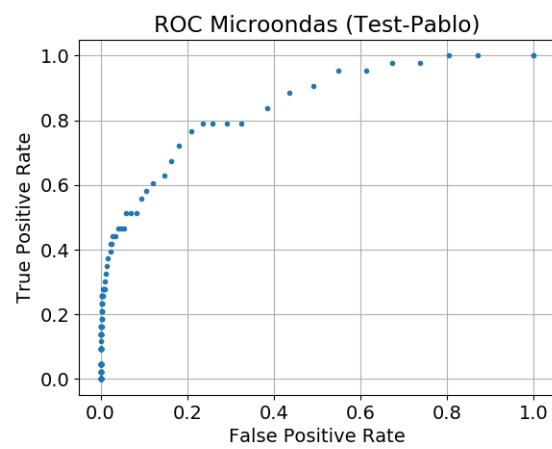


Figura E.53: ROC Microondas Ventanas Deslizantes Test casa Pablo. $AUC = 0,8482$

Apéndice E. Resultados de los algoritmos

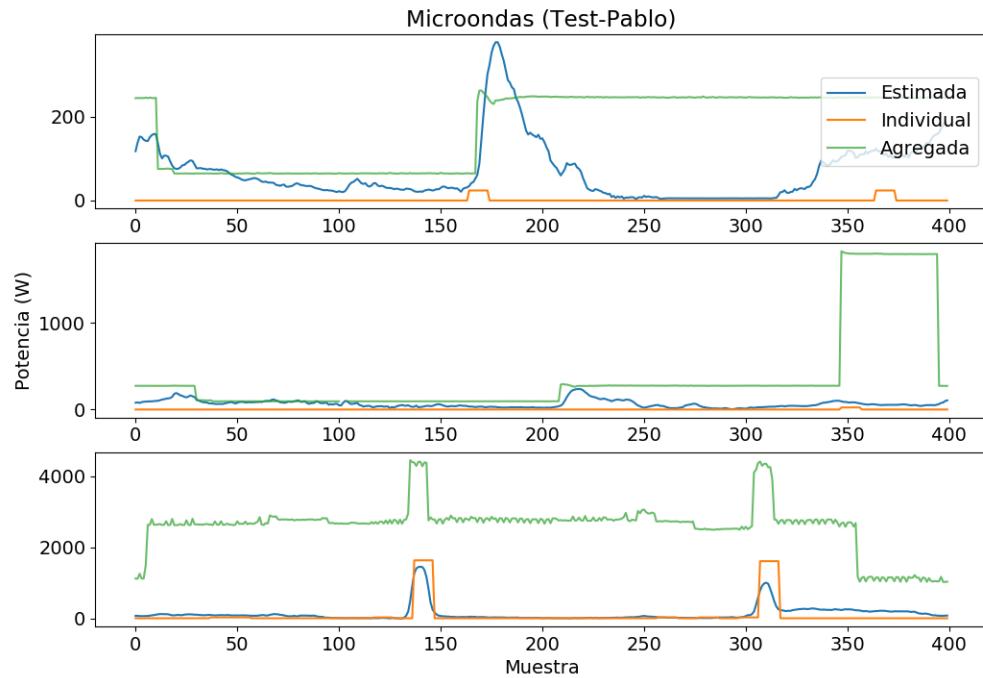


Figura E.54: Ejemplo Microondas Ventanas Deslizantes Test casa Pablo

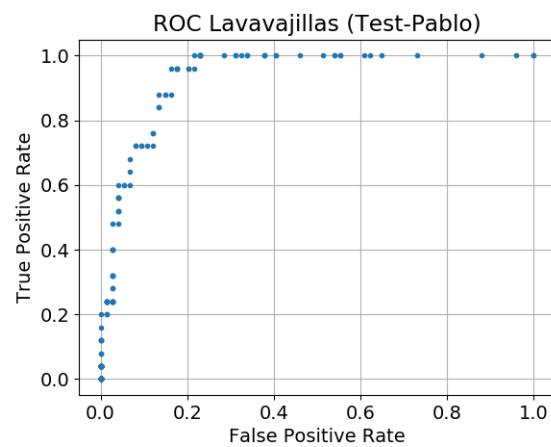


Figura E.55: ROC Lavavajillas Ventanas Deslizantes Test casa Pablo. $AUC = 0,9376$

