Engineering Applications of Artificial Intelligence 133 (2024) 108331

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

# Most influential feature form for supervised learning in voltage sag source localization

Younes Mohammadi [a,*], Boštjan Polajžer [b], Roberto Chouhy Leborgne [c], Davood Khodadad [a]

[a] Department of Applied Physics and Electronics, Umeå Universitet, 90187, Umeå, Sweden
[b] Faculty of Electrical Engineering and Computer Science, University of Maribor, 2000, Maribor, Slovenia
[c] Universidade Federal do Rio Grande do Sul, Osvaldo Aranha, 99, 90035-190, Porto Alegre, RS, Brazil

ARTICLE INFO

*Keywords:*
Voltage sag (dip)
Source localization
Supervised and unsupervised learning
Convolutional neural network
Time-sample-based features

ABSTRACT

The paper investigates the application of machine learning (ML) for voltage sag source localization (VSSL) in electrical power systems. To overcome feature-selection challenges for traditional ML methods and provide more meaningful sequential features for deep learning methods, the paper proposes three time-sample-based feature forms, and evaluates an existing feature form. The effectiveness of these feature forms is assessed using k-means clustering with $k = 2$ referred to as downstream and upstream classes, according to the direction of voltage sag origins. Through extensive voltage sag simulations, including noises in a regional electrical power network, k-means identifies a sequence involving the multiplication of positive-sequence current magnitude with the sine of its angle as the most prominent feature form. The study develops further traditional ML methods such as decision trees (DT), support vector machine (SVM), random forest (RF), k-nearest neighbor (KNN), an ensemble learning (EL), and a designed one-dimensional convolutional neural network (1D-CNN). The results found that the combination of 1D-CNN or SVM with the most prominent feature achieved the highest accuracies of 99.37% and 99.13%, respectively, with acceptable/fast prediction times, enhancing VSSL. The exceptional performance of the CNN was also approved by field measurements in a real power network. However, selecting the best ML methods for deployment requires a trade-off between accuracy and real-time implementation requirements. The research findings benefit network operators, large factory owners, and renewable energy park producers. They enable preventive maintenance, reduce equipment downtime/damage in industry and electrical power systems, mitigate financial losses, and facilitate the assignment of power-quality penalties to responsible parties.

## 1. Introduction

### 1.1. Background

The growing utilization of sensitive equipment, multiple inverter-based renewable energy generations in electrical distribution systems and microgrids (Dhara et al., 2024), along with the integration of electric vehicles (EVs) (Sundarakani et al., 2024), in modern electrical power systems underscores the importance of power quality. Among various disturbances, voltage sags (dips) are the most common short-term disruptions. As per the IEEE Standard 1564 (IEEE Guide for Voltage Sag, 1564-2014), a voltage sag occurs when the root mean square (RMS) voltage drops between 0.1 pu and 0.9 pu of the nominal voltage for a duration ranging from 0.5 cycles to 1 min. These sags can result from events such as transient short circuit faults, both symmetrical (Sym) or asymmetrical (Asym), the starting or loading of large induction motors (IMs), and the inrush current during transformer energization (TE). Voltage sags can lead to equipment trips, resulting in industrial and economic losses (Bhujade et al., 2023). According to (Bhujade et al., 2023), the annual cost of load loss due to voltage sags for an industrial distribution network was reported as $87.248 without photovoltaic (PV) penetration and $72.254 with a 78% PV penetration. Research on voltage sags remains a significant area of interest, especially given the increasing use of renewable energy resources (RERs) and EVs. Voltage sags can impact the quality of solar PV generation significantly (Ismail and Jamaludin, 2023) and the charging performance of EVs in the grid-to-vehicle mode (Zhu et al., 2023). Studies in this field have focused on aspects such as detecting the sag's start time (Depally et al., 2023), identifying their type (Huchche and Patne, 2023), pinpointing their source (Veizaga et al., 2023), and, crucially, voltage sag source localization (VSSL). The latter holds great significance, as it is the first step in

* Corresponding author. Håken Gullesons väg 20, 901 87, Umeå, Sweden.
*E-mail address:* Younes.mohammadi@umu.se (Y. Mohammadi).

https://doi.org/10.1016/j.engappai.2024.108331
Received 16 January 2024; Received in revised form 15 March 2024; Accepted 23 March 2024
Available online 2 April 2024

**Abbreviation**

| | |
|---|---|
| Acc | Accuracy |
| Asym | Asymmetrical |
| CBM | Current-based method |
| CNN | Convolutional neural network |
| Conv | Convolutional |
| DL | Deep learning |
| DR | Distance relay |
| DS | Downstream |
| DT | Decision tree |
| EL | Ensemble learning |
| IB | Instantaneous-based |
| IM | Induction motor |
| KNN | K-nearest neighbor |
| $LE_x$ | Line to earth (ground) fault in phase x∈{a,b,c} |
| $LL_{xy}$ | Line-line fault between phases x and y ∈{a,b,c}, x ≠ y |
| $LLE_{xy}$ | Line-line to earth fault between phases x and y ∈{a,b,c}, x ≠ y |
| LLL | Line-line-line fault |
| LLLE | Line-line-line to earth fault |
| LR | Logistic regression |
| ML | Machine learning |
| NN | Neural network |
| PB | Phasor-based |
| PQM | Power quality monitor |
| RBF | Radial basic function |
| RCC | Real current component |
| ReLu | Rectified Linear Unit |
| RF | Random forest |
| RS | Resistance sign |
| Seq | Sequence |
| SST | Slope of system trajectory |
| SVM | Support vector machine |
| Sym | Symmetrical |
| TE | Transformer energizing |
| TSB | Time-sample-based |
| US | Upstream |
| VSSL | Voltage sag source localization |

minimizing substantial productivity losses in typical industrial setups, providing preventive actions, improving overall power quality, and assigning accountability for sags caused by faults or TE events (Moghaddam, 2021; Sun et al., 2019). VSSL is even more critical, since these sags can propagate across interconnected networks, affecting distant loads far from the origin (Paul et al., 2020; Aljarrah et al., 2024). Consequently, there is a necessity to develop accurate methods for identifying the source location of voltage sags precisely. Otherwise, strategies such as optimal integration of PVs become necessary, to reduce the severity of voltage sags (Cebrian et al., 2024) or modern control approaches to integrate EV charging stations while addressing voltage sags (Rao et al., 2024). The VSSL methods could then be incorporated as directional functions in protection relays and electric power meters at connection points within production-customer systems. This is particularly pertinent, as preventive actions, in systems with a high penetration of RERs, EVs (in vehicle-to-grid mode), and energy storage systems, where source direction might be obscured (Wu et al., 2020). Recording synchronized waveforms in waveform measurement units (WMUs) (Mohsenian-Rad and Xu, 2023) and applying VSSL methods to them, enable accurate pinpointing of the source location for the same voltage sag. This process provides precise information about the nearest location of sag sources, serving as a fault locator in case of faults associated with the sag sources. In this context, machine learning (ML) emerges as a potent tool in VSSL analysis, capable of enhancing the accuracy (Acc) of fundamental methods which do not use ML, and reducing processing time. However, the challenge lies in selecting appropriate input features, labeling the voltage sag data, and choosing proper learning techniques.

