Supervised Machine Learning in Inter-Level, Ultra-Low Frequency Power Line Communications

Kushal Thapa[*], Stan McClellan[†], Damian Valles[§]

Ingram School of Engineering

Texas State University

San Marcos, TX, USA

email: [*]k_t260@txstate.edu, [†]stan.mcclellan@txstate.edu, [§]dvalles@txstate.edu

Abstract — **Power Line Communications (PLC) is a technology that uses power lines to carry communication data alongside electrical signals. This technology has a huge potential and low infrastructure cost due to the pre-existing and ubiquitous power grid. However, the electrical components in the grid and the highly dynamic properties of the grid make power lines a hostile medium for PLC, especially when the PLC signal path extends over multiple levels (current or voltage) of the grid. Subsequently, efficient transmission of the PLC signals from the transmitter end and effective demodulation at the receiver end are both challenging. In our research, we limit PLC transmission frequencies to the ultra-low spectrum and investigate supervised machine learning as a potential signal demodulation technique. Employing inter-level PLC architecture, we transmit and collect baseband-modulated data, then use various machine learning algorithms to recover the data. We also investigate various feature-based and featureless machine learning methods for PLC and conclude that feature-based methods provide better generalization for our dataset.**

Keywords - ultra-low frequency power line communications; machine learning; smart grid; feature based learning; featureless learning.

I. INTRODUCTION

The following journal article is an extended version of a previous conference publication [1] and is partially derived from a thesis report [2].

Power Line Communications (PLC) is a technology that enables the transmission of communication signals over power lines. This technology has been in use since the early 1920s for various telemetry and utility applications [3][4]. One major advantage of this technology over other communication media, such as phone lines, optical fibers, etc., is the ubiquity of available infrastructure, i.e., the wires of the power grid itself. However, the grid is also a primary hindrance to the widespread deployment of PLC because of its heterogeneous nature, and dynamic load and noise characteristics [5]. These factors pose difficult challenges in the efficient propagation of PLC signals.

The heterogeneity of the grid comes in the form of varying voltage and current levels, different impedance profiles and circuit architectures, and various grid devices such as power transformers. Transformers cause a particularly difficult problem in signal propagation, especially for higher-frequency PLC signals [6]. This problem can be circumvented by excluding these devices

altogether and bounding the PLC communication path within a single "level" of the power grid. In such an "intra-level" PLC, the system voltage, circuit architecture, and other properties of the power signal are consistent throughout the path of the communication signal, effectively simplifying the problem and allowing for the use of conventional communications techniques. As a result, intra-level PLC has developed into many applications such as powerline Local Area Network (LAN) [7] and baby monitors [8]. Unfortunately, this solution does not address the problem of signal propagation through the different levels of the power grid i.e., "inter-level" PLC.

One existing solution to this propagation problem in inter-level PLC systems is the use of additional equipment like signal repeaters to bypass power transformer interconnection points. These repeaters filter and catch the communication signal in one level and pass it to the next level so that the integrity of the signal is preserved [9]. Broadband over Power Line (BPL) technology, which is a duplex and high data-rate application of the inter-level PLC, is reliant on such devices [9]. However, the cost and complexity associated with developing, deploying, and maintaining these devices have effectively obviated the use of BPL in modern grid communications [10][11]. Another approach, which does not require inter-level bypass devices, is the transmission of Ultra-Low-Frequency (ULF) signals directly over the power line infrastructure. ULF signals do not attenuate as much compared to higher frequency signals, even after passing through distribution (power) transformers, which have a fairly limited passband. Thus, ULF signals, typically in the range of 150 Hz-1350 Hz [12], can be injected into the power lines to effect a low-bandwidth form of PLC. In fact, this approach is the main principle behind a venerable, simplex, low data-rate PLC application called "Audio Frequency Ripple Control" (AFRC) that has been used in many locales worldwide for load peak shaving, street lighting, etc. [12].

A major disadvantage of using ULF bands for communication is the low data rate. However, low-rate communications techniques have been shown to have substantial value in command & control, distributed system management, and other applications, which are distinct from the "bearer-plane" of conventional communications systems [13]–[15]. There is an urgent need for this kind of technology in the power sector. The existing power grids are failing because of the exponential increase in power demand over the last few decades and the resulting issues [16]–[18].

This problem is exacerbated by the disconnect between the power producers and power consumers, which leads to serious supply chain problems. ULF-PLC can help in bridging this disconnect by providing real-time, ubiquitous, low-rate communications facilities, thereby becoming a valuable element in the communication infrastructure of a continuously sensing and self-monitoring power grid called "the smart grid."

ULF applications like AFRC employ a simplex "downstream" format, i.e., the communication signal propagates from a High Voltage (HV) or Medium Voltage (MV) tier to a Low Voltage (LV) tier. A different approach to communication is an "upstream" method, where the signal originates in the LV tier and is collected at the MV/HV tier, typically a local substation [19][20]. For upstream inter-level PLC, assuming the communication signal effectively transits between the tiers of the distribution grid, then various Digital Signal Processing (DSP) techniques can be utilized to capture, demodulate, and decode the signal at the substation. One complexity of this approach is that DSP-based techniques would have to be unique to that the deployed source-to-destination power grid architecture. In many cases, this complexity would be exaggerated by a requirement to be tailored to individual signals or modulation techniques. This complexity is present because every element in the power grid introduces its own time-varying interference, which is coherent with the fundamental of the power signal. This issue is well-known in terms of distribution grid architecture, but not well-understood in terms of signal processing [21][22]. Thus, an approach driven by DSP alone is not scalable. To make ULF-PLC technology more universally applicable, Machine Learning/Neural Network (ML/NN) algorithms can be employed individually or in concert with conventional DSP-driven approaches [23]. The use of ML/NN algorithms may be particularly beneficial in this environment because they can estimate and dynamically adjust various parameters to correctly reconstruct the transmitted information.

Depending on the application, PLC can benefit from the full spectrum of ML types and techniques (see [23]). In this research into the use of ML/NN in ULF-PLC, supervised learning techniques are examined for two reasons: (1) the availability of labels for the training dataset, which provides an accuracy metric, and (2) the ease of implementation, which provides a path to deployment in a practical setting. The classification task of the ML algorithms was to classify a simple On-Off Keyed (OOK), upstream, inter-level ULF-PLC signal into its on and off states, thereby determining the sequence of information bits transmitted via the power line channel. Although we evaluated some unsupervised or reinforcement learning approaches, the clear benefit of supervised learning in this application provided substantial advantages. As part of the investigation, we examine PLC signal characteristics in the ULF band, explore various ways of processing PLC data to be used in ML/NN, and compare the performance of some common ML/NN algorithms in this domain.

The rest of this paper is organized as follows. Section II provides a brief survey and evaluation of the relevant literature. Section III describes the methodology used in the collection and processing of our inter-level, upstream ULF-PLC data. Section IV presents the ML/NN framework used in our experimental study. The results of these experiments, alongside their corresponding discussions, are presented in Section V. Finally, the overarching conclusions of the research are presented in Section VI.

II. LITERATURE SURVEY

This section surveys the existing body of research related to PLC. Although the archival literature covers a wide variety of research in various facets of PLC, we focus on low and ultra-low frequency, narrow-band PLC to survey existing implementation methods.

Korki et al. [24] developed a ULF-PLC model for an LV distribution power grid network involving mathematical analysis of the grid components. To evaluate the proposed model, Bit Error Ratio (BER) versus Signal-to-Noise Ratio (SNR) was investigated through numerical simulation. The simulation results showed that the model had acceptable BER (compared to other communication standards) for communication between neighboring nodes when SNR was above 15dB. As a result, the authors concluded that their analytical approach can satisfactorily model ULF-PLC in the LV region. This research applies a 'bottom-up' approach to PLC since it builds a simulation model incorporating the individual parts of the network. Its counterpart, a 'top-down' approach, uses measurement and data to estimate the nature of the PLC network. The bottom-up approach is harder to generalize and scale to arbitrary PLC networks because of the highly variable and dynamic nature of the PLC medium. In fact, the model proposed in [24] leverages a highly specific scope (e.g., PLC in a one-phase LV tier between one-hop nodes), which produces very limited results. Therefore, in the present work, a top-down method is used in hopes of producing a generalizable ULF-PLC solution applicable to a wider, less-specific grid domain.

Varadarajan et al. [25] investigated the channel and noise characteristics of the physical layer of PLC systems in the distribution grid. The characterization of channel and noise in the Federal Communications Commission (FCC) regulated PLC frequency band (9kHz to 150kHz) was done with empirical measurements from field tests using a modem, utility meters, and measurement equipment such as spectrum analyzer and oscilloscope. The researchers observed that the presence of distribution transformers causes severe frequency-selective attenuation of the PLC signals, which corroborates our observations. Hence, they concluded that focusing on the less attenuated (lower frequency) sub-bands can increase the quality and the coverage of the transmitted PLC signals. In contrast to [24], this study uses a top-down, empirical approach to characterize the PLC channel, which is simpler than an analytical bottom-up approach. However, the results discussed in [25] are not "constructive" and so do not provide a solution to overcome the challenges of PLC, instead providing observations and suggestions for implementation. In contrast, the present research considers applicable results from [25] (e.g., the effect of transformers

in ULF-PLC) while also exploring several approaches, which can comprise a solution framework to leverage and categorize generally applicable ML-based approaches.

Prasad et al. [26] evaluated the use of Power Line Modems (PLM) for detecting faults in the power grid. To do so, the authors collected raw data containing diagnostic parameters from the PLM, then extracted features from the raw data and used ML algorithms to identify and predict faults. This approach was able to detect faults with high accuracy and predict future faults with a lower, but satisfactory, accuracy. Similar to [24], this study uses mathematical analysis to model PLC architecture. Additionally, similar to [25], this study uses a top-down approach to detect faults through ML. In contrast to [24] and [25] however, [26] also presents a direct application of PLC in terms of useful grid diagnostics. However, [26] is limited in scope and future utility because the associated communication technology (i.e., BPL) has become obsolete [10][11]. Therefore, the present research attempts to evaluate technologies that may be leveraged in a more general implementation of ULF-PLC.

Nassar et al. [27] reviewed the signal processing techniques used to model 3-500 kHz range narrowband PLC (NB-PLC) and presented the local utility applications of NB-PLC using multi-path modeling and transmission-line modeling. In [27], PLC noise is characterized in terms of its periodicity and uncoordinated interference, and a framework for implementing Orthogonal Frequency-Division Multiplexing (OFDM) in NB-PLC is provided. The authors conclude that channel and noise modeling can be used to make robust PLC for two-way communication between the customers and the local utility in the higher frequency bands (e.g., not ULF-PLC). This study uses the channel and noise modeling approach similar to [24] while focusing on NB-PLC above 3kHz for utility applications. Although the approach and outcomes of [27] are comprehensive, no novel techniques or experimentation were introduced. Further, the applicability of spectrum above 3kHz is suspect in the distribution grid due to the large, series inductance presented by power transformers. In contrast, the present work focuses on ULF-PLC and attempts to leverage findings from the various approaches described in the literature.

In summary, two main methods of PLC implementation are prevalent: a 'bottom-up' approach, which leverages mathematical modeling and simulation and a 'top-down' approach, which leverages empirical measurement to develop useful models and evaluate technology performance. The bottom-up approach is network and application specific, subject to component modeling errors or omissions, cannot be easily scaled, and requires complex simulation. The top-down approach is more easily scalable, simpler to implement, and doesn't suffer from component modeling errors. Therefore, the objective of the present work is to employ a top-down approach, including a combination of DSP techniques and ML/NN to aid in discriminating inter-level, ULF-PLC communication signals in the distribution grid.

III. DATA

In this section, the methodology used to collect the raw data and the processes used to convert or transform the raw data to ML-ready datasets are described.