E.5. Resultado de test en ventanas deslizantes sobre NILM-UY

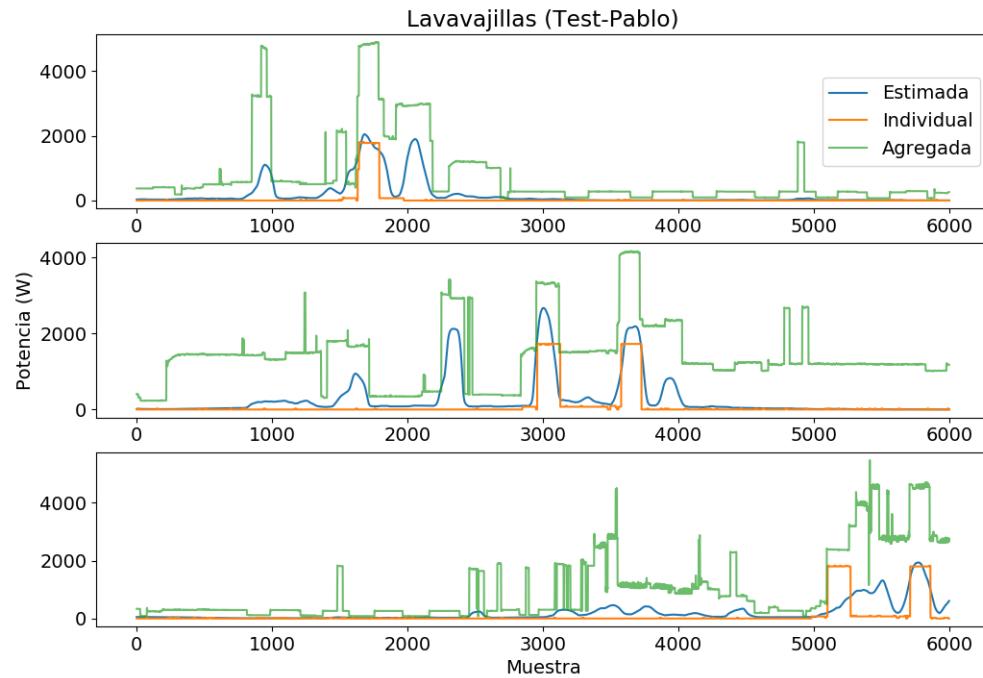


Figura E.56: Ejemplo Lavavajillas Ventanas Deslizantes Test casa Pablo

Apéndice E. Resultados de los algoritmos

Apéndice F

Comparación contra resultados de referencia

F.1. Resumen

En esta sección de evaluación se presenta una comparación de los resultados obtenidos por los autores implementando los modelos descritos en [19] contra los resultados reportados por Kelly en su artículo. La comparación se hace solamente para las métricas de desagregación. Lamentablemente, en [19] no se especifica si los resultados reportados son sobre los vectores de activaciones o si se evalúa con una ventana deslizante. Tampoco se reporta que largo de ventana se utiliza en evaluación y más importante aún, no se indica el umbral de potencia utilizado para las métricas de clasificación. Es por esto que solo se realiza la comparación en desagregación. Por el lado de los autores, la comparación será realizada sobre los vectores de activaciones, con los largos de ventana definidos en el Capítulo 7.

Las implementaciones asumen entrenamiento con datos sintéticos, señalados en la publicación de referencia como fuente de regularización. Los modelos a ser comparados son entonces, rectángulos con datos sintéticos, y autoencoder con datos sintéticos, contra las métricas de sus homónimos reportadas en [19].

F.2. Comparación sobre métricas de desagregación

Se realiza la comparación contra el trabajo de referencia sobre métricas de desagregación (también llamadas métricas de regresión). Las mismas son el MAE y REITE. Se observa que, para el caso de rectángulos, solamente en la heladera se consigue un MAE menor que el reportado por Kelly. Por otro lado, solamente en la heladera y en el microondas se encuentra un REITE mayor que el reportado por Kelly. La heladera en este caso obtiene mejor desempeño en una u otra implementación según la métrica que se esté observando. Se nota nuevamente que a pesar de definir el REITE como una cantidad no negativa, Kelly reporta valores negativos, probablemente ignorando el valor absoluto en su fórmula.

Apéndice F. Comparación contra resultados de referencia

Electrodoméstico	Implementación		Kelly	
	MAE (W)	REITE	MAE (W)	REITE
Jarra eléctrica	86	0.069	15	-0.32
Heladera	18	0.251	22	-0.07
Lavarropas	134	0.418	28	-0.65
Microondas	162	0.326	16	-0.23
Lavavajillas	122	0.239	30	-0.53

Tabla F.1: Resultados de desagregación en test visto para rectángulos, comparación implementación de referencia vs resultados reportados en [19]

Por otro lado, para el caso del autoencoder el MAE reportado por Kelly siempre es menor que el conseguido por los autores, aunque la heladera presenta un MAE similar en ambas implementaciones. Esta vez, a diferencia del caso de rectángulos, la heladera tiene un mejor desempeño en REITE, al igual que sucede con el lavavajillas. Se tiene que la métrica REITE da resultados similares para la jarra eléctrica.

Electrodoméstico	Implementación		Kelly	
	MAE (W)	REITE	MAE (W)	REITE
Jarra eléctrica	68	0.023	16	0.02
Heladera	26	0.153	25	-0.35
Lavarropas	148	0.433	44	0.18
Microondas	148	0.363	13	0.06
Lavavajillas	104	0.064	21	-0.34

Tabla F.2: Resultados de desagregación en test visto para autoencoder, comparación implementación de referencia vs resultados reportados en [19]

Para las casas no vistas en el entrenamiento se tiene, en rectángulos, el mismo comportamiento que para las casas vistas sobre MAE: la heladera es el único electrodoméstico que tiene mejor desempeño en la implementación realizada en el Proyecto. De la misma forma, la implementación del Proyecto continúa presentando un REITE más alto cuando se evalúa sobre la casa no vista. La métrica REITE da más baja sobre la implementación de Kelly para todos los electrodomésticos, exceptuando el lavarropas y el microondas. En esta ocasión es el lavavajillas que obtiene resultados similares en ambas implementaciones bajo esta métrica. En el caso del MAE, la heladera consigue resultados similares para ambas implementaciones.

En cuanto a la generalización, el REITE mejora al evaluar la implementación de Kelly sobre la casa no vista para dos electrodomésticos. Más contraintuitivo aún es que ningún electrodoméstico empeora su desempeño en MAE al evaluarse en una casa no vista para esta implementación. Por otro lado, en el caso del Proyecto, dos electrodomésticos mejoran el desempeño en MAE, aunque uno lo empeora. Dos electrodomésticos consiguen una mejora en REITE, y los otros tres deterioran su desempeño. La heladera, nuevamente, obtiene un similar error en MAE.