### 1.2. Literature review

Several studies have already been conducted regarding the VSSL task. These studies fall into two main categories: multi-monitoring and single-monitoring methods. Multi-monitoring methods, as exemplified by studies such as (Deng et al., 2021; Liu et al., 2023; Li et al., 2023a), employ data from multiple monitoring points for VSSL or identification of a faulty power line. However, utilizing these methods can be costly, due to the need for data centers and communication resources. Our research concentrates primarily on single-monitoring methods. These methods utilize measurements from a single point to determine the location of voltage sags, whether upstream (US) or downstream (DS), relative to that point. Many single-monitoring methods can also be extended to multi-monitoring scenarios by applying them to various measuring points. Furthermore, single-monitoring VSSL methods can be divided into three types: analytical methods, signal processing-based methods, and methods based on ML, as given in Tables 1–3, respectively. The analytical methods rely on specific criteria, such as changes in power, energy, current, voltage, impedance, and resistance before and after a voltage sag. These criteria help ascertain the direction of the sag source (DS/US). A review of the literature encompassing works related to these methods is presented in Table 1, covering the period from 2008 to the present. The Table's organization includes the method's rule, the type of variables used (three-phase (*abc*), positive/-negative sequence ( ± Seq), Clarke ($\alpha\beta$), Park ($dq$), phasor-based (PB), or instantaneous-based (IB)), method's name, treatment of measurement noise, investigated sag source, utilization of actual sags, and publication year. In Table 1, the abbreviations MDPE, MRCC, MSST, MCBM, and MDR refer to modified versions of their respective original methods. Furthermore, an analysis on the mentioned methods was conducted in (Solak and Sikorski, 2019) briefly and in (Mohammadi et al., 2017a) as detailed. The second type of single-monitoring methods entails the utilization of signal processing techniques, sometimes complemented by optimization algorithms. Table 2 offers an overview of these approaches, indicating the notably fewer number of publications on these methods in comparison to the analytical methods outlined in Table 1.

The third type employs ML approaches, including traditional techniques, shallow and deep neural networks. These approaches frame the VSSL problem as a binary classification problem to determine the sag source as DS/US. The ML-based methods can be divided further into two groups: methods utilizing a discrete subset of features sensitive to sag source location, and those employing sequences of voltage and/or current signals captured a few cycles[1] before and after a voltage sag. The primary focus of our research in this paper is on the latter group of methods. A comprehensive survey of these methods, spanning the period from 2007 to the present, is presented in Table 3.

While investigating ML methods, one significant challenge involves addressing imbalanced datasets, particularly in classification problems. Alongside methodologies such as the synthetic minority oversampling technique (SMOTE), the work referenced in (Lu et al., 2024) implemented a four-stage approach involving clustering, filtering, auto-encoding, and oversampling for the minority class in binary

[1] One cycle is 20 msec in 50 Hz systems and 16.66 msec in 60 Hz systems.

**Table 1**
Overview of analytical methods for VSSL.

| Employed feature' type (rules) | Component | Sag sources | Method | Noise included | Using actual sags | Published year | [Ref.] |
|---|---|---|---|---|---|---|---|
| Polarity of disturbance power and energy (DPE) | *abc* | Fault | DPE, reactive power (RP), Power-flow based | No | No | 2000, 2007, 2015 | (Parsons et al., 2000) ( Passos et al., 2015) |
| Polarity of slope sign of voltage vs. current changes | *abc* | Fault | Slope of system trajectory (SST) | No | No | 2003 | Li et al. (2003) |
| Polarity of the real current component (RCC) changes | *abc* | Fault | RCC | No | No | 2004 | Hamzah et al. (2004) |
| Polarity of the real part of incremental impedance | + Seq | Fault, IM | Resistance sign (RS) | No | Yes | 2005 | Tayjasanant et al. (2005) |
| Polarity of impedance magnitude and angle changes of a distance relay (DR) | + Seq | Fault | DR, Generalized DR | No | No | 2005, 2011 | (Pradhan and Routray, 2005; Yilin and Yonghai, 2011) |
| Polarity of phase change in sequence current (PCSC) | + Seq | Fault | PCSC | No | No | 2007 | Pradhan et al. (2007) |
| Polarity of power, current, and total harmonic distortion changes | *abc*, + Seq | Fault, IM, TE | – | No | No | 2008 | Ahn et al. (2008) |
| A generalized version of (Parsons et al., 2000; Li et al., 2003; Hamzah et al., 2004; Tayjasanant et al., 2005) | $\alpha\beta$ | Fault, IM, TE | $DPE_{\alpha\beta}$, $RCC_{\alpha\beta}$, $SST_{\alpha\beta}$, $RS_{\alpha\beta}$ | No | Yes | 2008 | Polajzer et al. (2008) |
| A statistical combination of (Li et al., 2003; Hamzah et al., 2004; Tayjasanant et al., 2005; Pradhan and Routray, 2005; Pradhan et al., 2007) | *abc*, + Seq | Fault | – | No | Yes | 2009 | Núñez et al. (2010) |
| Polarity of impedance ($Z$) changes | *dq* | Fault, IM | $Z_{dq}$ | No | No | 2010 | Shao et al. (2010) |
| Polarity of incremental impedance | − Seq | Fault | RS- | No | No | 2011 | Kanokbannakorn et al. (2011) |
| Polarity of the current magnitude and angle changes | + Seq | Fault | Current-based method (CBM) | No | No | 2012, 2013 | (Moradi et al., 2012; Moradi and Mohammadi, 2012, 2013) |
| A generalized version of (Moradi et al., 2012; Moradi and Mohammadi, 2012, 2013) | $\alpha\beta$ | Fault, IM, TE | $CBM_{\alpha\beta}$ | No | Yes | 2015 | Polajžer et al. (2015a) |
| A generalized version of (Moradi et al., 2012; Moradi and Mohammadi, 2012, 2013) | IB + Seq | Fault, IM, TE | $CBM_{+}$ | No | Yes | 2015 | Polajžer et al. (2015b) |
| Polarity of the real part of internal impedance | 3-phase | Fault, IM | Internal RS (IRS) | – | – | 2016 | Yi et al. (2016) |
| A generalized version of (Parsons et al., 2000; Hamzah et al., 2004; Tayjasanant et al., 2005; Pradhan and Routray, 2005; Yilin and Yonghai, 2011) | IB + Seq | Fault | $DPE_{+}$, $RP_{+}$, $RCC_{+}$, $SST_{+}$, $DR_{+}$ | – | – | 2017 | Mohammadi et al. (2017b) |
| A version of (Pradhan and Routray, 2005) applicable in systems with inverter-based distributed generators (IBDGs) | + Seq | Fault | Improved DR (IDR) | No | No | 2020 | Mohammadi and Leborgne (2020a) |
| A version of (Moradi et al., 2012; Moradi and Mohammadi, 2012, 2013) applicable in systems with IBDGs | + Seq | Fault | Improved CBM (ICBM) | No | No | 2020 | Mohammadi and Leborgne (2020b) |
| Modified version of (Parsons et al., 2000; Hamzah et al., 2004; Pradhan and Routray, 2005; Pradhan et al., 2007; Moradi et al., 2012; Moradi and Mohammadi, 2012, 2013) using the transient period of sags | IB and PB + Seq | Fault, TE | MDPE, MRCC, MSST, MCBM, MDR | No | Yes | 2022 | Mohammadi et al. (2022a) |
| Polarity of disturbances in active and reactive power (DQ, DP) | *abc* | Fault, IM | DP, DQ | No | No | 2022 | Jing and Ma (2022) |
| Cosine similarity between the phase current and line voltage of two other phases | + Seq | Fault, IM | Cosine similarity (CS) | Yes | No | 2022 | Saadat et al. (2022) |
| A total score containing a threshold obtained from some rules using maximum and minimum values of voltage and current signals | *abc* | Fault | Total score (TS) | No | Yes | 2022 | Yalman et al. (2022a) |
| Extraction of voltage sag applied for VSSL | *abc* | Fault | DPE using auxiliary PQMs | Yes | No | 2023 | Tan et al. (2023) |
| Voltage magnitude difference ($\Delta V$) between an HV and an MV bus exceeding a limit | *abc* | Fault | $\Delta V_{limit\%}$ | No | Yes | 2023 | Castello et al. (2023) |