A. Transmission and Collection

The data transmission and collection architecture of the inter-level, upstream ULF-PLC is shown in Figure 1. As shown in this figure, baseband modulated ULF-PLC signals were transmitted from a research lab (LV tier) through the distribution grid and collected at a remote substation (MV tier). To create the experimental datasets, an Ametek CS3000 programmable current source [28] was used to generate On-Off Keyed (OOK) current signals with frequencies varying from 690 Hz to 2010 Hz and various amplitudes. These signals were injected into the split-phase, LV (120VRMS) power line via a specially constructed narrowband filter. This filter is a transformer-based bandpass filter or "resonant tank," which blocks the power signal from entering the output port of the current source while allowing the signal generated by the current source to pass upstream into the power line. The signal injected into the power line first enters a split-phase distribution link typical of conventional "wall outlets," then into the local distribution grid via three-phase distribution transformers, which "step up" the voltage for longer distance transmission. At each "level change" (tier), the transformer jointly adjusts the voltage (upward) and current (downward), according to the structure of the transformer's coils (turns ratio). Some signature of the transmitted ULF-PLC signal is introduced into all three phases of the distribution and transmission links during this voltage transition as a result of magnetic cross-coupling within the transformer core [6]. In the substation, current transformers (CT) affixed to MV (13kVRMS) feeders sense the current disturbances introduced by the injected ULF-PLC signal. This current signal was then discretized using a conventional Data Acquisition device (DAQ) [29].

B. Raw Time-Series Data

The raw data captured using substation-resident, remote DAQ comprised of a set of three-phase time-series data sequences containing samples of the 60Hz power signal and its harmonics, the injected communication signal (ranging from 690Hz-2010Hz) and its harmonics, and time-variant noise at all frequencies. The acquired signals were sampled at 8kHz using 16-bits of resolution per sample (see more information about the capture parameters in Appendix B). The fundamental of the power signal dominates the raw data because of its extremely high amplitude relative to the injected signal, as shown in Figure 2. The three-phase raw time-series data plotted in this figure contain the PLC signal injected at 1595Hz, but this is not apparent from this plot alone. The vertical scale of the plot is in millivolts (mV) due to the nature of the output signals of the sensing current transformer where the full-scale of 1250mV corresponds to several hundred amperes, whereas the injected signal amplitude was lower than 1A.
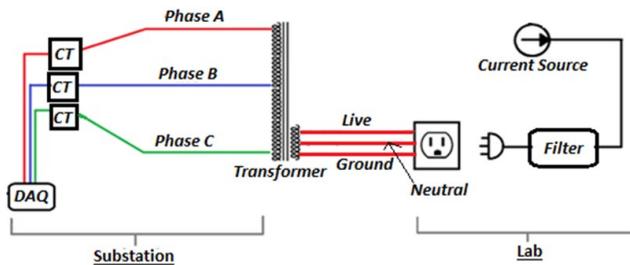
Figure 1.   Data transmission and capture architecture for upstream, inter-level ULF-PLC.
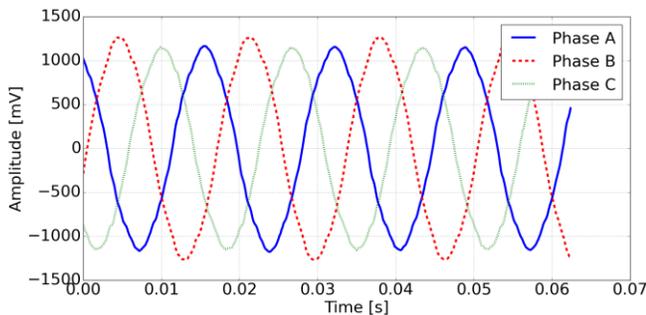
Figure 2.   Time-domain plot of three-phase data received at the substation with PLC frequency at 1595Hz. The vertical scale of the plot is in millivolts (mV) due to the nature of the output signals of the sensing current transformer (1250mV corresponds to several hundred amperes).

C.   Processing

Since the raw-time series data were predictably not very useful in our case, the next logical step was to transform the signal into other domains and/or extract useful information (or features) from it.

1)   Transformation

Transformation, for our purposes, is defined as the conversion of time-series data to frequency domain or time-frequency domain data. There is very little loss of information in these kinds of transformations as long as care is maintained in the representation and storage of the intermediate or final values. The frequency and time-frequency transformations of our original time-series data are described in subsequent sections.

a)   Frequency Domain

The time-domain raw data were transformed to the frequency domain using the Fast Fourier transform (FFT) [30]–[32]. Figure 3 shows a representative example of the frequency domain plot of a single-phase sequence containing a ULF-PLC signal transmitted at 1595Hz. As seen in this figure, the power signal and its harmonics dominate the spectral plot. Nonetheless, as indicated by the red arrow in Figure 3, a smaller but prominent peak is present at 1595Hz, which shows the presence of our transmitted PLC signal. However, the spectrum plot cannot show the time-varying nature of the signal, and thus does not provide us information about the OOK encoded data that were transmitted. Hence, a single frequency-domain transformation of the acquired signal may be effective in

detecting the presence or absence of the PLC signal but cannot be used to decode the time-varying information contained within it. Further, errors in framing synchronization in composing the frequency-domain transformation can produce unwanted or misleading side effects.

b)   Time-frequency Domain

The deficiencies of frequency domain transformation, i.e., the inability to decode OOK information in our case, can be mitigated by appropriate time-frequency analysis such as a spectrogram. The transformation methods used to produce a spectrogram are very similar to the frequency-domain conversion, except the FFT is applied to frames or sub-sequences within the sequence instead of the entire time-series sequence (e.g., Short-Time Fourier Transform, or STFT). This approach provides some indicative information regarding the presence or absence of the PLC signal in those time frames. Therefore, concatenating the spectral analysis of appropriately overlapped and windowed time-domain frames across the length of the data can be used to adequately estimate the spectral configuration of the data.

In practice, there are a few measures that can be taken during this time-frequency transformation to enhance the information contained in the raw data. Windowing the time-domain frames with window functions can minimize the spectral leakage effect [33]–[35]. Similarly, overlapping the time-domain frames prevents information loss caused by the tapered ends of the windowing functions. Parameters related to these techniques, such as window type, frame length, overlap length, as well as other parameters such as frequency resolution affect the quality of the spectrograms, and thus, their values need to be carefully chosen (see Appendix C for these values for present work). Since spectrograms can be produced using Fourier (or similar) transforms, the resulting datasets may have complex-valued representations, with both real and imaginary components. Equation 1 [36] shows the mathematical representation of a complex number 'z,' where 'a' is the real component, 'b' is the imaginary component, and 'i' is the indeterminate satisfying $i^2 = -1$.

$$z = a + i*b \qquad (1)$$
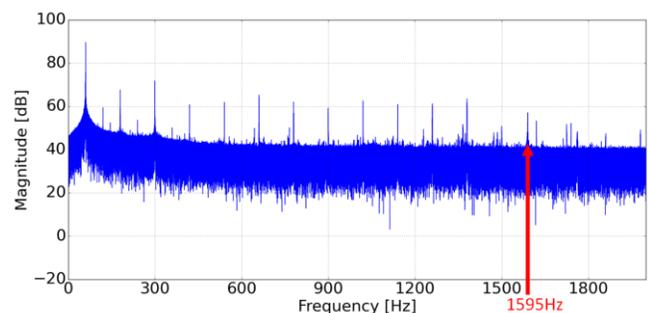
Figure 3.   Frequency domain plot of the raw data showing the presence of the ULF-PLC signal at 1595Hz (pointed by the red arrow). The highest peak in this plot shows the dominant power signal at around 60Hz. The odd harmonics of this power signal can be seen as the secondary peaks.

In ML/NN implementation, these complex-valued data pose a critical challenge since the traditional ML/NN algorithms and architecture are typically not structured to operate on complex-valued datasets. One way to overcome this challenge is to devise new ML/NN frameworks, which accommodate complex-valued inputs and have complex-valued weights or coefficients. This is an active area of research [37][38] but out of scope for the present work. An easier, more accessible approach is to use an appropriate, real-valued concatenated or pre-processed representation of the complex-valued data as input to conventional ML/NN structures. In this research, we utilize two different real-valued formats for the complex-valued transform data: the magnitude/phase or polar format, and the rectangular or cartesian format.

i) Magnitude/Phase (Polar) Format

The most common method of converting complex-valued data to real-valued data is by combining the rectangular or cartesian real and imaginary components via a polar transformation. In polar form, the magnitude of a complex number is represented by the square root of the sum of squares of the real and imaginary components, and the phase is represented by the arctangent of the ratio of imaginary to real components [36].

For transformed time-series data, taking the magnitude of each complex-valued element of a 2D complex-valued spectrogram produces a magnitude spectrogram. This magnitude spectrogram has the same dimensions as its parent complex-valued spectrogram but does not include phase information. Figure 4(a) shows a magnitude spectrogram of one of our captured ULF-PLC datasets where the PLC frequency is at 1595Hz. This figure shows a dotted band at around 1595Hz, which corresponds to the transmitted OOK PLC signals. The zoomed-in version of this band is shown in the inset black box of Figure 4(a), which clearly shows the 'On' and 'Off' states of the ULF-PLC signal. The solid bands at various frequencies correspond to the power signal and its odd harmonics (shown by the red arrows in Figure 4(a)), which are spaced 120Hz apart.

The PLC signal also produces its own harmonics, though not as strong as power signal harmonics. Interestingly, the "harmonics" of the ULF-PLC signal are harmonically offset from the ULF-PLC signal based on the fundamental frequency of the power signal, not the ULF-PLC signal. As a result, they are "echoes" or "images" of the ULF-PLC signal with the harmonic structure described by the power signal's frequency. If the ULF-PLC signal is injected in the frequency bands between the harmonics of the power signal, then a magnitude spectrogram can provide a good estimation of the transmitted information, especially OOK or amplitude-shifted ULF-PLC signal as shown in Figure 4(a). Therefore, the present work uses magnitude spectrograms as one of the ML/NN datasets to evaluate the accuracy of decoding the PLC signals.

As seen from 4(b), the phase spectrogram is very noisy, and the presence of the ULF-PLC signal is not visually apparent. Therefore, phase spectrograms were not used in the present work.

(a) Magnitude Spectrogram

(b) Phase Spectrogram
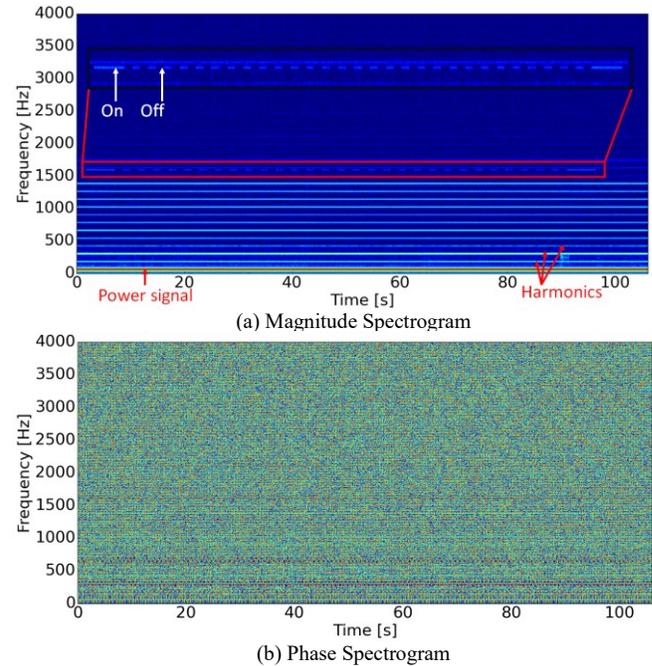
Figure 4. (a) Magnitude spectrogram of the captured data showing the transmitted OOK ULF-PLC signal with the dotted horizontal band at around 1595Hz. The solid bands at lower frequencies correspond to the power signal and its harmonics. (b) Phase spectrogram of the same data.

ii) Rectangular (Cartesian) Format

In addition to magnitude and phase spectrograms, other types of spectrograms (e.g., real and imaginary) can be generated as well by manipulating the components of the complex data. Plots in Figure 5 show the real spectrogram and imaginary spectrogram of the same dataset used for magnitude and phase spectrograms in Figure 4.