F.2. Comparación sobre métricas de desagregación

Electrodoméstico	Implementación		Kelly	
	MAE (W)	REITE	MAE (W)	REITE
Jarra eléctrica	89	0.129	7	0.03
Heladera	17	0.385	18	-0.13
Lavarropas	93	0.078	11	-0.74
Microondas	153	0.273	6	0.50
Lavavajillas	135	0.319	30	-0.31

Tabla F.3: Resultados de desagregación en test no visto para rectángulos, comparación implementación de referencia vs resultados reportados en [19]

En esta comparación se muestra que la implementación de la publicación de referencia obtiene peor REITE para la casa no vista, aunque en tres electrodomésticos el MAE mejora. Por otro lado, tres electrodomésticos obtienen mayor y menor error para MAE y REITE, respectivamente, en la implementación del Proyecto.

Electrodoméstico	Implementación		Kelly	
	MAE (W)	REITE	MAE (W)	REITE
Jarra eléctrica	87	0.168	6	0.13
Heladera	30	0.140	26	-0.38
Lavarropas	99	0.156	24	0.48
Microondas	135	0.241	9	0.73
Lavavajillas	175	0.256	24	-0.33

Tabla F.4: Resultados de desagregación en test no visto para autoencoder, comparación implementación de referencia vs resultados reportados en [19]

En resumen, las métricas reportadas para las implementaciones del Proyecto y del artículo de referencia no son comparables en general. Pocas veces se encuentran valores similares según ambas implementaciones, aunque siempre en una sola de las métricas (la otra difiere). Por otro lado, no puede observarse un deterioro en las métricas al evaluar sobre la casa no vista, lo cual resulta contraintuitivo. Muchas veces hay diferencias de hasta un orden de magnitud entre los errores obtenidos por las dos implementaciones. Este fenómeno puede ser producido por el mecanismo de evaluación (no aclarado en el artículo), por la forma en que se generaron los datos sintéticos, o por el largo de ventana que se define *a priori* al entrenar y no se indica al evaluar en el caso de Kelly, mientras que en el caso del Proyecto se define a partir de los datos de entrenamiento exclusivamente. Estas métricas muchas veces no reflejan cuán cerca está el algoritmo de comportarse como se espera (casos con REITE y MAE bajos pueden ser obtenidos por salidas constantes). Finalmente, resulta difícil comprender como la publicación de Kelly et al. muestra valores de MAE tan buenos pero con REITE tan elevados, hay ejemplos como el microondas en test no visto donde se reporta un error medio en potencia de 9 Watts pero un 73 % de error en la asignación de energía, estos resultados son de difícil interpretación.

Apéndice F. Comparación contra resultados de referencia

Apéndice G

Enfoques a futuro

G.1. Acerca de los algoritmos de AI

G.1.1. “*The bitter lesson*”

En su breve *post* con el nombre anterior, Richard Sutton habla de la ventaja que tienen los métodos genéricos por sobre aquellos específicos que involucran mucho esfuerzo de programación y conocimiento humano. Las siguientes tres frases ilustran su punto:

- “*The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin.*”
- “*(...) building in how we think we think does not work in the long run.*”
- “*breakthrough progress eventually arrives by an opposing approach based on scaling computation by search and learning*”

Aunque controversial, el *post* contiene una idea que los autores comparten: los resultados obtenidos por algoritmos de AI que son incapaces de mejorar con la escala serán superados. La escalabilidad refiere al poder de cómputo y a la cantidad de datos.

El problema de NILM, sin embargo, es un problema distinto a los problemas entendidos como los más desafiantes de la “inteligencia artificial”. Archivos de imagen o audio se conocen como no-estructurados. Una serie agregada de corriente, sin embargo, tiene menor complejidad. Se forma, a lo sumo, con varios electrodomésticos (probablemente menos que 100). Además, el comportamiento físico de la red eléctrica en una residencia ya ha sido modelado y estudiado, no así muchos problemas de clasificación que el cerebro humano resuelve sin mayores dificultades. Podríamos entender cómo identificar electrodomésticos y crear una serie de reglas que partan de algún modelado. Sin embargo, esto no es trivial en problemas como el de clasificación de imágenes.

Sin perjuicio de lo anterior, el problema de NILM sigue estando abierto. Se adquirió un Smappee, que supuestamente desagrega la energía consumida por los electrodomésticos de la casa en la que se colocó. No funciona de la mejor manera.

Tampoco debe entenderse como conclusión del artículo de Sutton que los investigadores deben sentarse a esperar que el poder de cómputo crezca. Hay valor en conseguir una buena solución más temprano que tarde. Pero se puede pensar una suerte de contrarrecíproco: aquellos métodos que con más poder de cómputo y más datos no mejoren su desempeño, no serán la solución del futuro.

Apéndice G. Enfoques a futuro

Los puntos a señalar en esta Subsección son

1. El problema de NILM no está resuelto aún.
2. Los algoritmos exitosos del futuro seguramente hagan fuerte uso de los datos y del poder de cómputo disponible.
3. Esto no implica descartar la inclusión de conocimiento a priori o descartar la programación de reglas en partes de la solución, ya que son caminos que, sin dudas, pueden ayudar al algoritmo que sea.