classification tasks. In addition to the ML methods reported commonly in the literature for VSSL, a diverse array of methods is applied to various aspects of electrical engineering, underscoring the broad applicability of ML beyond VSSL. For instance, in the realm of electrical power quality, the authors in (Zhang et al., 2024) proposed a combination of CNN and long short-term memory (LSTM) networks effective against adversarial attacks for classifying power quality disturbances. In the field of control systems (Zhuang et al., 2023), presents a controller for fault estimation and tolerance in a mobile robot with two independent driving wheels using Q-learning, a subset of reinforcement learning in ML. Furthermore (Wang et al., 2023), introduces an optimal learning control approach for linear systems. In the domain of resiliency (Fatima et al., 2024), provides an extensive review of ML applications in predicting power outages during hurricanes.

### 1.3. Research gaps, contributions, and applications

#### 1.3.1. Research gaps toward contributions

Investigating the 55 literature reported in Tables 1–3, the existing analytical techniques (listed in Table 1) demonstrate a certain level of

**Table 2**
Overview of signal-processing-based methods for VSSL.

| Employed features (rules) | Component | Sag sources | Used method | Noise included | Using actual sags | Published year | [Ref.] |
|---|---|---|---|---|---|---|---|
| A development of ( Parsons et al., 2000) | *abc* | Fault | Hilbert, S, S-(time-time) TT transforms | No | No | 2008, 2013,2020 | (Kong et al., 2008) (Ling Ai and Shareef, 2013) (Shareef et al., 2013) (Chen et al., 2020) |
| A standard sequence of sag waveform | *abc* | Fault, IM, TE | Phase space reconstruction + Fuzzy grey relational analysis | No | No | 2023 | Zeyu et al. (2023) |
| A standard sequence of sag waveform | *abc* | Fault, IM, TE | Particle swarm optimization (PSO) + correlation analysis | No | No | 2023 | Dehong et al. (1990) |

Acc consistently, averaging around 90% based on (Mohammadi et al., 2022b). Moreover, there is also clear potential for improving the performance of ML-based methods (presented in Table 3), even though recent studies have achieved enhanced Acc, with reported rates of 95% (Liu et al., 2022), 96.7% (Xu et al., 2022), and even up to 98% (Yalman et al., 2022b).

While the majority of the aforementioned methods, have been tested primarily against sags caused by faults, a small subset of the techniques (12 out of 55), such as (Polajžer et al., 2015a; Dehong et al., 1990; Junjian et al., 2023b), have considered sags originating from TE events. These TE-related sags tend to display distinctive harmonic imbalances, often characterized by prolonged recovery times (Peng et al., 2013), which impact sensitive loads and wind turbines significantly (Aristi et al., 2009; Mohseni et al., 2011). Consequently, these situations call for the implementation of supplementary compensation strategies (Molla and Kuo, 2020). Therefore, addressing VSSL stemming from TEs holds notable importance for network operators, particularly in cases where events like energization after auto-reclosing of protection relays lack accompanying information (Polajžer et al., 2007). It is also observed that most evaluations of these methods have centered predominantly around clean voltage sags, with only a handful of studies venturing into the realm of noisy signals (6 out of 55), such as (Tan et al., 2023; Mohammadi et al., 2022b; Junjian et al., 2023a). Real-world measurements inherently encompass varying levels of noise, which undoubtedly affect the effectiveness of the applied techniques.

Additionally, certain existing ML-based methods rely on intricate mathematical models (Wu et al., 2020). Moreover, some of the method's Acc has not been tested on large real-world systems (4 out of 55), such as (Junjian et al., 2023a, 2023b). A number of these methods also incorporate a threshold, such as the approach in (Yalman et al., 2022a), which may need adjustments based on the specific system characteristics. Besides, there is a gap in the development of ML-based techniques that rely on straightforward measurements obtainable from power quality monitors (PQMs). The existing literature (Khosravi et al., 2007, 2008; Meléndez et al., 2008) only utilizes the magnitude information of voltage and current waveforms, resulting in limited Acc. Incorporating angle information could provide more comprehensive insights into VSSL problems. These methods often depend on the magnitude and angle of voltage and current phasors recorded over several cycles before and after sags. This concept is what we refer to as "time-sample-based (TSB) features" or time-series.

In order to address these research gaps, this paper develops highly accurate ML methods to enhance the VSSL. The methods are designed to be applicable to sags caused by both power system faults and TEs, eliminating the need for a threshold and specific optimization algorithms, making it straightforward and versatile with low complexity and/or computational cost. In this context, to circumvent the feature selection challenges posed by traditional ML methods, the paper proposes utilizing input TSB features and strives to identify the prominent (most influential) ones. This approach differs from previous methods that rely on a subset of instantaneous-based or phasor-based scalar features, as was proposed in (Mohammadi et al., 2017c, 2021, 2022b; Yalman et al., 2022b). Sub-features of these types necessitate a feature selection step and the identification of a proper combination of features. Furthermore, their utilization demands greater hardware capacity for the computational tools applied to each individual feature, making them more time-consuming in real-world applications. Identifying the most influential feature form, as proposed in this paper, ensures the capture of all relevant information regarding the direction of the voltage sag source during the sag period within a few cycles, even for the short-period sags, instead of picking up only one value during sag or a difference value between the during a sag and pre-sag periods. This contributes directly to improved classification results and enhances the overall performance of the learning-based methods as well as provides more meaningful sequential features for deep learning (DL) methods with a possible less-complex architecture. It is noteworthy that the authors of this study already unearthed crucial information regarding the sag sources within the few transient periods following the starting voltage sags (Mohammadi et al., 2022a).

### 1.3.2. Contributions

A summary of the contributions of this paper into enhancing the VSSL are outlined as follows.

1. Proposing three distinct input feature forms with physical meanings, represented as TSB time series, addressing challenges in feature selection for traditional ML methods, and providing more meaningful sequential features for DL methods to learn from.
2. Identifying the most influential feature form by utilizing k-means clustering, capable of handling unlabeled voltage sag datasets.
3. Developing traditional supervised ML methods, such as DT, SVM, RF, KNN and an EL approach as well as a DL method, i.e., a 1D convolutional neural network (1D-CNN) with a layered feature extractor structure.
4. Addressing measurement errors by introducing varying levels of noise (40, 30, 20, and 15 dB) into the generated sag dataset. This augmentation aims to render the dataset more akin to a real-measurement dataset and train a compact model, distinguishing it from the methodology employed in (Mohammadi et al., 2022b), where multiple models were derived across different noise levels.
5. Identifying the most accurate supervised method(s) across the most influential feature form, regardless of sag source types and noise levels, in a comparative study.
6. To the best of the authors' knowledge, our framework of experiment, incorporating the proposed input TSB features, and especially the designed 1D-CNN combined with the most influential TSB feature form, has not previously been employed in the literature for VSSL problems.

### 1.3.3. Applications

Recognizing the most accurate ML-based method(s), driven by the most influential input feature form within the context of time-series data, provides versatile applications in both online and offline scenarios for VSSL, as outlined below:

Online applications.