(a) Real spectrogram

(b) Imaginary spectrogram
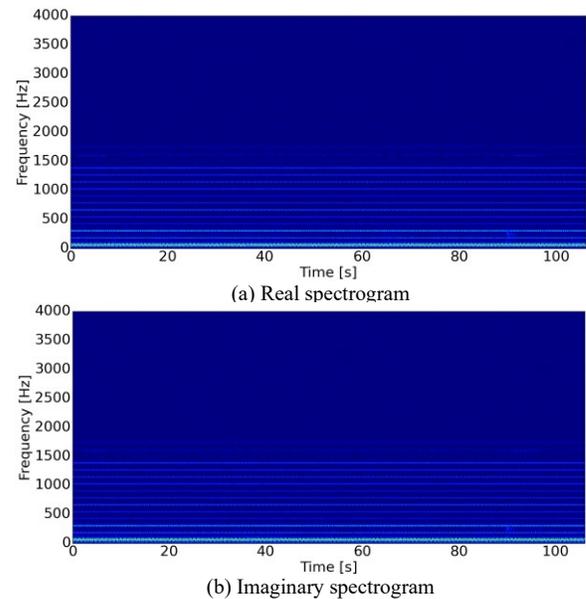
Figure 5. Real spectrogram (a) and imaginary spectrogram (b) of captured data with ULF-PLC signal at 1595 Hz.

The real and imaginary spectrograms show a trace of the PLC signal at 1595Hz; however, this band is not as distinct as the magnitude spectrogram plot in Figure 4. This is expected since the real and imaginary spectrograms completely ignore the other component, resulting in a huge loss of information. However, a combination of these two 2D spectrograms in a 3D space could potentially be better than a magnitude spectrogram since this would be a direct representation of complex-valued data as shown in Equation 1. This 3D spectrogram is hereby referred to as 'Rectangular spectrogram' in the present work because of its correspondence to the rectangular form of the complex-valued transform results.

Figure 6 shows in detail how the rectangular spectrogram (and other spectrograms) may be generated from the raw time-series data (as suggested in [39]). First, the 1D raw-time series data are divided into multiple overlapping frames, which are then windowed. Next, these time-series frames are transformed into frequency space using FFT. The 1D complex-valued FFT products for all frames are then combined to get a 2D complex spectrogram. From this complex spectrogram, real, imaginary, magnitude, and phase (not shown) spectrograms are produced. Finally, the real and imaginary spectrograms are stacked to form a 3D rectangular spectrogram.

In previous work [40], the rectangular spectrogram has been shown to produce better accuracy results than the magnitude spectrogram, especially when the signal of interest is dominated by out-of-band interferers such as the power signal fundamental, which is present in PLC transmissions. Therefore, we used 3D rectangular spectrogram datasets for ML/NN analysis in the present work as well.

2) Feature Extraction

Besides the spectrograms, other signal features including amplitude envelope, zero-crossing rate, spectral centroid, etc. can also be used to extract target information from the received ULF-PLC signals. For the present work, three signal features are used: (1) amplitude envelope, (2) root-mean-square energy, and (3) spectral centroid. These features are briefly described in subsequent sections.

a) Amplitude Envelope

The amplitude envelope of a signal describes the variation in the amplitude of the signal over time [41]. It effectively traces the outline of the signal in the time domain and is loosely representative of the instantaneous energy in the signal [42]. In our case, the complete amplitude envelope is not useful since the 60Hz power signal and its harmonics dominate all other superimposed signals, as described previously. As is typical in communications systems, the narrowband amplitude envelope, centered on specific sub-bands, can be more
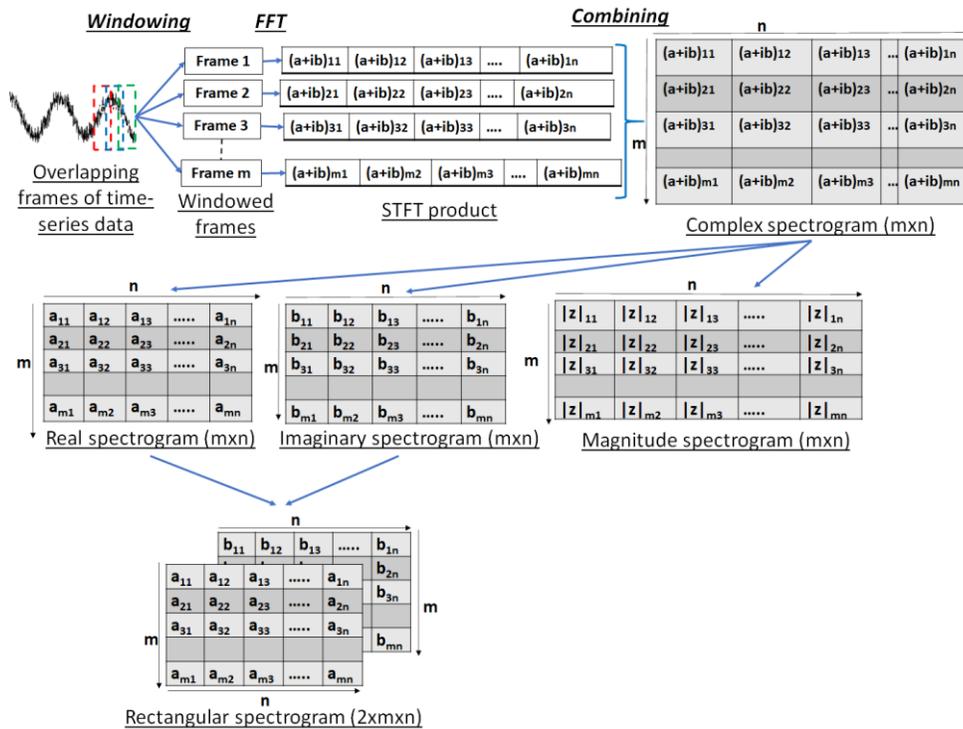
Figure 6. Process of generating rectangular spectrogram from raw time-series data. As shown here, the 1D raw-time series data are first divided into multiple overlapping frames. Then, each of these frames are windowed by a window function. Next, these time-series frames are transformed into frequency space using Fast Fourier Transform (FFT). The 1D complex valued FFT products for all frames are then combined to get a 2D complex spectrogram. From this complex spectrogram, real (denoted by axy), imaginary (bxy) and magnitude spectrogram ($|z|_{xy}$) are produced. Finally, the real and imaginary spectrograms are stacked to form a 3D rectangular spectrogram.

useful. To approximate this approach and create a useful training/testing dataset, we filtered the raw frames of the powerline signal with non-overlapping bandpass filters of 100Hz bandwidth (1Hz-100Hz, 101Hz-200Hz,…,2901-3000Hz) to create thirty sub-band representations of the signal. Clearly, the amplitude envelope of certain sub-bands contains the injected PLC signal, so the amplitude envelope can be useful.

b) Root-Mean-Square Energy

An energy of a digital signal is defined as the sum of the square of its magnitude [43]. Root-Mean-Square Energy (RMSE) is the square root of the mean energy of a signal, as in Equation 2 [44] where x_i is the ith sample of signal x and N is the total number of samples. This is particularly useful for signals that are "balanced" or have zero mean in the time domain.

$$RMSE = \sqrt{\frac{(x_1^2 + x_2^2 + \cdots + x_N^2)}{N}} \qquad (2)$$

For our raw (full-band) PLC data, the energy (or instantaneous energy in a frame) would again be dominated by the power signal. Hence, we frequency separated the frames as described in Section III.C.2)a) and calculated RMSE for each of the thirty narrowband signals of each frame. Similar to the amplitude envelope, we were expecting variations in the RMSE in the frequency range containing ULF-PLC signals (for example, 1501-1600Hz for the 1595Hz PLC signal) to provide useful information about the PLC signal.

c) Spectral Centroid

The spectral centroid is the measure of the center of mass of the signal's spectrum [45]. Unlike time-domain features such as amplitude envelope and RMSE, the spectral centroid is a frequency-domain feature. The spectrum of our full-band signal contains a primary peak near 60Hz and large secondary harmonic peaks at odd multiples of 60Hz because of the dominant power signal. As shown in Figure 3, whenever the OOK PLC signal is active ('On'), a peak is present at that frequency. The presence or absence of the PLC signal shifts the center of mass of the spectral representation, thereby providing a classification measure of the state of the transmitted PLC signal.

D. Dataset

The raw data, obtained from DAQ at the substation, were composed of twenty-three individual files in the WAV format. Each of these files contained three-phase, time-domain PLC data (such as shown graphically in Figure 2) with a different baseband PLC transmission frequency for each file ranging from 690Hz to 2010Hz (Appendix B). The remaining processing steps in the present work consider only data from 'Phase A.' However, in subsequent research, the comparison between phases after injection of a communication signal on a single phase can also produce some useful insight [16].

The bit duration or baud rate of the injected OOK signal, i.e., the original length of the individual on or off state representing a '0' or '1' bit, in the experimental dataset was about 2 seconds (16,000 samples) when interpreted via conventional "Return-to-Zero (RZ)" line encoding [46]. To enhance the dataset for ML/NN training, we interpreted the transmissions as "Non-Return-to-Zero (NRZ)" with 1,000x repeats per bit, thus providing a substantially larger training set.

Features of the oversampled dataset such as amplitude envelope, RMS energy, and spectral centroid were extracted and compiled into 'Dataset 1' (see Figure 7), a 2D dataset with 61 feature columns - 30 each for amplitude envelope and RMS energy, and 1 for spectral centroid - and 18,400 rows (800 per file*23 files). Similarly, 2D magnitude spectrogram and 3D rectangular spectrograms (Appendix C) were constructed for the signals and combined into 'Dataset 2' and 'Dataset 3' (see Figure 7). Even though, as seen in Figure 2, the unprocessed, full-band time-series data does not seem to provide any indication of the transmitted PLC signal due to the strong out-of-band interference of the power signal, we compiled this raw data into 'Dataset 4' to be certain that the ML/NN evaluation came to the same conclusion.
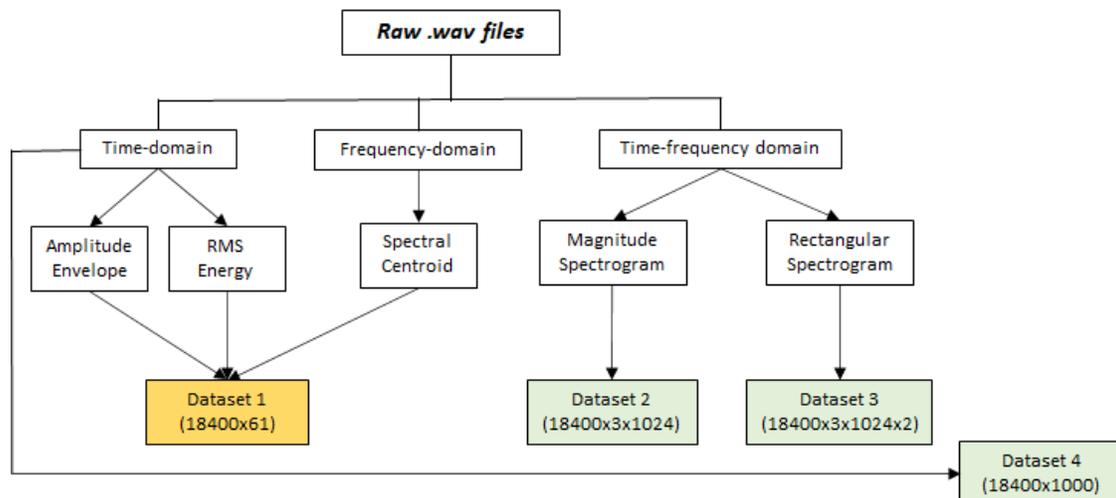
Figure 7.   Flow chart showing the summary of steps taken to convert the raw data to ML/NN ready datasets.

The flow chart in Figure 7 shows the summary of these steps (see Appendix E for a more detailed flow chart). In this figure, the orange highlights, i.e., 'Dataset 1,' comprises feature-based data while the light green highlights, i.e., 'Dataset 2,' 'Dataset 3' and 'Dataset 4,' comprise featureless data. The summary of these datasets is shown in Table I. The difference between our use of feature-based versus featureless datasets is described in section IV.