G.1.2. Falencias del aprendizaje supervisado

El enfoque de aprendizaje supervisado siempre va a estar limitado por los datos (etiquetados) de los que se disponga. Hay gente interesada en el *active learning*, que es muy promisorio. Esto implicaría etiquetar estratégicamente algunos datos, y hacer efectiva la colaboración hombre-máquina. Esta es una de las esperanzas de los algoritmos que se basen en datos etiquetados. El otro enfoque que permite que los algoritmos de aprendizaje supervisado escalen es el *self-supervised learning*. En este enfoque la máquina se encarga de etiquetar los datos para luego utilizarlos. Según LeCun, esto puede entenderse como la utilización de un subconjunto de las entradas para predecir otro subconjunto. No ha habido trabajos en el área con este enfoque, es probable que no sea un problema apropiado para el mismo.

La desventaja de depender de datos etiquetados es que recabar bases de datos es costoso. Esa base de datos ha de ser extensa y frecuentemente actualizada para que realmente se logre un buen desempeño. Una opción para nada descartable es la de simular. Por otro lado, ninguna de las publicaciones revisadas se preocupa por la utilización de los datos recabados durante la ejecución de la solución. Se puede conjutar que hay un tope de residencias que pueden disponer de un dispositivo de NILM (por ejemplo, la mitad de las ubicadas en cierta ciudad), pero se está muy lejos de tener tantos dispositivos de NILM como residencias. Esto hace que, en la práctica, la cantidad de datos con los que una empresa que distribuya dispositivos puede llegar a contar es inmensa, si dicha empresa se hiciera con solo una parte del no explotado mercado.

G.1.3. Distintos enfoques

Los autores creen que hay tres posibles enfoques a la hora de encarar el problema de NILM. En primer lugar se encuentra el enfoque tradicional, que involucra detección de flancos, clustering, etc. Este enfoque puede tener parte de aprendizaje supervisado, o no. Este enfoque permite tener una referencia de desempeño, pero es limitado porque no integra información contextual y una vez implementado, es incapaz de mejorar (por ejemplo, no se mejora el conjunto de características extraídas).

Un segundo enfoque es el adoptado en este Proyecto. Se basa fuertemente en aprendizaje supervisado, intentando estimar una función que abarca varios pasos, tomando como entrada la serie temporal y teniendo como salida el consumo. Esta función puede implementarse de varias maneras (autoencoder y rectángulos en el caso del Proyecto). Este método es capaz de escalar siempre que haya suficientes datos. En particular, tendrá buen funcionamiento sobre aquellos electrodomésticos que tenga en su base de datos, y no necesariamente se extenderá a electrodomésticos desconocidos. Esta es la principal falencia de estos enfoques. Este segundo enfoque podría, eventualmente, incluir algo del estilo meta-learning, donde se apunte a una buena generalización *out-of-sample*, en lugar de simplemente apuntar a generalizar bien en validación. Se afirma que métodos que se basen en cantidades mayores o menores de electrodomésticos individuales etiquetados siempre tendrán un tope de desempeño establecido por el dataset

G.2. Algoritmo no-supervisado

en cuestión. Superado este tope el método pertenecerá a la categoría de algoritmos genéricos (como sería meta-learning). Una forma de generar muchos datos y asegurar una buena capacidad de generalización es sintetizando series temporales, existiendo formas simples, como la implementada en el Proyecto.

La otra alternativa es un enfoque no supervisado, en particular, un enfoque de aprendizaje por refuerzo. La intención es buscar una solución que mejore con el tiempo. Una solución que mejore con el tiempo que pasa en una casa será muy buena al largo plazo para esa casa. También es posible ir más allá, encontrando una solución que no solamente mejore a medida que pasa el tiempo en una sola casa, sino en todas las casas que comparten el algoritmo. Un algoritmo que sea cada vez mejor, en todos los sitios a la vez, utilizando todos los datos a la vez. Un algoritmo del estilo estará obligado a tener cierta capacidad de generalización y se adaptará en un tiempo no muy largo a cada nueva casa.

G.2. Algoritmo no-supervisado

Es inevitable que un algoritmo del estilo utilice mucho de programación de reglas e información a priori. En otro caso, el algoritmo debería aprender innecesariamente las dinámicas de fondo de la red eléctrica de un hogar (por ejemplo, darle como entrada una señal sinusoidal implicaría complejidad innecesaria). Por otro lado, hay un tope de desempeño que solo puede ser superado con algoritmos genéricos o con una base de datos exhaustiva. En comentarios en la web se lee que los programadores de Sense tenían dificultades en poder detectar la marca de un auto eléctrico que no estaba en su base de datos. Esto indica que su enfoque es exhaustivo. Toda esa inversión de horas hombre es sumamente costosa. La idea es que si un algoritmo aprende lo suficientemente rápido, esta inversión es evitable.

G.2.1. Solución genérica

Ilústrese la solución genérica: se dispone de muchos datos, sin etiquetas. No importa qué patrón es qué, el problema de clasificación desagregado ha sido estudiado y es mucho más simple que el agregado. El problema grande es reconocer electrodomésticos (patrones) sin necesariamente etiquetarlos por su nombre. Comparado con el problema anterior, el problema de qué nombre se les asigna a dichos electrodomésticos es simple.

Si se quiere que haya aprendizaje ha de haber una función a optimizar. El problema admite una función de costo bastante intuitiva. Dado que no se pueden conocer las etiquetas, o fuentes, que producen la curva agregada, no es posible medir el desempeño en desagregación como precisión, asignación correcta de energía, o error cuadrático respecto a la fuente. Sin embargo, sí se puede utilizar un concepto similar al de un autoencoder: hacer que la detección permita reconstruir la serie temporal con gran fidelidad. Cabe notar que nunca se elude del todo la inclusión de conocimiento a priori. La métrica objetivo, de por sí, ya indica un camino. La función de costo deberá ser, necesariamente, un error de reconstrucción, ya que el único *ground truth* del que se dispone es la serie agregada. Esta función de costo es uno de los componentes indispensables de la solución.