**Table 3**
Overview of ML-based methods for VSSL.

| Feature category | Employed feature' type | Component | Sag sources | Used method | Noise included | Using actual sags | Published year | [Ref.] |
|---|---|---|---|---|---|---|---|---|
| TSB (Sequential) | A sequence of RMS voltages and currents | *abc* | Fault | Multivariable principal component analysis (MPCA) + multilayer perceptron (MLP)/radial basic function (RBF)/decision tree (DT) | No | Yes | 2007 | Khosravi et al. (2007) |
| | Similarity between a sequence of RMS three-phase voltage and currents | *abc* | Fault | Principal component analysis (PCA) | No | Yes | 2008 | Meléndez et al. (2008) |
| | A sequence of RMS voltages and currents | *abc* | Fault | MPCA + case-based resonance (CBR) | No | Yes | 2008 | Khosravi et al. (2008) |
| Discrete subset of individual features | A 40-member subset of the ones in ( Parsons et al., 2000; Li et al., 2003; Moradi and Mohammadi, 2012, 2013; Polajžer et al., 2015b; Hamzah et al., 2004; Tayjasanant et al., 2005; Pradhan and Routray, 2005; Yilin and Yonghai, 2011; Pradhan et al., 2007; Polajzer et al., 2008; Shao et al., 2010; Moradi et al., 2012) | IB + Seq, PB + Seq, PB − Seq, $\alpha\beta$ | Fault | Support vector machine (SVM) | No | No | 2017 | Mohammadi et al. (2017c) |
| | A combination of 4 features from ( Parsons et al., 2000; Li et al., 2003; Hamzah et al., 2004; Tayjasanant et al., 2005) and sag magnitude, phase-jump, harmonic content, imbalance, and slope | *abc* | Fault | Feed-forward neural network (FNN) | Yes | No | 2020 | Liu et al. (2020) |
| | Using (Parsons et al., 2000) regarding complex sags | *abc* | Fault | Atomic algorithm and random matrix theory | No | No | 2020 | Wu et al. (2020) |
| | RP changes, impedance angle during a sag, incremental resistance, current magnitude changes | IB + Seq, PB + Seq | Fault | Ensemble learning (EL) on DT learners | No | No | 2021 | Mohammadi et al. (2021) |
| | A 5-member subset from (Parsons et al., 2000; Li et al., 2003; Hamzah et al., 2004; Tayjasanant et al., 2005) | *abc* | Fault | Convolutional neural network (CNN) | No | No | 2021 | Kai et al. (2021) |
| | A 17-member subset of PB features, an 11-member subset of IB features, and a 28-member subset of all PB and IB features from (Mohammadi et al., 2017a, 2017b; Parsons et al., 2000; Polajžer et al., 2015a, 2015b; Tayjasanant et al., 2005; Pradhan and Routray, 2005; Polajzer et al., 2008; Núñez et al., 2010; Shao et al., 2010; Kanokbannakorn et al., 2011; Moradi and Mohammadi, 2013) | IB + Seq, PB + Seq, PB − Seq, $\alpha\beta$ | Fault, TE | SVM, logistic regression (LR), DT, random forest (RF), k-nearest neighbor (KNN), and EL on all | Yes | Yes | 2022 | Mohammadi et al. (2022b) |
| | Sag magnitude, fundamental frequency, Standard Deviation (Std.) frequency curve, RMS of the fundamental frequency, maximum value of phase-jump curve regarding complex sags | *abc* | Fault, IM, TE | CNN | No | No | 2022 | Liu et al. (2022) |
| | A 4-member subset of features from ( Parsons et al., 2000; Li et al., 2003; Hamzah et al., 2004), and an eigenvalue as US/DS | *abc* | Fault | Sparrow search-optimized RBF neural network | No | No | 2022 | Xu et al. (2022) |
| | A subset of features using maximum and minimum values of voltages and currents | *abc* | Fault | Artificial neural network (ANN), SVM, KNN, Naive Bayes (NB) | No | Yes | 2023 | Yalman et al. (2022b) |
| | An 8-member subset of voltage features as a time ratio of sag, no. of mutation points, waveform coefficient, Kurtosis, Skewness, Logarithmic energy entropy, Falling slope of base amplitude frequency, and Std. | *abc* | Fault, IM, TE | Whale-optimized SVM with an RBF kernel | Yes | No | 2023 | Junjian et al. (2023b) |
| | A 16-member subset of features in (Junjian et al., 2023b) | *abc* | Fault, IM, TE | Improved optimized FNN (Extreme learning machine (ELM)) | Yes | No | 2023 | Junjian et al. (2023a) |
| | Mean spectral radius, Std. of the characteristic eigenvalue, maximum eigenvalue, and minimum eigenvalue of a state data matrix | *abc* | Fault | Random matrix theory-CNN | No | No | 2023 | Wu et al. (2023) |

(*continued on next page*)

**Table 3** (*continued*)

| Feature category | Employed feature' type | Component | Sag sources | Used method | Noise included | Using actual sags | Published year | [Ref.] |
|---|---|---|---|---|---|---|---|---|
| | A three-member subset of features in ( Parsons et al., 2000; Li et al., 2003; Hamzah et al., 2004) | *abc* | Fault | RF | No | No | 2023 | Li et al. (2023b) |
| | Degree of voltage sag impact, grid structure, and fault type | *abc* | Fault | MLP and transfer learning (TL) | No | Yes | 2023 | Li et al. (2023c) |

1. **Directional function in directional relays and fault indicators**: Operating as part of a protection relay, addressing voltage sags originating from power system faults in real-time.
2. **Detection of US faults' location in radial electrical power systems**: Offering an alternative to traditional overcurrent and distance relays, making decisions for both permanent and temporary power system faults.
3. **Point of common coupling (PCC) applications**: At the interface between any electricity production and customer systems, such as electrical distribution system and large factories/RERs, parks/microgrids, imposing penalties for distribution system generating voltage sag propagated to customer area and reducing production losses.
4. **Functional module in voltage sag monitoring systems**: Enables real-time source localization in conjunction with voltage sag characteristics in PQMs.
5. **Electrical power networks equipped with WMUs** (Ahmadi-Gorjayi and Mohsenian-Rad, 2023): Particularly applicable when unitizing input data from WMUs operating at high reporting times (e.g., 256 samples per cycle) (Ye et al., 2023), facilitating the referencing of the same voltage sag source recorded among multiple WMUs, and aiding in the precise estimation of the nearest area to the voltage sag's origin location. This application aligns with research (Mohsenian-Rad and Xu, 2023) that refers to it as 'pinpointing the source location of an event,' especially for short-time voltage sags. Contrastingly, phasor measurement units (PMUs) are unable to store all voltage sag information within a few cycles, offering insufficient input for VSSL methods.

Offline applications.