TABLE I.   DATASETS AND THEIR BRIEF DESCRIPTION

Dataset Name	Type	Description
Dataset 1	Feature-based	Uses amplitude envelope, RMS energy, and spectral centroid
Dataset 2	Featureless	2D magnitude spectrogram
Dataset 3	Featureless	3D rectangular spectrogram
Dataset 4	Featureless	1D time-series data

IV.   MACHINE LEARNING AND NEURAL NETWORK SETUP

ML/NN can be categorized based on various criteria, such as forms of learning (supervised, unsupervised, reinforcement), types of tasks (classification or prediction), learning techniques (batch or online), etc. [47][48]. Although not a traditional way of classifying ML/NN, a distinction based on data processing requirements can be important because of its implication in realistic applications. Under this criterion, ML/NN can be broadly divided into two categories: feature-based learning and featureless learning. In the current work, we implement both feature-based and featureless ML/NN on our PLC data, comparing and contrasting the performance of these two methods based on various metrics. In this section, the two methods are discussed along with a brief description of their experimental setup, their expected performance, and their relation to our datasets.

A.   Feature-based ML/NN

In feature-based learning, useful features are extracted from the raw data during pre-processing. These features are then compiled into a dataset before being fed into the ML/NN algorithms. The advantage of this approach is that large datasets with lots of noise and irrelevant features can be transformed and concentrated into smaller datasets, reducing computing time and complexity. However, the ML performance largely depends on the quality and consistency of the extracted features. The study of the data and the knowledge of the domain is vitally important in this type of ML, and thus, substantial effort in ML applications is spent on these data study and feature planning stages.

For the current work, 'Dataset 1' was used for feature-based ML/NN, as indicated in Figure 7 and Table 1, with a 70:30 training:testing configuration. To form the models with various supervised algorithms, Python Sci-kit learn [49] was used for Logistic Regression (LR) [50], Support Vector Machines (SVM) [51], and Decision Tree (TREE) [52]. The hyperparameters for these algorithms were optimized using the grid search [53] method. Accuracy, precision [54], recall [55], and f1 scores [56] were computed to evaluate and compare these various models' training and testing performance. Learning curves [57] were plotted and evaluated to ensure the models were not overfitting or underfitting. Confusion matrices [58] were also plotted to visualize the accurate label versus the predicted label.

In general, LR does well on linear classification. Our OOK PLC signals are linearly separable. However, the background power line channel, with the power signal and its coherent harmonics, is highly non-linear and introduces coherent distortion. Hence, if the featureless or unprocessed datasets were used on LR, we expected this algorithm to struggle. The feature extraction process we have utilized filters out these background non-linearities to some extent. Therefore, LR should still be able to do reasonably well with feature dataset. SVM, on the other hand, has both linear and non-linear kernels and TREE is a non-linear classifier. Therefore, we expected both of these algorithms to perform well using our feature-based dataset.

Besides these basic "one neuron" ML models, a multi-neuron, multilayer Neural Network (NN) or Artificial Neural Network (ANN) model was also tested using Python's Tensor Flow [59] and Keras [60]. The various hyperparameters of these NN models were optimized by manual trial and error (see Appendix D). Accuracy scores, loss and validation curves, and confusion matrices were generated to evaluate this model's performance, which was compared with the other ML models. Since ANN is also a non-linear classifier, we expected this method to provide high classification accuracy. For a low dimensional, low complexity feature dataset such as our PLC data, ANN may be overly complex and could potentially lead to overfitting. However, ANNs are more versatile than LR, SVM, and TREE because of their multi-neuronal, multi-layer structure, and thus, could also be used in ULF-PLC data, which has more complex signal characteristics.

B.   Featureless ML/NN

In contrast with feature-based learning, featureless learning does not require explicit feature extraction during data pre-processing. Although the ML or NN algorithms still require features, these features are extracted from the raw data by the algorithms themselves thereby removing the burden from the operator. This translates to a more generalizable, repeatable, and uniform ML analysis. This type of ML is more scalable since human intervention is less important when scaling to datasets of different properties. Unfortunately, featureless learning also requires more computational power, and the ML is more prone to noise.

ANN was used for our featureless datasets because of its simple architecture, ease of implementation, scalability, and most importantly, its ability to process multi-dimensional datasets. The multi-neuron nodes and fully connected layers of ANN can process complex information, such as the 2D and 3D spectrogram datasets, where the traditional ML algorithms fail. Therefore, they are an ideal candidate to evaluate our featureless ULF-PLC datasets.

The hyperparameters of the ANN including the number of hidden layers and the number of neurons in each layer were kept the same as for feature-based ANN (Appendix D). Accuracy scores, loss and validation curves, and

confusion matrix were generated for the optimized models for each dataset to evaluate the performance of the models and compare them across datasets.

Within our featureless datasets, we expected the time-series dataset to fail for the reasons presented in Section III.B., and for the magnitude and rectangular spectrogram datasets to produce better outcomes due to the tradeoff between the quality and the quantity of information they contain. For example, the magnitude spectrogram captures the On/Off amplitude information of OOK signals more clearly and compactly. However, as described in Section III.C.1)b)ii), the rectangular spectrogram contains more information about the transmitted PLC signal. Therefore, the accuracy and other performance measures of these two datasets in a NN model depend on how well the NN fits the respective data. With enough optimization, we expected both these datasets to perform similarly.

V. RESULTS AND DISCUSSIONS

The results of our research are divided into three main sections based on the type of dataset used in ML/NN. Section V.A. covers the results generated from feature-based dataset, Section V.B. covers the results from the use of featureless dataset, and Section V.C. compares the feature-based and featureless datasets.

The experimentation using both feature-based and featureless ML/NN is further divided into three cases, as listed in Table 2, each representing a unique real-life scenario. These scenarios arise because of the nature of the power signals. The frequency of the fundamental power signal in the PLC channel is not constant but oscillates close to 60Hz. This power signal produces strong odd harmonics at 120Hz spacings, which also oscillate corresponding to the fundamental. These dominant signals, therefore, occupy a wide range of variable frequencies in the ULF-PLC spectrum. This problem is exacerbated by additional noise of the channel, which can present time-variant interference to the ULF-PLC signals. Therefore, locating clean frequency bands for transmitting a signal in the ULF spectrum can be challenging.

Case 1, as listed in Table 2, simulates the scenario in which the frequency band to send the PLC signal is known and consistent. This would require a great deal of knowledge of the channel, so that the PLC frequency band does not overlap with the pre-existing harmonics (or the noise), which are also time-varying. Therefore, this is an unlikely scenario; however, if implemented correctly, is the easiest for the receiver to process.

TABLE II.   THREE EXPERIMENTAL CASES SIMULATING THREE UNIQUE FREQUENCY SCENARIOS IN ULF-PLC

Case	Description
Case 1	Baseband ULF-PLC signal frequency is always the same, and thus known
Case 2	Baseband ULF-PLC signal frequency changes within a set of known values
Case 3	Baseband ULF-PLC signal frequency changes within a range

Case 2 is similar to Case 1 in the sense that it simulates a scenario where the PLC frequency is known. However, in this case this frequency is not constant, but changes frequently to some known set of values (e.g., frequency hopping [61][62]). This case is built on the assumption that there is a technology present on the transmitter side that facilitates frequency hopping by continuously scanning the power line spectra and identifying an "open" channel from a set of allocated channels. This case is more realistic than Case 1 because it accounts for the variable nature of the channel and the ability to dynamically select open subsections of the low-frequency spectrum. However, this approach still relies on the assumption that the at least some of the frequency bands or sub-channels in the pre-defined set is always optimal for transmission. As a result, effective implementation of a frequency-hopped ULF-PLC system would depend in large part on dynamic spectral analysis of the channel.

Finally, Case 3 simulates a scenario in which the PLC frequency band is unknown as well as dynamic. Like Case 2, this case is built on the assumption that there is system that dynamically identifies an open channel. However, unlike Case 2, the open channel frequency does not need to be from a pre-determined set (but will be within the ULF range). This case also has interesting application in the observation or extraction of unknown events or the presence of anomalous signals in the ULF-PLC bands.

For both types of datasets and all three cases, the performance of the models is evaluated primarily using the metric of training accuracy, which essentially reflects the system's BER, an important metric in digital communication. BER is the ratio of error bits to the total number of transmitted bits. Thus, BER is the "unit complement" of accuracy, i.e., BER + accuracy =100%. Therefore, higher accuracy translates to lower BER, which in turn means that communication is more efficient.

A. Feature-based Dataset

As described in Section III.D. and illustrated in Figure 7, we constructed the feature-based dataset ('Dataset 1') by combining the amplitude envelope, RMS energy, and spectral centroid features.

1) Case 1- Known and Static Frequency

For Case 1, the PLC frequency is known. To process, we divided 'Dataset 1' into individual frequency subsets, then trained and tested ML/NN within each subset. The data in all these frequency subsets were standardized (separately), the hyperparameters of LR, SVM, and TREE algorithms were optimized, and then fitted with the corresponding models.

The training to test ratio for each experiment was set to 70:30. Since, each frequency subset was trained and tested separately, there were 560 training samples and 240 test samples for each Case 1 experiment. After fitting the models with these training data, the models were tested with test data. The process was repeated ten times (with randomized training test split for each iteration) to explore the variance of the performance of these models. Figure 8 shows the accuracies (mean accuracy curves and the 95%

confidence interval from ten iterations) of the four models from these test data at various frequencies. The inset graph within Figure 8 shows the frequencies, from 930Hz to 1650Hz, where the model test accuracies were consistently over 90%.

As seen in Figure 8, the lower and higher end of the experimental spectrum both have low test accuracies across all models. This is due to the limits of the coupling filter as well as the bandpass filtering nature of the channel for the PLC signals. At lower range of frequencies (below 870Hz in the graph), the harmonics of the fundamental power signal is very strong, and thus, the PLC signal is severely distorted. As frequency increases, these harmonics die off, opening dynamic subchannels into which the PLC signals can be introduced. However, beyond a certain frequency (after 1710Hz in the graph), the PLC signals are heavily attenuated by the transformers and other grid components, including the coupling filter, causing poor PLC output at the receiver. Therefore, there is a frequency window where the PLC signal transmission is optimal. Figure 8 shows that, in our case, this window is between 930Hz to 1650Hz. Within this window, the test accuracies are fairly high (mostly above 95%) and stable for all models, as shown by the inset graph in Figure 8.

In summary, this Case 1 experimentation shows that with sufficient knowledge of the channel and with a set of known PLC frequency, upwards of 95% accuracies can be obtained using feature-based ML/NN. High accuracy of the models means low BER, which translates to high efficiency in communication.

2) Case 2- Known but Dynamic Frequency

For Case 2, we took 'Dataset 1' as a whole for training and testing, instead of separating it with respect to frequency as described for Case 1. Hence, 'Dataset 1' was randomly split into training and test set, and the ML/NN models were fitted with the training set and evaluated with the test set. The dataset is more complex in this case for ML/NN to generalize because the PLC frequency is not the same throughout, and thus, the weight for the feature column keeps changing during learning process. This results in slower convergence than Case 1, as shown in Figure 9.
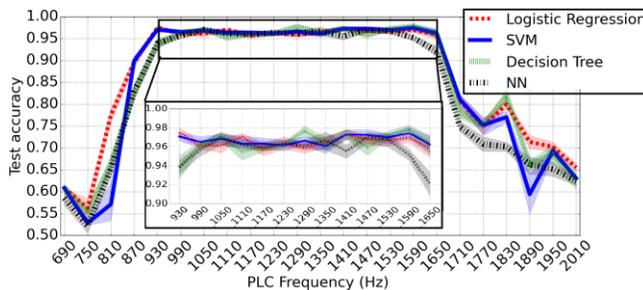
Figure 8. Graph showing test accuracy curves of Logistic Regression, SVM, Decision Tree and Neural Network models fitted with various ULF-PLC signal frequency dataset (subsets of Dataset 1). The shadowed region of the curve represents the 95% confidence interval of the accuracies. The inset graph shows a zoomed-in version of the curves for frequencies from 930Hz to 1650Hz.