Otro componente indispensable de la solución es el que la hace útil. Se habla de reconstrucción, y la reconstrucción se construye a partir de algunos valores. Estos valores han de ser tales que incluyan información relevante. Por ejemplo, un algoritmo que estime rectángulos tiene información útil. A partir de las estimaciones de potencias media, inicios y fines, se intenta reconstruir la señal agregada. La reconstrucción ha de ser a partir de información relevante y entendible.

Apéndice G. Enfoques a futuro

Por último, el método debe tener una forma de aprender. Dada una función de costo, modificar los parámetros y observar el nuevo valor es un enfoque posible. Este tipo de optimización es necesaria para que el algoritmo aprenda continuamente. Pueden usarse métodos más o menos complejos, y no siempre se garantiza que el tiempo de aprendizaje sea razonable. Tener una forma de aprender eficaz es de vital importancia.

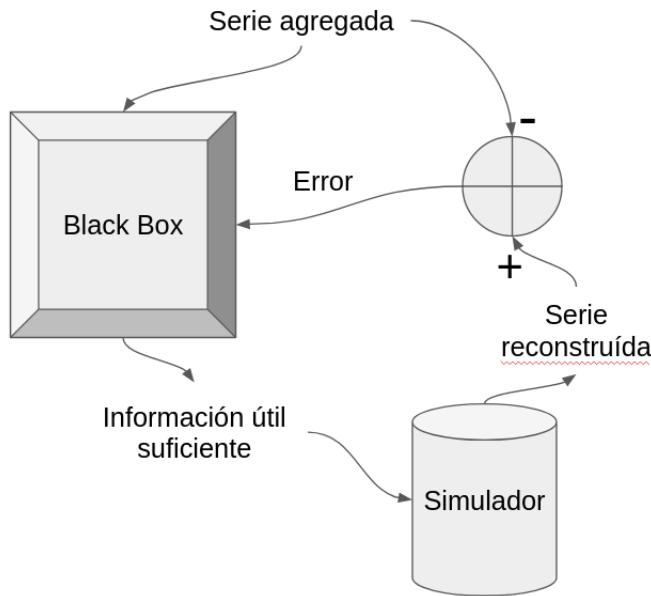


Figura G.1: Solución genérica

La Figura G.1 ilustra el tercer enfoque. La estructura básica del algoritmo es la siguiente: consume los datos de la corriente agregada, y opcionalmente del voltaje, de una o más casas. A partir de los datos, logra llegar a la estimación de los valores deseados (a definir). A partir de los valores deseados, reconstruye la serie agregada. Se compara la serie agregada sintetizada con la serie agregada real y se utiliza la señal de error para aprender (ajustar parámetros).

Sin dudas, la formulación de la solución como expuesta arriba es demasiado genérica. No se apunta aquí a implementar la solución ni a resolver cada uno de sus componentes. Solamente se pretende proponer la visión de lo que debería ser una posible solución escalable en NILM, que eventualmente, dados suficientes datos y poder de cómputo, pueda llegar a reproducir fielmente la serie agregada real a partir de las estimaciones convenientes.

Más abajo se presenta un método que cumple con estos tres requerimientos, aunque falla en la eficacia del aprendizaje. El método utiliza una función de costo en reconstrucción. Utiliza información utilizada para reconstruir que es útil y entendible. También es optimizable. Su falencia es que la optimización lleva demasiado tiempo debido a que no hay una forma eficaz de aprender. En particular, no se puede realizar descenso por gradiente.

G.2. Algoritmo no-supervisado

G.2.2. Una implementación fallida

Sequence to Point Binary Autoencoder

Una posible propuesta es implementar un autoencoder del estilo *sequence to point*: codificar en N electrodomésticos (variables binarias) toda la información, y recomponer un punto de la señal original simplemente combinando linealmente estas variables. A esta propuesta se le llamará “autoencoder binario”.

En general, diferentes hogares tendrán diferentes electrodomésticos. Lo que se puede hacer es entrenar un modelo con varias cabezas, donde el encoder sea compartido y el decoder sea específico para cada casa. En este sentido, se pueden aprender pesos específicos para la combinación lineal del vector binario para cada hogar. Sin embargo la obtención del vector binario es aprendida desde muchas casas a la vez. Eventualmente, entrenando en todas las casas a la vez, mediante minimización alternada, por ejemplo, ese vector binario consistirá en indicadores de presencia de los electrodomésticos más relevantes que son comunes a las casas.

La idea detrás de esta formulación es que la señal agregada puede componerse como suma de electrodomésticos con cierta potencia estática. Esto son los electrodomésticos tipo 2 presentados en el Capítulo 2, de característica ON-OFF. El encoder debería aprender a reconocer activaciones con cierta similitud, que pesadas luego por los pesos específicos de cada casa, redunden en una buena reconstrucción.

Implementación

En esta parte se comentará acerca de la implementación realizada, que consistió en un problema de juguete: descomponer en bits algún número entero.

Primeramente se implementó el autoencoder binario en PyTorch. El principal problema de este enfoque es que el binarizar el espacio latente impide la propagación del gradiente hacia atrás. Es por esto que se exploraron dos caminos.

El primero de ellos mediante el uso de la librería nevergrad, de Facebook Research. Esta librería está orientada a la optimización de funciones por métodos independientes del gradiente. Entre ellos hay algoritmos genéticos y optimización bayesiana. Con esta librería se pudo codificar enteros de hasta 3 bits en menos de 10 minutos, pero enteros de 8 bits no parecían manejables, por lo que se descartó el enfoque.