1. **Identification of auto-energized transformers** (Polajžer et al., 2007): Pinpointing the locations of TE events accurately following the auto-reclosing of protection relays causing random (unplanned) sags; subsequently notifying the system operator.
2. **PCC applications between electrical transmission and distribution systems**: Valuable for transmission system operators, facilitating the imposition of penalties on the distribution systems with

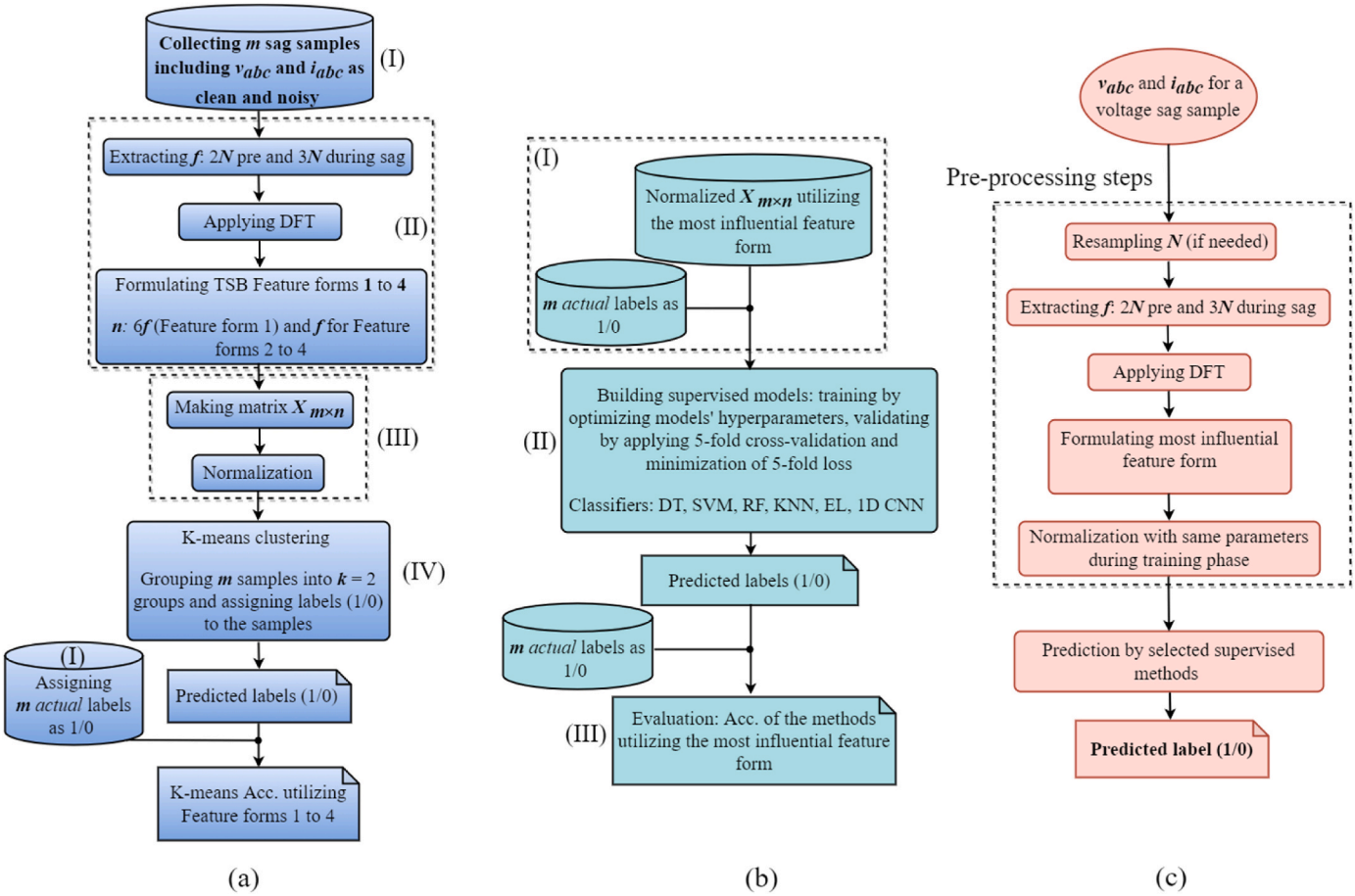


**Fig. 1.** Flow of suggested framework in enhancing VSSL. (a) Identifying the most influential feature form; (b) Identifying the most accurate supervised method utilizing the most influential feature form; (c) Deployment of the selected model to predict a label for unseen sag samples.

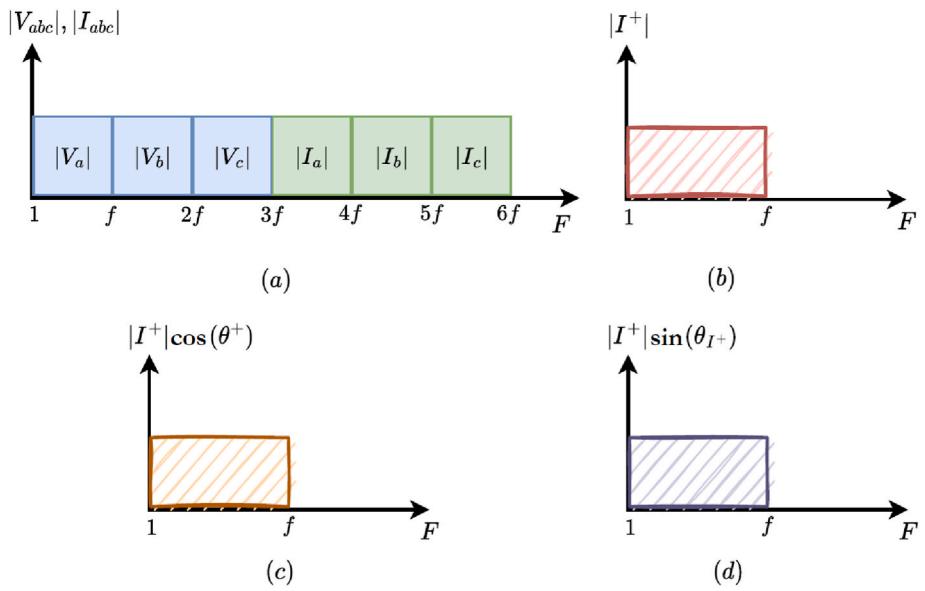


**Fig. 2.** The evolution of the formulated input feature forms (a) Form 1; (b) Form 2; (c) Form 3; (d) Form 4.

significant RERs, when addressing sags occurring in transmission power lines due to power system faults in distribution levels.

### 1.4. Work organization

The paper is structured as follows: **Section 2** explains the experimental framework employed in this study. **Section 3** provides an overview of the dataset used to generate the training and test subsets. In **Section 4**, the outcomes of the developed methods are outlined, following the framework mentioned in Section 2, including the results from extensive numerical simulations and testing via a few field measurements in a real electrical power system. **Section 5** offers a comprehensive discussion, and, lastly, **Section 6** concludes the study.

## 2. Framework of experiment

The experiments in this paper are categorized into three main parts: (a) Identifying the most influential feature form among four forms; (b) Identifying the most accurate supervised method(s) using the most influential feature form; (c) Deployment design for real-time implementation. The complete flow of the experiment is shown in Fig. 1. A detailed explanation of each part is given further in Sections 2.1–2.3.

### 2.1. Part a: Identifying the most influential feature form among four forms

I. Collecting voltage sag samples and assigning labels

The first step in part *a* of the suggested framework involves creating an appropriate dataset (referred to as (I) in Fig. 1a). This can be achieved either through synthetic data obtained from numerical simulations, or by utilizing real-world measurements collected from the electrical power networks. In our study, due to the challenges in real-world measurements, we obtained the data through numerical simulations. Following this, we assigned a label of DS (1) or US (0) to each sag sample. A label of 1 indicates a sag source downstream of the PQM recording the sag, whereas a label of 0 refers to a sag source upstream of the PQM. A more detailed explanation of this process is provided in Section 3, generation of a dataset.