(a) Case 1 at frequency of 1290Hz

(b) Case 2 with full 'Dataset 1'

(c) Case 2 with trimmed 'Dataset 1'
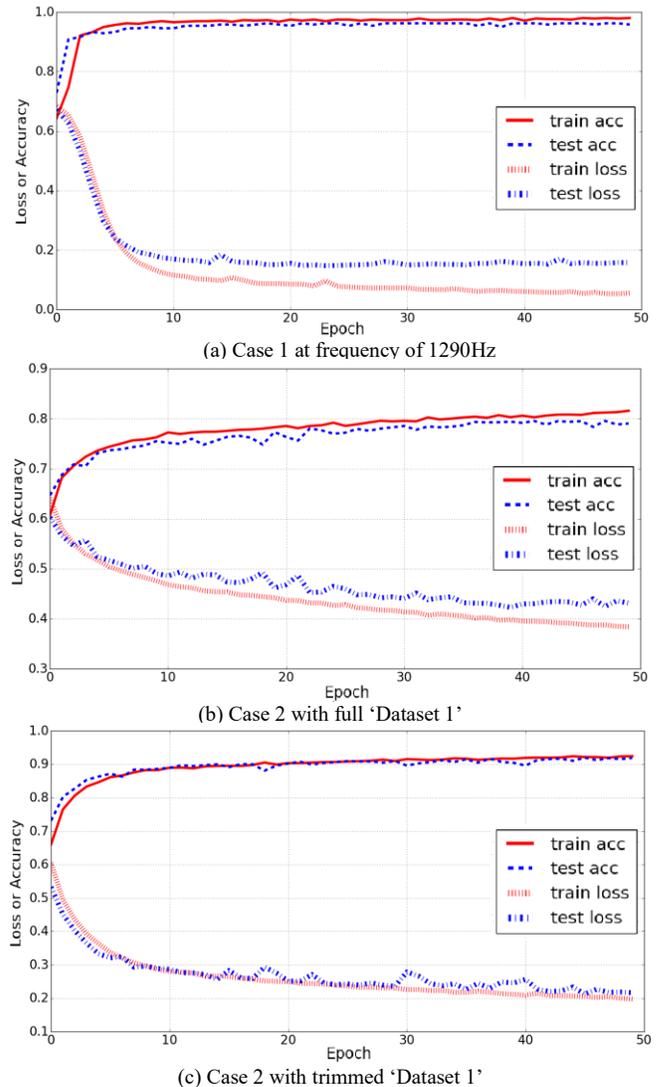
Figure 9. Graphs showing training and test (or validation) accuracy and loss curves of identical NN for (a) Case 1 with frequency of 1290Hz, (b) Case 2 with full dataset, and (c) Case 2 with trimmed dataset. The trimmed dataset for (c) is from 870Hz to 1710Hz.

As seen in Figure 9(a), the accuracy and loss curves of both training and test sets converge within the first few epochs for Case 1. However, for Case 2, the convergence takes more epochs. Figure 9(b) is for the Case 2 dataset with full range of frequencies, i.e., the complete 'Dataset 1,' which shows that the loss and accuracy curves do not stabilize within 50 epochs. In addition, the accuracies are much lower at the end of the training than in Case 1 (approximately 80% compared to around 95% for Case 1). This is because of the presence of 690Hz-810Hz and 1770Hz-2010Hz frequency data within 'Dataset 1,' which contains distorted PLC signals as shown in Figure 8. The data corresponding to these frequencies dilute the dataset, thereby causing low accuracies in Case 2. Therefore, for a fairer comparison, we took out these frequency data from the dataset for further Case 2 analysis, resulting in a "trimmed" dataset for Case 2. Figure 9(c) shows the

accuracy and loss curves of this trimmed 'Dataset 1.' In addition to the increased accuracies, trimming the dataset resulted in enhanced loss and accuracy curves.

After trimming 'Dataset 1' to include only the frequency data from 870Hz-1710Hz (i.e., the passband of our PLC channel), next we fitted the four algorithms with this updated dataset. The training and test accuracies, precision, recall and F1 scores of these models are listed in Table 3.

As seen in Table 3, SVM had the highest scores in all performance metrics for this trimmed 'Dataset 1': Case 2. TREE and NN displayed similar performance, while LR outcomes were substantially worse with approximately 10% difference in every metric compared to SVM. Further, observing the recall values, the NN model is competitive with the SVM due to the recall capabilities to classify true labels very well and has a minimal number of samples that were "false negatives."

To analyze the possible causes of this discrepancy between the various algorithms, we constructed a 2D plot with decision regions for each of these models. The two feature columns for these 2D plots were selected using backwards Sequential Feature Selection (SFS) [63] to give two most influential feature columns for each model. Figure 10 shows these plots for LR, SVM and TREE models. The axis names in these plots are in the format *'Feature abbreviation'* followed by *'Frequency sub-band.'* RMSE indicates 'RMS Energy' feature while APEV indicates 'Amplitude Envelope.'

As seen in Figure 10, LR displays a linear and continuous boundary, while the decision boundary of SVM is continuous and curved, and the decision boundaries of TREE are straight but discontinuous. These decision boundaries separate the two OOK signal states (i.e., On and Off or 1 and 0); therefore, the more flexible these boundaries are, the better the accuracy of the models' classification. The linearity of LR can sometimes prevent it from fitting some training data, whereas the SVM and TREE are more flexible, and thus, can have higher classification accuracy. However, SVM and TREE are more prone to overfitting for the same reason. In our current case, i.e., Case 2, there is no overfitting (or underfitting) on any of our models as shown by the small difference between training and testing accuracies in Table 3.

TABLE III.   PERFORMANCE OF LR, SVM, TREE AND NN WITH TRIMMED 'DATASET 1.' THE VALUES SHOWN IN THE TABLE ARE THE MEAN +/- STANDARD DEVIATION FROM TEN ITERATIONS

Algorithms	Training accuracy	Testing accuracy	Precision	Recall	F1 score
LR	0.81790 +/- 0.00328	0.81870 +/- 0.00398	0.84334 +/- 0.00522	0.80534 +/- 0.00808	0.82387 +/- 0.00429
SVM	0.92355 +/- 0.00293	0.90826 +/- 0.00388	0.93420 +/- 0.00595	.88838 +/- 0.00473	0.91070 +/- 0.00370
TREE	0.90768 +/- 0.00478	0.86604 +/- 0.00980	0.91220 +/- 0.02532	.82626 +/- 0.02008	0.86661 +/- 0.00889
NN	0.90571 +/- 0.00313	0.90078 +/- 0.00409	0.91201 +/- 0.01740	0.89916 +/- 0.02417	0.90507 +/- 0.00517

(a)   LR
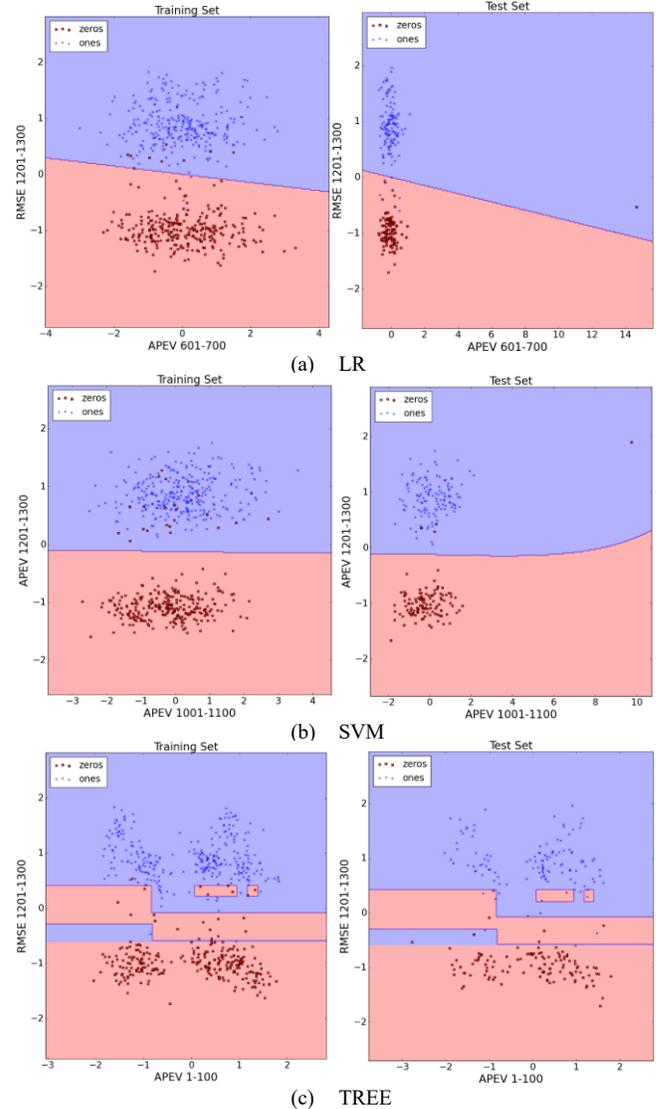
(b)   SVM

(c)   TREE

Figure 10. Plots showing training and test data (standardized), plotted over two optimal feature columns, along with the decision boundary and shaded target regions of (a) LR (b) SVM and (c) TREE models. The axis names are in the format 'Feature abbreviation' followed by 'Frequency sub-band.'

3) Case 3- Unknown and Dynamic Frequency

In Case 3, the ML/NN is trained with a range of frequencies, and the test is done with some frequency within or close to this range, but not present in the training set. Therefore, the only prior information needed in this scenario for real-life implementation is the general frequency range containing the PLC signal and a training set, which incorporates this same range, or set of potential channels/subchannels. The expectation is that the ML/NN will train within this range and be able to generalize well enough to classify signals, which lie in frequencies or subchannels that were not specifically trained for recognition. To simulate this scenario, we separated 'Dataset 1' into individual frequency subsets or subchannels, used all subchannels except one subset for training, and tested on the unused subchannel. Our objective with this case was to test

how well ML/NN models can generalize on data with untrained PLC frequency and to observe what effect the frequency has in the case of ambiguous subchannel isolation. As with Case 2, including the complete 'Dataset 1' for this case would dilute the dataset and cause low training accuracy. In practical use cases, the frequency bounds or range of subchannels can be specified. Therefore, for Case 3, we used trimmed 'Dataset 1' as in Case 2.

Figure 11 shows the test accuracies of this experiment for the four algorithms (LR, SVM, DT, and NN). In the figure, frequencies shown on the X-axis indicate the subchannel that was not contained in the training set. As a result, the results shown in Figure 11 indicate the accuracy of the ML/NN approach for transmissions where that specific sub-band was used in testing but was not used in training.

As seen in Figure 11, all of the algorithms had some success in generalizing to the test data with untrained PLC sub-bands. However, the performance is not consistent for any algorithms across the whole spectrum. The low accuracies at the extremes, i.e., 870Hz and 1710Hz can be explained by the fact that the models are trained using sub-bands higher **or** lower from the testing sub-band, but not both higher **and** lower. Therefore, the model only has the upper or lower set of sub-bands to generalize the test set. This causes low accuracy. Extending this similarity or bounding measure argument, we expect the sub-bands near the middle of the range to have the highest accuracy, decreasing to either side. The curves in Figure 11 promise a similar trend; however, this general trend seems to be contradicted by two big dips around sub-bands centered at 1050Hz and 1350Hz. Upon closer examination of the models at these two sub-bands, we discovered that the inconsistencies are due to the lack of robustness of the models themselves. The architecture and hyperparameters of each model were held constant for all sub-bands and all test scenarios, i.e., the models are not optimized for specific training set. This approach simulates the real-life implementation of this case, where the models are trained across a broad set of subchannels and cannot be optimized for the unknown subchannel. Therefore, our models were not optimized for certain test sets, and they failed at those frequencies. Despite this setback, the models showed high accuracy in the majority of tested sub-bands, and thus showed promising generalization behavior. Further optimization of these models or use of more complex models may be useful in rectifying the inconsistencies.
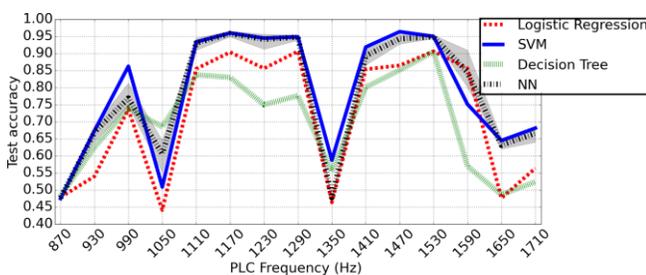
B. Featureless Dataset

In addition to using a feature-based dataset in ML/NN, we also investigated the use of featureless datasets. We did so by constructing 1D time-series, 2D magnitude spectrogram and 3D rectangular spectrogram datasets, as discussed in Section III.D. NNs were used on each dataset because of their ability to accept multi-dimensional input data. The multi-neuronal architecture of NNs is also more complex than one-neuron structures of LR, SVM or TREE, which could aid in parsing through our noisy featureless datasets. The structure and hyperparameters of the NN for the time-series and magnitude spectrogram dataset were identical to the NN used for the feature-based dataset (Appendix D). For the rectangular spectrogram however, an additional convolution layer and a max pool layer were added (as part of manual optimization) before the fully connected NN layers. The complete architecture of this convolutional neural network (CNN) is shown in Figure 12.