Por otro lado se intentó suavizar el requerimiento de que el código sea binario. Para ello se programó un autoencoder tradicional, pero se incluyó en la función de costo un término que fuerza al espacio latente a ser binario.

La función de costo original incluye la clásica función de costo de reconstrucción $L_r = (\hat{y} - y)^2$. La función de costo total queda como sigue:

$$L = L_r + \alpha \sum_i (|l_i - 1/2| - 1/2)^2 \quad (G.1)$$

donde l_i es la variable indexada por i del espacio latente, que se desea binarizar. Teóricamente, cuando existe una solución con código binario que sea óptima de L_r , este método debería encontrarla, ya que esta solución también minimizaría L . Sin embargo, aparecen problemas de optimización.

El problema de optimización surge de quedarse en mínimos locales dados por un código con valores cercanos a 0 o a 1 pero no del todo cercanos. Un pequeño error alrededor de 1 para un bit que esté pesado por un gran valor permite ajustar la salida a gusto.

Con la intención de solucionar este problema se comenzó por optimizar comenzando con el valor de α en 0. Una vez que el error dejaba de bajar, se aumentaba el valor de α mínimamente, para volver a esperar la convergencia. Este método se probó mientras se analizaba:

Apéndice G. Enfoques a futuro

1. la función de costo global L
2. la función de costo de reconstrucción L_r
3. la función de costo binaria L_b ($L = L_r + \alpha L_b$)

Luego se modificó este método de forma que solo aumentara el valor de α cuando se llegaba a dejar de disminuir L_r y además L_b fuera mínimo. Operando con este tipo de trucos se lograron desempeños razonables con valores cercanos a 0 y a 1, pero el problema del párrafo anterior seguía presente, aunque en menor medida.

El problema es que al aumentar α , el gradiente no apunta en la dirección de cambiar algún bit, que es lo necesario para escapar el mínimo local. Cuando α es grande, el costo de alejar un bit del valor binario que produce un desempeño subóptimo es demasiado grande. El método de optimización basado en el gradiente no tiene esto en cuenta. Optimizadores que logren superar esta desventaja quizás permitan hacer de este método algo potente.

G.3. Black Box

En esta Sección se presentarán lugares donde los autores consideran puede incluirse programación de reglas explícitas. Hay diversos lugares donde esto puede aumentar el desempeño. Este enfoque de programación de reglas no es escalable por sí solo, pero sin dudas sirve para hacer más eficiente un método que, en su completitud, sí aprende.

Uno de los lugares donde puede ser útil aportar cambios de dominio, por ejemplo pasando la señal al dominio de la frecuencia, es en la extracción de características. Un bloque de extracción de características como etapa de procesamiento de los datos es indispensable si se quiere aportar conocimiento a priori del problema. Esto se estudió en varias publicaciones (y en el Capítulo 5).

Otra componente interesante a definir de forma imperativa (aunque no quita que sea aprendida por fuera del algoritmo), es la detección de flancos. La información de cuándo hay un cambio en consumo es sin duda, valiosa, y no es necesario que el algoritmo tenga que aprender a reconocerlo. Nuevamente, una detección de flancos para electrodomésticos individuales se implementó en el Capítulo 5.

Por otro lado, cualquier otro algoritmo de NILM puede ser útil y aportar al desempeño. Es por ello que es previsible que haya un bloque de fusión o combinación de varios algoritmos. Esta etapa de mezcla es previa a la estimación de los valores deseados.

En la definición de los valores deseados (información útil) hay inclusión de conocimiento a priori. Se podría aclarar, como implica la propuesta del *sequence to point binary autoencoder*, la cantidad de electrodomésticos de los cuales se quiere una activación. Este tipo de aclaraciones pueden verse como exigencias a la salida, o como información a la entrada del sintetizador o simulador (en este caso sería “hay solamente N electrodomésticos”). El carácter del sintetizador puede ser más o menos complejo: podría hacer, para un solo punto, una combinación lineal de las componentes de un vector binario (como se presentó arriba). O también podría estimar inicio, fin, y potencia media de N electrodomésticos (de un solo estado). O estimar el estado, de consumo conocido, de cada electrodoméstico (multi-estado), además de su momento de activación. Estimar estados, consumos por estado, activaciones y número de electrodomésticos, en un enfoque aún “rectangular”: la señal será la suma de rectángulos. O en un caso extremo, la sintetización puede implicar estimar toda la serie desagregada.

En general, al sintetizador o simulador se le pasan “modelos” de electrodomésticos. Lo que se tiene que estimar, inevitablemente, es cuándo se activan. Surge un problema con aquellos electrodomésticos cuya carga varía en el tiempo dependiendo, por ejemplo, de la temperatura ambiente. El problema de detectar cuánto están consumiendo dos calentadores (por ejemplo, dos AC) que cambian su potencia en el tiempo dependiendo

G.3. Black Box

de la temperatura es un problema sumamente complejo que no parece ser abordable por un modelo simplista como el de los “rectángulos”. De hecho, estos electrodomésticos pertenecen al tipo 4 (continuamente variables), que ningún método conocido logra detectar con efectividad.

G.3.1. Bloques o subcomponentes

Los problemas a resolver pueden plantearse individualmente, aunque están muy interrelacionados. El primero de ellos es la detección de flancos. Es probable que sea de utilidad y es un paso indispensable para muchos algoritmos de NILM, tanto supervisados (que requieren entrenamiento con medidas que no se consiguen en el hogar del usuario o información del usuario), como no supervisados. Hay varias publicaciones dedicadas exclusivamente a esto.