II. Proposing and formulating input TSB feature forms

This stage is referred to as (II) in Fig. 1a. After collecting the voltage sag data recorded by the PQMs, a few cycles of voltage ($v_{abc}$) and current ($i_{abc}$) waveforms are needed to be selected, both before and during the sag. Here, the parameter $f$ represents the number of recorded instances before and during the sag. In this study, considering the number of samples per cycle as $N$, $f$ encompasses $2 \times N$ samples before and $3 \times N$ samples during a voltage sag. Subsequently, a one-cycle running window discrete Fourier transform (DFT) (Gasquet and Witomski, 1999) is employed to the fundamental component of voltage and current signals to extract their phasors.[2] The next step involves formulating the necessary sequential input features. Four types of input feature forms are generated for each voltage sag sample. The selected input feature forms include one derived from existing literature, termed Feature form 1. Feature forms 2 and 3 were designed based on observations from voltage sag events recorded in a PQM (Fig. 2). Voltage sags, namely, act as energy sinks; consequently, during such events, the DS (US) origin causes an increase (decrease) in currents, as well as an in-active power flow observed in the PQM compared to before the voltage sag occurred. The proposed Feature form 4 integrates both the magnitude and angle of the current, capturing the increasing behavior for the current and the decreasing behavior for the angle associated with voltage sags originating from DS sources. Detailed explanations of feature forms 1 to 4 are provided on the following.

a) Feature form 1

By extracting the RMS values from 3-phase current and voltage phasors, a series of TSB RMS instants ($|V_a|, |V_b|, |V_c|, |I_a|, |I_b|, |I_c|$) is generated, as shown in Fig. 2a. The physical concept behind this feature form is rooted in the distinctions of RMS voltage and current for a DS/US sag sources, as introduced initially in (Li et al., 2003). However, the TSB variation, denoted as Feature form 1, was explored initially in (Khosravi et al., 2007, 2008; Meléndez et al., 2008) w ithin the concept of ML methods.

b) Feature form 2

[2] Phasor represent magnitude and angle of a sinusoidal waveform.

Positive sequence components[3] are prevalent in all types of voltage sags, whether symmetrical or asymmetrical. The existing literature, such as (Pradhan et al., 2007; Moradi et al., 2012; Moradi and Mohammadi, 2012, 2013; Polajžer et al., 2015a, 2015b; Mohammadi and Leborgne, 2020b), demonstrate that in VSSL problems, current information is more advantageous than voltage. Hence, a sequence of instants is produced by extracting only the magnitude of the positive sequence components ($|I^+|$) from 3-phase current phasors. Formulating current variation, as indicated in Feature form 2 (Fig. 2b), is adapted and transformed for the first time in this paper into TSB features, aiming to assess its suitability for VSSL applications assisted by ML methods.

c) Feature form 3

The physical concept of the real (active) current component was proposed as a rule in the analytical method for VSSL in (Hamzah et al., 2004) or incorporated as a part of a subset of scaler features applied to ML methods in (Mohammadi et al., 2017c, 2021, 2022b). By extracting the magnitude and angle of the positive sequence components from 3-phase current and voltage phasors, a series of real current component instants ($|I^+|\cos(\theta^+)$) is produced, as depicted in Fig. 2c, where $\theta^+$ is the angle between the voltage and current positive sequence phasors. Thus, the earlier analytical method (Hamzah et al., 2004) is, for the first time, adapted and transformed in this paper into TSB features, referred to as Feature form 3.

d) Feature form 4

The physical concept of current angle only was initially introduced for VSSL in (Pradhan et al., 2007) or incorporated as a part of a subset of scalar features applied to ML methods in (Mohammadi et al., 2017c, 2021, 2022b). By extracting the magnitude and angle of the positive sequence components from only current phasors, a sequence of $|I^+|$ and $\theta_{I^+}$ instants are produced. However, to integrate both into a single feature form easily applicable to ML methods, two forms were proposed, ($|I^+|\sin(\theta_{I^+})$) and ($|I^+|\cos(\theta_{I^+})$). The former, as TSB imaginary (reactive) current component instants, was chosen, and introduced for the first time in this paper, as a quite unique TSB feature form, referred to as Feature form 4, shown in Fig. 2d. This selection over the instants ($|I^+|\cos(\theta_{I^+})$) was because the ML methods (employed in this study) yielded significantly improved results with this feature form. Note that here $\theta^+$ refers to the angle of the positive sequence current, itself. Hence, the Feature form 4 relies solely on current information, potentially reducing the cost of measurement devices in scenarios where only current transformers are present.

III Making an input matrix and normalization

The four feature forms are used as input samples $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]$, $i = 1, \ldots, m$, where $n$ represents the number of sequential features that will be used further to generate the dataset matrix $X_{m\times n} = [x_1, x_2, \ldots, x_m]^T$ with $m$ rows screening the total number of samples. When using Feature form 1, the value of $n$ is $6f$, while for the other feature forms, it is equal to $f$. In order to ensure that each of the $n$ features in the matrix dataset $X$, composed of input feature forms 1 to 4, contributes equally, a row Z-score normalization (Bagnall et al., 2017) is applied to each $x$ element of matrix $X$ across all $m$ samples as $[(x_i - \mu(x_i))/\sigma(x_i)]$. Here $\mu$ and $\sigma$ are the mean and Standard Deviation of each row of $X$. Following this, a similar normalization is applied to each column of the matrix $X$, like the approach explored in (Olshen and Rajaratnam, 2010). This is because using the row-column normalization approach has shown a more effective results for the developed ML methods in this paper, as compared to only using one of them for the sequential data analyzed in this study. The mentioned stages are referred to as (III) in Fig. 1a.

IV. K-means clustering to analyze input feature forms

The motivation behind proceeding with an unsupervised method stems primarily from (Morais et al., 2010), wherein a significant volume of electrical power system events, voltage sags included, is recorded without labels. This poses a constraint for supervised methods, which function exclusively with labeled data. Furthermore, labeling numerous voltage sags captured by PQMs globally presents a challenge, due to the time-consuming nature of the process, potential errors, and the occasional lack of definitive labels. When analyzing feature forms 1 to 4, although other unsupervised methods such as k-medoid or density-based spatial clustering of applications with noise (DBSCAN) could be utilized, k-means (Mohammadi et al., 2022c, 2022d; Mohammadi et al., 2024) was chosen due to its simplicity and demonstrated effectiveness in addressing real-world engineering problems (Miraftabzadeh et al., 2023). K-means operates by calculating the distances between voltage sag samples and cluster centers. Consequently, depending on the complexity or simplicity of the input dataset (including the type of feature form), it may cluster the data with difficulty and lower accuracy, or with ease and higher accuracy. Therefore, in stage (IV) in Fig.1a, k-means is applied to normalized matrix $X$ composed of the feature forms explained in stage II above. The number of clusters in k-means is selected as fixed here as $k = 2$ since, in the VSSL problem, there are only two meaningful classes or labels: DS labeled by 1 and US labeled by 0. The initial centroids are then obtained using the *k-means++* initialization schema. Each input normalized sample $x_i$ ($i = 1, \ldots, m$) is assigned to a cluster ($\omega_{ij}$) with the shortest squared Euclidean distance ($d_{ij}$) to one of two centroids $\mu_j$ ($j = 1, 2$), Eq. (1).

$$\omega_{ij} = \begin{cases} 1, & if\ j = \underset{j}{\operatorname{argmin}} d_{ij} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

The k-means minimizes Eq. (2), which is the summation of the distances of each $x_i$ to its corresponding $\mu_j$.