As seen in Figure 12, each sample of the rectangular spectrogram is 3x1024x2, where 2 is the number of channels (i.e., stacked real spectrogram and imaginary spectrograms). The first operation is convolution with 64 3x3 filters producing 1x1022x64 feature maps that are fed into the max pool layer with pool size 1x3 and produces output of 1x340x64. This output is flattened to a 1D array of size of 21,760 samples and fed into the fully connected NN for training and testing. This NN is identical to the ones used in the feature-based dataset (Appendix D). Subsequent sections describe the comparative analysis of these three featureless datasets.

1) Case 1- Known and Static Frequency

In Case 1, the NN/CNN is trained and tested within a particular frequency subset or set of subchannels/sub-bands. The test accuracy results of the NN/CNN models are presented in Figure 13.
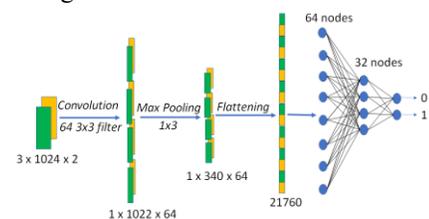
Figure 12. CNN architecture showing the various layers of operation on our rectangular spectrogram dataset.
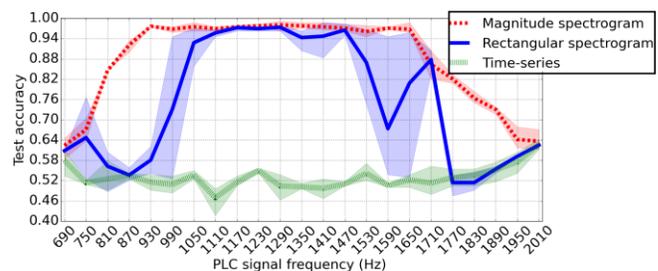
Figure 13. Graph showing test accuracy curves (with 95% confidence interval) of NN models fitted with magnitude spectrogram (Dataset 2), rectangular spectrogram (Dataset 3) and time-series (Dataset 4) datasets containing ULF-PLC signals of various frequencies.

Figure 11. Test accuracies for Logistic Regression, SVM, Decision Tree and Neural Network for Dataset 2: Case 3. The PLC frequency on the x-axis represents the test frequency.

The time-series dataset exhibited poor performance, as expected due to the strong out of band interference (i.e., the power signal). Between the two spectrogram formats, the magnitude spectrogram models exhibited better overall performance, and particularly better accuracy than the rectangular spectrogram models, especially in the lower and the higher end of the spectrum. The magnitude spectrogram models also had a much narrower and more consistent confidence interval than the rectangular spectrogram. The confidence interval result is especially interesting since a larger confidence interval implies larger variance, and hence, the rectangular spectrogram dataset seems to have higher variance in its accuracy results. Higher variance suggests more unpredictability and less repeatability in the performance, which is undesirable.

This discrepancy in performance between magnitude spectrogram and rectangular spectrogram formats can be explained by the type and quantity of information each dataset contains. The magnitude spectrogram, by definition, contains only the magnitude or energy information of the signal, which is directly related to the signal amplitude. So, for OOK signals, the magnitude spectrogram more clearly represents modulation transitions, thus simplifying the task for the NN. In contrast, the rectangular spectrogram holds more information about the signal, including phase data, which could be advantageous in some use cases such as phase-based modulation (e.g., Phase Shift Keying) or a combination of phase and amplitude modulation (e.g., Quadrature Amplitude Modulation). However, in the approaches discussed here, the larger size of the rectangular dataset causes data dilution resulting in lower performance compared to magnitude spectrogram models.

2) Case 2- Known but Dynamic Frequency

As in Case 2 of the feature-based ML/NN analysis {Section V.A.2)}, the NN/CNN is trained with the trimmed featureless datasets (without separating frequency subsets). The objective in this approach is to explore which featureless datasets perform the best when trained with samples containing PLC signals of varying frequencies. Table 5 summarizes the result of this experiment (with ten iterations of each model).

As shown in Table 5, the magnitude spectrogram models outperform other models. The time-series models fail again due to the nature of the unprocessed time-series PLC data and the presence of the strong out-of-band interferer. Interestingly, the recall score of the time-series dataset is very close to 1. This does not necessarily mean that the time-series model performed well. Figure 14 shows the confusion matrix of one of these time-series models, which explains why the recall score was high even though the accuracy was low. As shown in this figure, the time-series model predicted all samples as '1' (On), which explains the low accuracy. The recall is calculated as shown in Equation 3. The true positives are the accurately classified positives (top left quadrant in Figure 14) while the false negatives are misclassified negatives (bottom left quadrant). In our case, both values are zero, giving 0÷0, which causes error, but is overwritten as 1 during processing. In this way, the recall was very high.

$$Recall = \frac{Truepositives}{Truepositives + Falsenegatives} \quad (3)$$

TABLE IV.   PERFORMANCE OF MAGNITUDE SPECTROGRAM (DATASET 2), RECTANGULAR SPECTROGRAM (DATASET 3) AND TIME-SERIES (DATASET 4) NN/CNN MODELS. THE VALUES SHOWN IN THE TABLE ARE THE MEAN +/- STANDARD DEVIATION FROM TEN ITERATIONS

Dataset type	Training accuracy	Testing accuracy	Precision	Recall	F1 score
Time-series	0.52595 +/- 0.00000	0.52583 +/- 0.00000	0.52583 +/- 0.00000	1.00000 +/- 0.00000	0.68924 +/- 0.00000
Magnitude spectrogram	1.00000 +/- 0.00000	0.95503 +/- 0.00326	0.96010 +/- 0.00413	0.95415 +/- 0.00512	0.95710 +/- 0.00314
Rectangular spectrogram	0.76138 +/- 0.15679	0.73655 +/- 0.14313	0.74463 +/- 0.14841	0.88257 +/- 0.08874	0.79102 +/- 0.07624
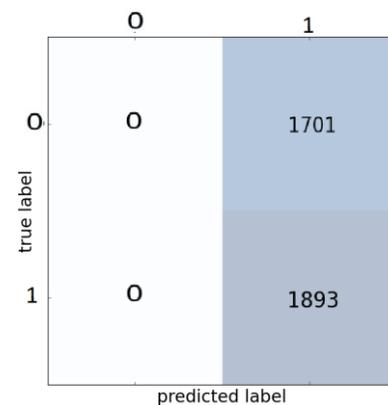
Figure 14. Confusion matrix of a featureless time-series model. This model was fitted with trimmed 'Dataset 4' for Case 2 analysis.

3) Case 3- Unknown and Dynamic Frequency

In Case 3, the NN/CNN models are trained using every sub-band within the dataset (Datasets 2, 3 and 4) except one, and tested with the one that was left out of the training process. The objective in this approach was to observe if the NN/CNN featureless models can generalize during the training process to be able to perform well on test data with untrained PLC sub-bands. The result of this experimentation is summarized in Figure 15.
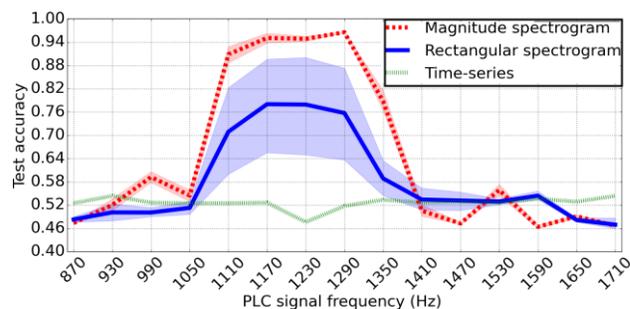
Figure 15. Test accuracy curves of NN/CNN models fitted with magnitude spectrogram, rectangular spectrogram, and time-series dataset for Case 3 analysis.

As seen in Figure 15, the time series data again failed, as expected. The magnitude and rectangular spectrogram formats, however, had some success in generalizing to a new frequency data subset. The bell-shaped accuracy curves of both rectangular and magnitude spectrograms support the "similarity measure" hypothesis formulated previously in Case 3 of feature-based datasets in Section V.A.3). Reiterating, it seems that the ML/NN models are able to generalize the best (and perform well) to a sub-band when trained on the greatest number of similar sub-bands, and in particular sub-bands on "both sides." The sub-bands near the edge of the passband have the fewest "similar" sub-bands for training whereas the sub-bands in the middle of the passband have the most, and hence, the accuracies for edge cases should be the lowest with increasing accuracy moving towards the middle. The curves for both spectrogram datasets shown in Figure 15 indeed show that the accuracies are low at the edge frequencies and greatest at mid frequencies, supporting our hypothesis. For practical implementation, this means that the training frequency range, for this type of unknown and dynamic PLC frequency scenario, should be wide enough so that the most probable PLC sub-bands or test frequencies lie somewhere in the middle. This structure would ensure good classification accuracy (or low BER), and hence, efficient communication.

C. Feature-based Vs Featureless Learning

In Sections V.A. and V.B., the results of feature-based and featureless ML/NN/CNN were discussed separately. In the present section, these two methods are compared directly based on the previous results of NN models from each category, as shown in Figure 16. NN is used to compare the two methods because an identical NN (with same hyperparameters) was used with feature-based as well as featureless datasets. For featureless method, the magnitude spectrogram is used as the dataset format of choice (i.e., Datasets 2) because of its overall best performance as shown by results in Section V.B.

Figure 16(a) reveals that the featureless NN model had slightly higher accuracies than the feature-based NN model when testing these models with a trained parameter. In this case, NN models were able to find the relevant features

using extracted features for the feature dataset and feature maps for featureless dataset. However, the information contained within the relevant feature maps is truer to the original raw signal than the lossy extracted features due to a reduction in steps required for data pre-processing. Therefore, the featureless NN models had marginally better performance than feature-based NN.

On the other hand, when the models were asked to generalize to an untrained parameter, as in Case 3 shown in Figure 16 (b), the feature-based NN models perform better overall than featureless NN. This result may be due to the fact that the NN model needed to be more complex for the featureless dataset on account of its greater size (3x1024 per sample for magnitude spectrogram dataset whereas 1x61 for the feature dataset), higher dimensionality (2D for magnitude spectrogram while 1D for feature dataset) or the noise within the dataset (higher noise level in the magnitude spectrogram dataset than the feature dataset). Another probable cause could be that the NN model overfitted the featureless dataset. Overfitting might not have been a problem with the feature-based dataset because of the limited number of features. However, for the featureless dataset, the number of features (in the form of feature maps) is much greater, and thus the weights are spread out during training, causing overfitting of the training data. This overfitting then results in poor generalization for untrained samples.