Otro bloque de importancia es el de la extracción de características. Mapear la serie temporal a una serie de características, o mapear cada flanco a un espacio de características, puede ser muy útil si en ese nuevo espacio la discriminación es más sencilla. Este tema también ha sido extensamente estudiado.

El bloque del algoritmo de desagregación propiamente dicho ha sido el centro de atención del área. Solamente para una casa, con o sin input del usuario, con o sin contar con datos etiquetados. Es sin duda algo que merece seguir siendo estudiado, pero es deseable que el algoritmo sea capaz de recibir una señal de error y cambiar su funcionamiento en función de ella, es decir que aprenda, y que la modificación de sus parámetros no llegue a un óptimo demasiado poco expresivo (como lo haría un modelo muy simple).

Por otro lado, la salida de varios algoritmos puede combinarse. La combinación de estimaciones es una estrategia usual en el procesamiento de señales, y puede encararse con algoritmos de *machine y reinforcement learning*.

Por último está el sintetizador o simulador. Es ese bloque que recibe las estimaciones y trata de reconstruir la señal real. Es también donde se puede colocar mucha de la información a priori y del modelado (asignar número de electrodomésticos, transitorios característicos, etc.). No se han encontrado muchas publicaciones en este sentido.

Cada uno de esos bloques podría recibir una señal de error y mejorar. Un sistema es un todo, y es plausible que el cambiar muchos parámetros a la vez no ayude a mejorar su desempeño global. Es perspicaz hacer optimización de parámetros bloque a bloque. Y es también un desafío abierto el incluir las señales de error de forma de que el *feedback* sea significativo, útil y eficiente.

Apéndice G. Enfoques a futuro

Referencias

- [1] Audio injector. <http://www.audioinjector.net/>.
- [2] Sense. <https://sense.com/>.
- [3] M.E.; Ocneanu A.; Benitez D.; Moura J.M.F. Anderson, K.D.; Bergés. Event detection for non intrusive load monitoring. In *In Proceedings of the 38th Annual Conference on IEEE Industrial Electronics Society*, pages 3312–3317. IEEE, 2012.
- [4] Roberto Bonfigli, Emanuele Principi, Marco Fagiani, Marco Severini, Stefano Squartini, and Francesco Piazza. Non-intrusive load monitoring by using active and reactive power in additive factorial hidden markov models. *Applied Energy*, 208:1590–1607, 2017.
- [5] Agendaless Consulting and Contributors. Supervisor. <http://supervisord.org/>, 2004 - 2020.
- [6] Leen De Baets, Chris Develder, Tom Dhaene, and Dirk Deschrijver. Detection of unidentified appliances in non-intrusive load monitoring using siamese neural networks. *International Journal of Electrical Power & Energy Systems*, 104:645–653, 2019.
- [7] Leen De Baets, Chris Develder, Tom Dhaene, Dirk Deschrijver, Jingkun Gao, and Mario Berges. Handling imbalance in an extended plaid. pages 1–5, 2017.
- [8] Chris Evans. vsftpd. <https://security.appspot.com/vsftpd.html>, 2011 - 2015.
- [9] Marisa Figueiredo. *Contributions to electrical energy disaggregation in a smart home*. PhD thesis, 2014.
- [10] Adrian Filip. Blued: A fully labeled public dataset for event-based nonintrusive load monitoring research. In *2nd Workshop on Data Mining Applications in Sustainability (SustKDD)*, page 2012, 2011.
- [11] Jingkun Gao, Suman Giri, Emre Can Kara, and Mario Bergés. Plaid: a public dataset of high-resoultion electrical appliance measurements for load identification research: demo abstract. In *proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pages 198–199. ACM, 2014.
- [12] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256, 2010.
- [13] George William Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.
- [14] Ming Jin, Ruoxi Jia, and Costas J Spanos. Virtual occupancy sensing: Using smart meters to indicate your presence. *IEEE Transactions on Mobile Computing*, 16(11):3264–3277, 2017.