$$\min \sum_{j=1}^{2} \sum_{i=1}^{m} \omega_{ij} d_{ij},\ d_{ij} = \left\| x_i - \mu_j \right\|_2^2 \tag{2}$$

Once all samples $x_i$ are assigned, the centroids within a cluster are then updated, Eq. (3).

$$\mu_j = \frac{\sum_{i=1}^{m} \omega_{ij} x_i}{\sum_{i=1}^{m} \omega_{ij}}, j = 1, 2 \tag{3}$$

Iterations of Eqs. (1)–(3) continue until convergence is achieved. To determine the Acc of the k-means, initially, label 0 or 1 is assigned to the samples placed within the two obtained clusters. This can be done by checking a few samples and corresponding actual labels that already exist from simulations. After stage (IV), the remaining samples are labeled accordingly. Finally, the Acc of the method is calculated by comparing the assigned labels with the actual labels. K-means is applied to all four feature forms (Forms 1 to 4), and the models' Acc is evaluated. In addition to evaluating accuracies, a t-distributed stochastic neighbor embedding (t-SNE) visualization schema is utilized to reduce the dimensionality of the high-dimensional data ($n = 6f$ or $n = f$) into two dimensions ("Dim 1" and "Dim 2"), offering insights into the most influential feature form. The results in Section 3 demonstrate that Feature form 4 holds the most significance. This is attributed to its inherent separation within the dataset, enabling k-means to group the data effectively. In conclusion, in the next section (part b of the

[3] Three-phase phasors can be decomposed into positive, negative and zero sequence components.

suggested framework), an investigation is conducted of Feature form 4 within the context of supervised methods. Furthermore, Feature form 1, which has already been explored in the literature (Khosravi et al., 2007, 2008; Meléndez et al., 2008), will be used to facilitate a comparison between the findings from applying Feature form 4 and Feature form 1. This comparison will provide further insights into the performance and relevance of Feature form 4.

### 2.2. Part b: Identifying the most accurate supervised method utilizing the most influential feature form

I. Input for supervised methods

Given the most influential feature form obtained from part a, normalized matrix $X$ utilizing this feature form (and form 1 just for comparison goals) along with the $m$ real labels as 1 and 0 are employed as input for the supervised methods. This step is referred to as (I) in Fig. 1b.

II. Building supervised ML and DL models

In this phase, denoted as (II) in Fig. 1b, the supervised ML and DL models utilized in this study undergo training with optimal hyperparameters across cross-validation folds, aimed at minimizing classification loss. Below, we provide a detailed explanation of the different components involved in building these models.

**Cross-validation technique**: We adopted a 5-fold cross-validation approach for both the ML and DL models to ensure a robust evaluation of the model's performance across different subsets of the data. This technique involved partitioning the matrix $X$ with $m$ rows into five distinct folds, using each fold sequentially as a validation set while training the model on the remaining four folds. This process was repeated five times, each with a different fold serving as the validation (test) set, thereby allowing every data point in our dataset to be used for both training and validation. This method is particularly beneficial in assessing the models' generalization capability, and helps mitigate overfitting by ensuring the model is validated against multiple, diverse data splits.

**Classification loss**: The classification loss computes the average loss across the folds. Within each validation fold, the loss is calculated between samples in the validation set, using a model trained on samples from the corresponding training folds. For traditional ML methods, the 5-fold loss is determined as the misclassification error (4.1), where $\delta$ represents the relationship between the actual label $y$ and the predicted label $\widehat{y}$. However, for the deep CNN, binary cross entropy (logistic loss) (4.2) is utilized, to minimize the error between the actual label $y$ and the expected label $p(y)$, a probability value of a sample belonging to a class ranging from 0 to 1. In (4.1) and (4.2), $c$ denotes the fold index, $ic$ signifies a sample index within fold $c$, and $N_c$ represents the number of samples per fold.

$$\frac{1}{5}\sum_{c=1}^{5}\frac{1}{N_c}\sum_{ic=1}^{N_c}\delta(y_{ic},\widehat{y}_{ic}),\ \delta(y_{ic},\widehat{y}_{ic})=\{1, y_{ic}\neq\widehat{y}_{ic}|0, otherwise\} \tag{4.1}$$

$$\frac{1}{5}\sum_{c=1}^{5}\left(-\frac{1}{N_c}\sum_{ic=1}^{N_c}y_{ic}.\log(p(y_{ic}))+(1-y_{ic}).\log(1-p(y_{ic}))\right) \tag{4.2}$$

**Hyperparameter Tuning**: Bayesian optimization is utilized to identify the optimal supervised models and their hyperparameters for traditional ML methods, aiming to minimize the 5-fold loss (4.1). Hyperparameters encompass elements such as the kernel type in SVM and the number of neighbors in KNN. As for the 1D CNN model, a grid-search approach was employed for parameter search and optimization. This method minimizes the 5-fold loss (4.2) iteratively, while meeting the evaluation metric, i.e., accuracy. Notable hyperparameters in this context include the learning rate and batch size, among others. Selection of the models was based mainly on the optimal hyperparameters in conjunction with monitoring for signs of overfitting through training and validation loss observations.

**Classifiers:** The classifiers include fundamental ML techniques such as DT, SVM, RF, KNN, EL using SVM, RF, and KNN, as well as a DL method (CNN). The selection of traditional ML methods was motivated by their general efficiency and relative simplicity compared to neural networks. In contrast, the CNN was chosen as a powerful DL method, to explore the possible complexities of dealing with different input feature forms, regardless of their accuracy. A short explanation of each supervised method is as follows.

(a) DT

DT (Breiman et al., 2017) is a hierarchical method that involves distinct steps leading to specific decisions. It employs a tree-like structure to represent decision paths, achieved through induction and pruning steps. During the induction step, the tree structure is constructed, while, in the pruning step, the tree's complexities are reduced to prevent overfitting. To map inputs to outputs, the algorithm traverses each path through various tree branches, reaching a final decision. The primary decision function used in DT utilizes the Gini index $G_i$ (5) to estimate the node impurity.

$$G_i = 1 - \sum_{j=1}^{2} {v_{ij}}^2, (i=1,\ldots,m, j=1,2) \tag{5}$$

The value $v_{ij}$ represents the ratio of class $j$ instances among the training data in the *i-th* node of the DT.

(b) SVM

SVM (Mohammadi et al., 2021) is a widely used binary classification method that projects data into a higher-dimensional space to identify an optimal hyperplane for class separation. In this process, the support vectors represent the coordinates in this new $n$-dimensional system. SVM employs various kernels, such as Radial Basic Function (RBF), Polynomial, Linear and sigmoid. After testing these kernels, a Polynomial kernel was selected due to its superior Acc on our data. The strength of SVM lies in its ability to perform well regardless of the number of features $n$; however, its slow training speed makes it less suitable for online applications. In the SVM-Polynomial, the main decision function (6) is employed to make predictions.

$$\widehat{y} = \text{sign}\left(\sum_{i=1}^{m} a_i y_i \left(x_i^T x_i + 1\right)^d\right), d > 0 \tag{6}$$

Here, $d$ represents the Polynomial degree, and $y_i \epsilon \{0, 1\}$ stands for the actual label corresponding to the *i-th* sample $x_i$.