VI. CONCLUSION

In this research study, an inter-level, upstream PLC network architecture was created; using this network architecture, OOK ULF baseband PLC signals were transmitted and captured; the captured data was processed to create feature-based and featureless ML/NN datasets. These datasets were used in various ML/NN formats to classify the temporal On/Off states of the transmitted OOK ULF-PLC signal. Using this pipeline, three broad questions related to the use of ML/NN in ULF-PLC data were explored. First, ULF-PLC signal characteristics were evaluated. Second, methods for pre-processing complex-valued communications signals were explored for use in conventional ML algorithms. Finally, the efficacy of a small

(a) Frequency analysis: Case 1- known and static frequency   (b) Frequency analysis: Case 3- unknown and dynamic frequency
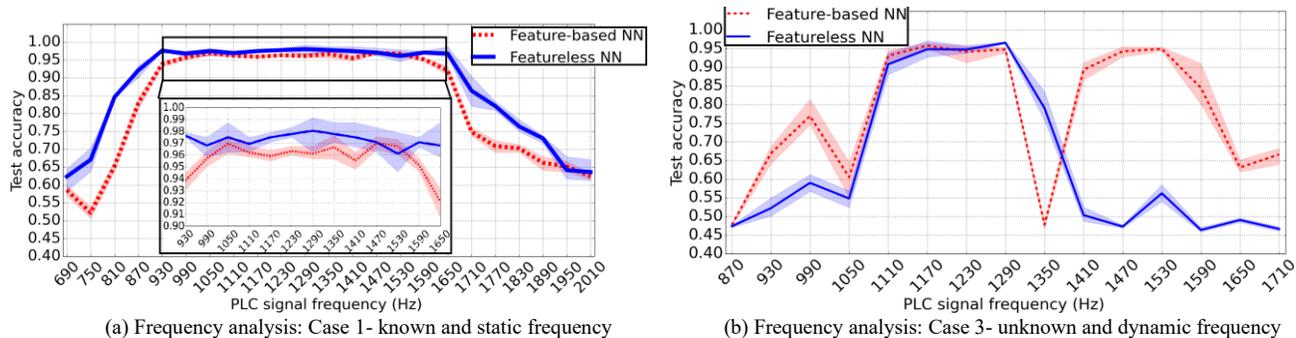
Figure 16. Graphs showing test accuracy curves (with 95% confidence interval) of Neural Network models fitted with feature dataset and magnitude spectrogram-based featureless datasets. (a) Case 1 frequency analysis using feature-based 'Dataset 1' and featureless 'Dataset 2' and (b) Case 3 frequency analysis using 'Dataset 1' and 'Dataset 2'.

collection of supervised ML techniques in decoding ULF-PLC communication data was evaluated.

A. PLC Signal Characteristics

As presented in Section V.A.1) (Figure 8) and Section V.B.1) (Figure 13), the optimal frequency range for inter-level ULF-PLC appears to be confined in a distinct passband around 930Hz-1650Hz. Below this band, the strong odd harmonics of the fundamental power signal interfere directly with the PLC signal, while above this band, grid components such as power transformers attenuate the signals substantially, causing poor signal propagation. This passband frequency range is dependent on various other factors such as the grid architecture (between the PLC signal transmitter and receiver), the coupling filter, etc. Therefore, frequency range as presented in this research may be relatively common but are not universally applicable. However, the bandpass nature of the power line channel is indeed characteristic of ULF-PLC.

B. Dataset Characteristics

One of the primary objectives of this study was to explore how the raw ULF-PLC data could be used in ML/NN and what types of data pre-processing might be most desirable or effective. To that end, we investigated two different methods: feature-based and featureless.

Between the three featureless datasets, the time-series format was the least effective because of the dominance of the power signal in this domain as an out-of-band interferer. When the same data was transformed to time/frequency representation, the ML/NN performance improved significantly. The magnitude spectrogram format performed better due to its efficient representation of the OOK ULF-PLC data. If more sophisticated frequency and/or phase shift methods are used for modulating data into the ULF-PLC channel, then the use of rectangular spectrogram formats is likely to outperform magnitude spectrogram formats, as concluded in our previous research [40]. Therefore, there are merits to all of these featureless datasets, in this ULF-PLC application and beyond. In particular the ability of the system to recognize data based on various sub-band configurations is compelling.

The outcomes from feature-based and featureless datasets were compared directly using an identical NN. The results of this comparison are discussed in Section V.C. and shown in Figure 16. Interestingly, the featureless magnitude spectrogram models had higher test accuracy than the feature-based models when the test data had the same PLC signal parameter (frequency) as the training data. However, when the test data had completely different signal parameters, the feature-based models were more consistent in providing high classification accuracies. As a result, we conclude that NN models overfit with the larger featureless dataset, causing it to generalize poorly on the test data with untrained parameters. Therefore, feature-based methods may be more appropriate when generalization is desired.

In addition to testing accuracy, comparison of feature-based and featureless methods using metrics such as the complexity of the ML models, time of processing, scalability, etc. may be valuable outcomes. In our tests using the ULF-PLC dataset, feature-based models had similar performance to featureless models despite employing simpler algorithms. The use of simpler models along with smaller dataset of the feature-based method resulted in a requirement for substantially less computer processing power and training time. However, the feature-dataset also requires substantial data pre-processing, which requires time and knowledge of the domain, and which can omit unknown, but relevant features. In contrast, although featureless methods require more complex ML/NN models, they are also easier to scale to varying architectures and applications and require less domain-specific knowledge to optimize. Further, they are capable of detecting features and/or outcomes that may have been previously unknown. Therefore, feature-based methods are desirable when training time and computer processing power are a consideration, while featureless methods are attractive when scaling and ease of deployment are the main concerns, and when the dataset environment may not be completely known.

C. ML/NN Algorithm Characteristics

To evaluate the ULF-PLC dataset, we utilized three supervised ML algorithms: Logistic Regression (LR), Support Vector Machine (SVM) and Decision Tree (TREE). We also used Neural Network (NN) and Convolutional Neural Network (CNN) models. The comparison between LR, SVM, TREE and NN, presented in Section V.A., indicates that the classification accuracies of these models were very similar, with SVM slightly better in most cases due to its non-linear capability. However, this also meant that SVM models needed much more stringent regularization to avoid overfitting, which increases implementation complexity. In contrast, NN and CNN may be more suitable for featureless datasets because of their relatively more complex architectures and native ability to process multi-dimensional datasets. For the same reason, NN and CNN approaches are also more flexible and scalable.

APPENDIX

- *Appendix A: Github Page*

The data files and Python scripts used for in this thesis can be found at:
https://github.com/kushal-thapa/ML_for_PLC_thesis

- *Appendix B: Features of the raw ULF-PLC data*

Number of raw PLC files: 23
File type: .wav
Length of each file: 100 seconds
PLC signal amplitude: 1A
PLC frequency: 690Hz-2010Hz with 60Hz spacings
Number of channels: 3
Sampling rate: 8000 samples per second

- *Appendix C: Spectrogram Parameters*

Frame size = 500 samples
Frame overlap = 250 samples
FFT size =1024
Window = Hanning

- *Appendix D: Optimized hyperparameters of the NN for Dataset 1*

No. of hidden layers		2
Hidden layer 1	**No. of nodes**	64
	Activation function	Relu
Hidden layer 2	**No. of nodes**	32
	Activation function	Relu
Output layer	**No. of nodes**	2
	Activation function	Softmax
Learning rate		0.001
Optimizer		Adam
Loss		Sparse categorical crossentropy
Epochs		50
Batch size		16

- *Appendix E: Experimentation steps*
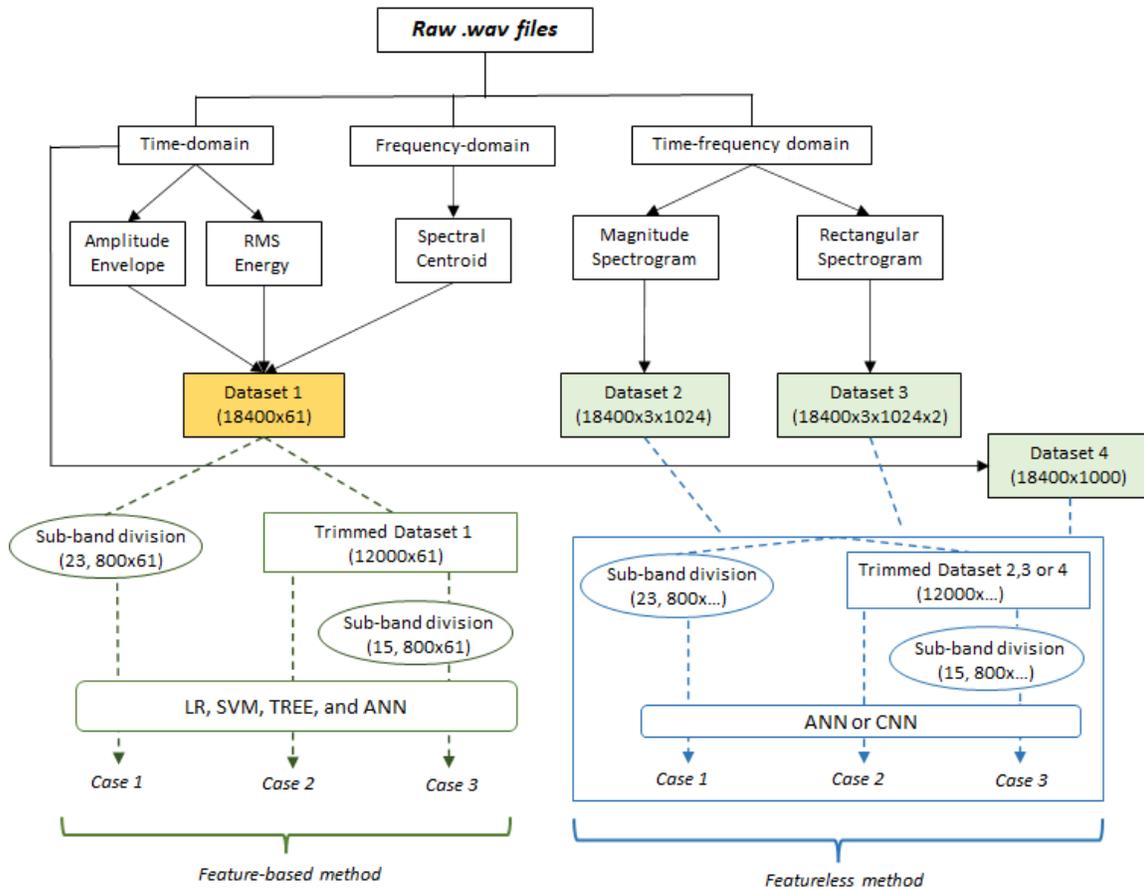
Figure Appendix 1: Extension of Figure 7 showing the creation of datasets and the subsequent use of those datasets in various cases in feature-based and featureless methods.

REFERENCES

[1] K. Thapa, S. McClellan, and D. Valles, "Supervised Machine Learning in Digital Power Line Communications," in *The Sixteenth International Conference on Digital Telecommunications*, Porto, Apr. 2021, pp. 16–21. [Online]. Available: https://www.thinkmind.org/index.php?view=article&articleid=icdt_2021_1_40_18004 (accessed Nov. 28, 2021).

[2] K. Thapa, "Digital Signal Processing and Machine Learning Applied to Power Line Communications," Unpublished, Texas State University, San Marcos, Texas, 2021. [Online]. Available: https://digital.library.txstate.edu/handle/10877/14044 (accessed Nov. 28, 2021).

[3] K. Dostert, "Telecommunications over the power distribution grid–possibilities and limitations," *IIR-Powerline*, vol. 6, no. 97, 1997.

[4] M. Schwartz, "Carrier-wave telephony over power lines: Early history [History of Communications]," *IEEE Commun. Mag. Commun. Mag. IEEE IEEE Commun Mag*, vol. 47, no. 1, pp. 14–18, Jan. 2009, doi: 10.1109/MCOM.2009.4752669.

[5] K. Sharma and L. M. Saini, "Power-line communications for smart grid: Progress, challenges, opportunities and status," *Renew. Sustain. Energy Rev.*, vol. 67, pp. 704–751, Jan. 2017.

[6] S. U. Ercan, O. Ozgonenel, Y. E. Haj, C. Christopoulos, and D. W. P. Thomas, "Power line communication design and implementation over distribution transformers," in *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, Nov. 2017, pp. 190–194.