Referencias

- [15] Matthias Kahl, Anwar Ul Haq, Thomas Kriechbaumer, and Hans-Arno Jacobsen. Whited-a worldwide household and industry transient energy data set. In *3rd International Workshop on Non-Intrusive Load Monitoring*, 2016.
- [16] Matthias Kahl, Anwar Ul Haq, Thomas Kriechbaumer, and Hans-Arno Jacobsen. A comprehensive feature study for appliance recognition on high frequency energy data. In *Proceedings of the Eighth International Conference on Future Energy Systems*, pages 121–131. ACM, 2017.
- [17] Jack Kelly, Nipun Batra, Oliver Parson, Haimonti Dutta, William Knottenbelt, Alex Rogers, Amarjeet Singh, and Mani Srivastava. Nilmtk v0.2: A non-intrusive load monitoring toolkit for large scale data sets. In *The first ACM Workshop On Embedded Systems For Energy-Efficient Buildings at BuildSys 2014*, Memphis, USA, 2014.
- [18] Jack Kelly and William Knottenbelt. The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes. *Scientific data*, 2:150007, 2015.
- [19] Jack Kelly and William J. Knottenbelt. Neural NILM: deep neural networks applied to energy disaggregation. *CoRR*, abs/1507.06594, 2015.
- [20] J Zico Kolter and Matthew J Johnson. Redd: A public data set for energy disaggregation research. In *Workshop on data mining applications in sustainability (SIGKDD), San Diego, CA*, volume 25, pages 59–62, 2011.
- [21] Thomas Kriechbaumer, Daniel Jorde, and Hans-Arno Jacobsen. Waveform signal entropy and compression study of whole-building energy datasets. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, pages 58–67, 2019.
- [22] Hong Yin Lam, GSK Fung, and WK Lee. A novel method to construct taxonomy electrical appliances based on load signaturesof. *IEEE Transactions on Consumer Electronics*, 53(2):653–660, 2007.
- [23] Brian Langenberger. Python audio tools. <http://audiotools.sourceforge.net>, 2016.
- [24] Steven B Leeb, Steven R Shaw, and James L Kirtley. Transient event detection in spectral envelope estimates for nonintrusive load monitoring. *IEEE Transactions on Power Delivery*, 10(3):1200–1210, 1995.
- [25] Roger Light. Mosquitto: server and client implementation of the mqtt protocol. *Journal of Open Source Software*, 2(13):265, 2017.
- [26] Yu-Hsiu Lin, Men-Shen Tsai, and Chin-Sheng Chen. Applications of fuzzy classification with fuzzy c-means clustering and optimization strategies for load identification in nilm systems. In *2011 IEEE international conference on fuzzy systems (FUZZ-IEEE 2011)*, pages 859–866. IEEE, 2011.
- [27] Chao Liu, Adedotun Akintayo, Zhanhong Jiang, Gregor P Henze, and Soumik Sarkar. Multivariate exploration of non-intrusive load monitoring via spatiotemporal pattern network. *Applied energy*, 211:1106–1122, 2018.
- [28] Qi Liu, Kondwani Michael Kamoto, Xiaodong Liu, Mingxu Sun, and Nigel Linage. Low-complexity non-intrusive load monitoring using unsupervised learning and generalized appliance models. *IEEE Transactions on Consumer Electronics*, 65(1):28–37, 2019.
- [29] Stephen Makonin and Fred Popowich. Nonintrusive load monitoring (nilm) performance evaluation. *Energy Efficiency*, 8(4):809–814, 2015.

Referencias

- [30] Preetum Nakkiran, Gal Kaplun, Yamini Bansal, Tristan Yang, Boaz Barak, and Ilya Sutskever. Deep double descent: Where bigger models and more data hurt. *arXiv preprint arXiv:1912.02292*, 2019.
- [31] Shwetak N Patel, Thomas Robertson, Julie A Kientz, Matthew S Reynolds, and Gregory D Abowd. At the flick of a switch: Detecting and classifying unique electrical events on the residential power line (nominated for the best paper award). In *International Conference on Ubiquitous Computing*, pages 271–288. Springer, 2007.
- [32] IEEE Power and Energy Society. Ieee standard definitions for the measurement of electric power quantities under sinusoidal, nonsinusoidal, balanced, or unbalanced conditions. *IEEE Std 1459-2010 (Revision of IEEE Std 1459-2000)*, pages 1–50, March 2010.
- [33] Hamed Nabizadeh Rafsanjani, Changbum R Ahn, and Jiayu Chen. Linking building energy consumption with occupants’ energy-consuming behaviors in commercial buildings: Non-intrusive occupant load monitoring (niolm). *Energy and Buildings*, 172:317–327, 2018.
- [34] James Robert. pydub. <https://pypi.org/project/pydub>, 2015.
- [35] Brian C. Ross. Mutual information between discrete and continuous data sets. *PLOS ONE*, 9(2):1–5, 02 2014.
- [36] Antonio Ruano, Alvaro Hernandez, Jesus Ureña, Maria Ruano, and Juan Garcia. Nilm techniques for intelligent home energy management and ambient assisted living: A review. *Energies*, 12(11):2203, 2019.
- [37] Nasrin Sadeghianpourhamami, Joeri Ruyssinck, Dirk Deschrijver, Tom Dhaene, and Chris Develder. Comprehensive feature selection for appliance classification in nilm. *Energy and Buildings*, 151:98–106, 2017.
- [38] SoX. Sound exchange. <http://sox.sourceforge.net>, 2015.
- [39] KH Ting, Mark Lucente, George SK Fung, WK Lee, and SYR Hui. A taxonomy of load signatures for single-phase electric appliances. In *IEEE PESC (Power Electronics Specialist Conference)*, pages 12–18, 2005.
- [40] Robert Wall. *A Report on the properties of the YHDC current transformer and its suitability for use with the OpenEnergyMonitor system*. Open Energy Monitor, 2018. <https://learn.openenergymonitor.org/electricity-monitoring/ct-sensors/yhdc-sct-013-000-ct-sensor-report>.
- [41] Robert Wall. *A Report on the properties of the YHDC current transformer and its suitability for use with the OpenEnergyMonitor system*. Open Energy Monitor, 2018. <https://learn.openenergymonitor.org/electricity-monitoring/ct-sensors/files/YhdcCTReportIss6.pdf>.
- [42] Hong-Tzer Yang, Hsueh-Hsien Chang, and Ching-Lung Lin. Design a neural network for features selection in non-intrusive monitoring of industrial electrical loads. In *2007 11th International Conference on Computer Supported Cooperative Work in Design*, pages 1022–1027. IEEE, 2007.
- [43] K. Yumak and O. Usta. A controversial issue: Power components in nonsinusoidal single-phase systems. In *2011 7th International Conference on Electrical and Electronics Engineering (ELECO)*, pages I–157–I–161, Dec 2011.
- [44] Michael Zeifman and Kurt Roth. Nonintrusive appliance load monitoring: Review and outlook. *IEEE transactions on Consumer Electronics*, 57(1):76–84, 2011.

Esta es la última página.
Compilado el lunes 1 junio, 2020.
<http://iie.fing.edu.uy/>