[7] A. K. Schurig, "Universal lan power line carrier repeater system and method," US5818821A, Oct. 06, 1998. [Online]. Available: https://patents.google.com/patent/US5818821A/en (accessed Nov. 28, 2021).

[8] R. Du and Q. Xiao, "Baby Monitor," US20090212926A1, Aug. 27, 2009. [Online]. Available: https://patents.google.com/patent/US20090212926A1/en (accessed Nov. 28, 2021).

[9] P. An, X. Chen, J. Wang, S. Zhou, and X. Shan, "Analysis on Application of Repeater Technology in Powerline Communications Networks," in *2007 IEEE International Symposium on Power Line Communications and Its Applications*, Mar. 2007, pp. 273–277. doi: 10.1109/ISPLC.2007.371136.

[10] K. Bode, "BPL Takes Posthumous Beating With IBEC Closure - ISP Blames Tornadoes, Though BPL Failures Played a Role," *DSL Reports*. http://www.dslreports.com/shownews/BPL-Takes-Posthumous-Beating-With-IBEC-Closure-117677 (accessed Nov. 28, 2021).

[11] J. Tanner, "You can't kill BPL if it's already dead | Telecom Ramblings," Jan. 13, 2012. https://www.telecomramblings.com/2012/01/you-cant-kill-bpl-if-its-already-dead/ (accessed Nov. 28, 2021).

[12] D. Dzung, I. Berganza, and A. Sendin, "Evolution of powerline communications for smart distribution: From ripple control to OFDM," in *2011 IEEE International Symposium on Power Line Communications and Its Applications*, Apr. 2011, pp. 474–478. doi: 10.1109/ISPLC.2011.5764444.

[13] "Intelligent Platform Management Interface Specification Second Generation." Intel Corporation, Hewlett-Packard Company, NEC Corporation, Dell Inc., Oct. 01, 2013. [Online]. Available: https://www.intel.com/content/www/us/en/products/docs/servers/ipmi/ipmi-second-gen-interface-spec-v2-rev1-1.html (accessed Nov. 28, 2021).

[14] "Spacecraft Onboard Interface Systems–Low Data-Rate Wireless Communications for Spacecraft Monitoring and Control." Council of the Consultative COmmittee for Space Data Systems (CCSDS), May 2013. [Online]. Available: https://public.ccsds.org/Pubs/882x0m1.pdf (accessed Nov. 28, 2021).

[15] D. Cooper and T. Jeans, "Narrowband, low data rate communications on the low-voltage mains in the CENELEC frequencies. I. Noise and attenuation," *IEEE Trans. Power Deliv. Power Deliv. IEEE Trans. IEEE Trans Power Deliv.*, vol. 17, no. 3, pp. 718–723, Jul. 2002, doi: 10.1109/TPWRD.2002.1022794.

[16] S. McClellan, D. Valles, and G. Koutitas, "Dynamic Voltage Optimization Based on In-Band Sensors and Machine Learning," *Appl. Sci.-BASEL*, vol. 9, no. 14, Jul. 2019, doi: 10.3390/app9142902.

[17] A. Narayan, "California's electric grid needs an overhaul.," *Bus. J. Serv. Fresno Cent. San Joaquin Val.*, no. 325799, pp. 18–18, 1/1/2021 2021.

[18] S. Littlechild and L. Kiesling, "Hayek and the Texas blackout," *Electr. J.*, vol. 34, no. 6, p. 106969, Jul. 2021, doi: 10.1016/j.tej.2021.106969.

[19] S. McClellan, "Suppression of interference in power and communications signals," US9876516B2, Jan. 23, 2018. [Online]. Available: https://patents.google.com/patent/US9876516B2/en (accessed Nov. 28, 2021).

[20] J. J. LoPorto and S. McClellan, "Intelligent power system and methods for its application," US8519832B2, Aug. 27, 2013. [Online]. Available: https://patents.google.com/patent/US8519832B2/en (accessed Nov. 28, 2021).

[21] J. Arrillaga and N. R. Watson, *Power system harmonics*. John Wiley & Sons, 2004.

[22] J. Niitsoo, I. Palu, J. Kilter, P. Taklaja, and T. Vaimann, "Residential load harmonics in distribution grid," in *2013 3rd International Conference on Electric Power and Energy Conversion Systems*, Oct. 2013, pp. 1–6. doi: 10.1109/EPECS.2013.6713054.

[23] A. M. Tonello, N. A. Letizia, D. Righini, and F. Marcuzzi, "Machine Learning Tips and Tricks for Power Line Communications," *IEEE Access*, vol. 7, pp. 82434–82452, 2019, doi: 10.1109/ACCESS.2019.2923321.

[24] M. Korki, N. Hosseinzadeh, H. L. Vu, T. Moazzeni, and C. H. Foh, "A channel model for power line communication in the smart grid," in *2011 IEEE/PES Power Systems Conference and Exposition*, Mar. 2011, pp. 1–7. doi: 10.1109/PSCE.2011.5772561.

[25] B. Varadarajan, Il Han Kim, A. Dabak, D. Rieken, and G. Gregg, "Empirical measurements of the low-frequency power-line communications channel in rural North

America," in *2011 IEEE International Symposium on Power Line Communications and Its Applications*, Apr. 2011, pp. 463–467. doi: 10.1109/ISPLC.2011.5764442.

[26] G. Prasad, Y. Huo, L. Lampe, A. Mengi, and V. C. M. Leung, "Fault Diagnostics with Legacy Power Line Modems," in *2019 IEEE International Symposium on Power Line Communications and its Applications (ISPLC)*, Apr. 2019, pp. 1–6. doi: 10.1109/ISPLC.2019.8693385.

[27] M. Nassar, J. Lin, Y. Mortazavi, A. Dabak, I. H. Kim, and B. L. Evans, "Local Utility Power Line Communications in the 3–500 kHz Band: Channel Impairments, Noise, and Standards," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 116–127, Sep. 2012, doi: 10.1109/MSP.2012.2187038.

[28] "California Instruments CS Series - AC Current Source," *Ametek Programmable Power*. www.powerandtest.com/power/ac-power-sources/cs-series (accessed Nov. 28, 2021).

[29] "DI-1100 4-channel USB Data Acquisition Starter Kit," *DATAQ Instruments*. https://www.dataq.com/products/di-1100/ (accessed Nov. 28, 2021).

[30] S. Bochner, K. Chandrasekharan, and K. Chandrasekharan, *Fourier transforms*, no. 19. Princeton University Press, 1949.

[31] E. O. Brigham and R. Morrow, "The fast Fourier transform," *IEEE Spectr.*, vol. 4, no. 12, pp. 63–70, 1967.

[32] K. R. Rao, D. N. Kim, and J. J. Hwang, *Fast Fourier transform : algorithms and applications.* Springer, 2010.

[33] D. A. Lyon, "The discrete fourier transform, part 4: spectral leakage," *J. Object Technol.*, vol. 8, no. 7, 2009.

[34] A. Breitenbach, "Against spectral leakage," *Measurement*, vol. 25, no. 2, pp. 135–142, 1999.

[35] V. Mathews and D. Youn, "Spectral leakage suppression properties of linear and quadratic windowing," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 32, no. 5, pp. 1092–1095, 1984.

[36] T. M. Apostol, *Mathematical analysis.*, Second edition. Addison-Wesley Pub. Co., 1974.

[37] J. S. Dramsch, M. Lüthje, and A. N. Christensen, "Complex-valued neural networks for machine learning on non-stationary physical data," *Comput. Geosci.*, vol. 146, p. 104643, Jan. 2021, doi: 10.1016/j.cageo.2020.104643.

[38] H. G. Zimmermann, A. Minin, and V. Kusherbaeva, "Comparison of the complex valued and real valued neural networks trained with gradient descent and random search algorithms," 2010, pp. 213–218.

[39] J. Shima, "Weak signal processing systems and methods," 10879946, Dec. 29, 2020. [Online]. Available: http://patft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=10879,946.PN.&OS=PN/10879,946&RS=PN/10879,946 (accessed Nov. 28, 2021).

[40] K. Thapa, S. McClellan, and D. Valles, "An Evaluation of Neural Network Performance Using Complex-Valued Input Data," in *ICDT 2021 : The Sixteenth International Conference on Digital Telecommunications*, Porto, Apr. 2021, pp. 27–31. [Online]. Available: https://www.thinkmind.org/index.php?view=article&articleid=icdt_2021_1_60_18006 (accessed Nov. 28, 2021).

[41] T. Smyth, "Amplitude Envelopes," *Amplitude Envelopes*. http://musicweb.ucsd.edu/~trsmyth/sinusoids171/Amplitude_Envelopes.html (accessed Nov. 28, 2021).

[42] R. G. Lyons and D. L. Fugal, *The essential guide to digital signal processing.* Prentice Hall, 2014.

[43] B. Boashash, Ed., "Chapter 4 - Advanced Time-Frequency Signal and System Analysis," in *Time-Frequency Signal Analysis and Processing (Second Edition)*, Oxford: Academic Press, 2016, pp. 141–236. doi: 10.1016/B978-0-12-398499-9.00004-2.

[44] S. Tjoa, "Energy and RMSE," *musicinforetrieval.com*. https://musicinformationretrieval.com/energy.html (accessed Nov. 28, 2021).

[45] J. M. Grey and J. W. Gordon, "Perceptual effects of spectral modifications on musical timbres.," *J. Acoust. Soc. Am.*, vol. 63, pp. 1493–1500, May 1978.

[46] K. W. Cattermole, "Principles of digital line coding.," *Int. J. Electron.*, vol. 55, no. 1, p. 3, Jul. 1983.

[47] S. C. H. Hoi, D. Sahoo, J. Lu, and P. Zhao, "Online learning: A comprehensive survey," *Neurocomputing*, vol. 459, pp. 249–289, Oct. 2021.

[48] J. Brownlee, "14 Different Types of Learning in Machine Learning," *Machine Learning Mastery*, Nov. 10, 2019. https://machinelearningmastery.com/types-of-learning-in-machine-learning/ (accessed Nov. 28, 2021).

[49] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. 85, pp. 2825–2830, 2011.

[50] "sklearn.linear_model.LogisticRegression," *scikit learn*. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html (accessed Nov. 28, 2021).

[51] "sklearn.svm.SVC." https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html (accessed Nov. 28, 2021).

[52] "sklearn.tree.DecisionTreeClassifier." https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html (accessed Nov. 28, 2021).

[53] "sklearn.model_selection.GridSearchCV." https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html (accessed Nov. 28, 2021).

[54] "sklearn.metrics.precision_score." https://scikit-learn.org/stable/modules/generated/sklearn.metrics.precision_score.html (accessed Nov. 28, 2021).

[55] "sklearn.metrics.recall_score." https://scikit-learn.org/stable/modules/generated/sklearn.metrics.recall_score.html (accessed Nov. 28, 2021).

[56] "sklearn.metrics.f1_score." https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html (accessed Nov. 28, 2021).

[57] "Plotting Learning Curves — scikit-learn 0.24.2 documentation." https://scikit-learn.org/stable/auto_examples/model_selection/plot_learning_curve.html (accessed Nov. 28, 2021).

[58] "sklearn.metrics.confusion_matrix." https://scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion_matrix.html (accessed Nov. 28, 2021).

[59] M. Abadi *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *ArXiv Prepr. ArXiv160304467*, 2016.

[60] A. Gulli and S. Pal, *Deep learning with Keras*. Packt Publishing Ltd, 2017.

[61] K. M. Dostert, "Frequency-hopping spread-spectrum modulation for digital communications over electrical power lines," *IEEE J. Sel. Areas Commun. Sel. Areas Commun. IEEE J. IEEE J Sel. Areas Commun*, vol. 8, no. 4, pp. 700–710, May 1990, doi: 10.1109/49.54466.

[62] D. J. Torrieri, *Principles of spread-spectrum communication systems.*, Fourth edition. Springer, 2018.

[63] "sklearn.feature_selection.SequentialFeatureSelector — scikit-learn 0.24.2 documentation." https://scikit-learn.org/0.24/modules/generated/sklearn.feature_selection.SequentialFeatureSelector.html (accessed Nov. 28, 2021).