(c) RF

RF (Belgiu and Drăguţ, 2016) differs from a single DT as it utilizes multiple DTs. This bagging-ensemble method is particularly beneficial for classifying data in massive datasets and assessing the importance of each feature in the final decision. RF is preferred over DT because it offers higher Acc and addresses the overfitting issue that DT can encounter effectively. However, due to the multiple DTs involved, RF tends to be slower than other models, making it less suitable for real-time applications. The main decision function in RF (7) calculates a weighted average of predictions across all DTs in a forest. Each DT's prediction is weighted in $W_h(x_i, \acute{x}) \geq 0$ based on the relevance of its input sample $x_i$ to a new sample $\acute{x}$ in the same tree.

$$\widehat{y} = \frac{1}{n_{tree}} \sum_{h=1}^{n_{tree}} \sum_{i=1}^{m} W_h(x_i, \acute{x}) y_i \tag{7}$$

(d) KNN

KNN (Zhang et al., 2017) is a basic and widely used classifier. This technique groups data with similar characteristics, based on the number of neighboring samples specified by $K$ without assuming any specific data distribution. KNN is used commonly in online data mining and pattern recognition tasks. The weighted variant of KNN employs a majority-weighted voting $c_{jmax}$ (8) to determine the class label assigned to a data point $x_q$. This determination is mased when $x_q$ is surrounded by $K$ neighbors, denoted as $x_l$, each situated at a distance $d(x_q, x_l)$ from $x_q$. The function $\delta(c_j, y_l)$ takes two parameters and returns 1 if they are equivalent, and 0 otherwise. Here, $c_j$ signifies the label associated with the *j-th* class, where $j$ can take values 1 or 2.

$$\hat{y} = c_{jmax} = \underset{c_j \epsilon \{0,1\}}{\operatorname{argmax}} \sum_{l=1}^{K} \frac{1}{d\left(x_q, x_l\right)^2} \delta\left(c_j, y_l\right) \tag{8}$$

(e) EL

As an additional supervised method, this study proposes an EL approach that incorporates the classifiers of SVM, RF, and KNN. The output decisions of all these classifiers are combined with the majority voting rule, resulting in a single output for the EL. Thus, the predicted label for a sample $x_i$ is determined by the label that receives the majority of predictions by the SVM, RF, and KNN models for each of the 5 folds. Introducing the EL enhances the probability of more accurate classification results because each classifier behaves differently and operates independently.

(f) Structured deep CNN

In this study, the input feature forms are in the format of one-dimensional signals or time series data. Consequently, a specifically designed one-dimensional convolutional neural network (1D-CNN) is employed to handle the matrix $X$, which comprises feature forms 1 and 4. Although other neural network models such as the LSTM (Cen et al., 2023), are potential candidates, CNN, previously utilized for tasks within the domain of VSSL (Kai et al., 2021; Liu et al., 2022; Wu et al., 2023), has demonstrated satisfactory results and was thus selected for our study. CNN operates on a layer-by-layer feature extraction architecture, enabling it to capture crucial features from the input feature forms effectively.

The 1D CNN typically comprises an input layer, an output layer, and several hidden layers. The relationship between these layers and the applications of 1D-CNN are detailed in (Kiranyaz et al., 2021). The input layer takes the Feature forms 1 or 4 as input, while the output layer, acting as the classification layer, assigns labels of 1 or 0 to the input feature forms. The hidden layers, which consist of Feature form 1 (or 4), are composed of 4 (or 5) Convolutional (Conv) 1D layers, 3 (or 5) Pooling layers, and 3 (3) Dense layers, along with 1 (1) dropout layer with a probability rate, all arranged sequentially. The Conv layers employ shared weights to construct filters, each with a small receptive field for extract features. Local connections and shared weights are the fundamental principles behind the Conv layers. The pooling layers are utilized for down-sampling to summarize extracted features and are typically positioned after each convolutional layer. The max-pooling function is favored over average-pooling due to its superior performance. Dense layers, also known as fully connected (FC) layers, contain learnable parameters positioned at the end of the network, categorizing the input based on features extracted from convolutional and pooling layer sequences. The dropout layer assigns zero to inputs to prevent over-fitting. The complete CNN architecture is provided in Table 4, where optimal values are obtained through hyperparameter tuning, as explained further in Section 4.2. a. The rectified linear unit (ReLU) activation function is applied after the Conv and Dense layers (except the last one in Table 4, where Sigmoid was used) to facilitate faster convergence (Nair and Hinton, 2010). The filter size for the last layer is set to 1, which is appropriate for our binary classification task, where the labels are either 1 or 0.

**Table 4**
A deep designed architecture of 1D-CNN for datasets composed of feature forms 1 and 4.

| Feature form 1 just for comparison goals (input size: 6*f*) | | | Feature form 4 as the most influential one (input size: *f*) | | |
|---|---|---|---|---|---|
| Layers | Filter sizes | Kernel size | Layers | Filter sizes | Kernel size |
| Conv1+ReLu | 32 | 3 | Conv1+ReLu | 8 | 3 |
| MaxPool1 | … | 2 | MaxPool1 | … | 2 |
| Conv2+ReLu | 64 | 3 | Conv2+ReLu | 16 | 3 |
| MaxPool2 | … | 2 | MaxPool2 | … | 2 |
| Conv3+ReLu | 128 | 3 | Conv3+ReLu | 32 | 3 |
| MaxPool3 | … | 2 | MaxPool3 | … | 2 |
| Conv4+ReLu | 256 | 3 | Conv4+ReLu | 64 | 3 |
| Dense1 | 64 | … | MaxPool4 | … | 2 |
| Dropout1 (0.63) | | | Conv5+ReLu | 128 | 3 |
| Dense2+ReLu | 32 | … | MaxPool5 | … | 2 |
| Dense3+Sigmoid | 1 | … | Dense1 | 16 | … |
| | | | Dropout1 (0.65) | | |
| | | | Dense2+ReLu | 8 | … |
| | | | Dense3+Sigmoid | 1 | … |

During the training process of the CNN, the weights $w$ (9) of each layer are updated using an optimization method. The commonly used mini-batch adaptive moment estimation (Adam) optimization method (Wang et al., 2022) is employed in this study.

$$w_t = w_{t-1} - \alpha \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t + \epsilon}}, \widehat{m}_t = \frac{m_t}{1-\beta_1^t}, \widehat{v}_t = \frac{v_t}{1-\beta_2^t}, m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t,\ v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{9}$$

where $t$ is the time parameter, $m$ and $v$ are moving averages, $g$ denotes the gradient on the current mini-batch, $\beta_1, \beta_2 \in [1, 0)$ represents the decay rate of the moving mean index, $\alpha$ corresponds to the learning rate, and $\varepsilon$ is a positive small number.

After examining a few different CNN structures, the current architecture, explained in Table 4, was selected, due to its desirable performance. Having the predicted labels for all $m$ voltage sag samples, and comparing them with the actual labels, the Acc of methods is calculated utilizing the most influential feature form (and Feature form 1). Therefore, the highest-Acc supervised methods and the most accurate one are also identified.

III. Evaluation metric

To evaluate the effectiveness of the learning methods developed in this study (refer to Fig. 1b), in identifying the most influential feature form and the most accurate supervised method(s), the main criterion used was Acc (10). $N_{correct}$ and $N_{total}$ represent the number of true predictions and the total number of observations, respectively. Eq. (10) is also used for the k-means clustering evaluation.

$$\text{Acc}\ (\%) = \frac{N_{correct}}{N_{total}} \times 100 \tag{10}$$

In VSSL problems, there is no inherent priority between the two classes, as both are equally crucial to be detected accurately, depending on the specific application, as elaborated in Section 1.3. c. Therefore, in addition, F1-score metrics were employed for both the DS and US classes to provide a realistic evaluation, particularly in distinguishing between these two imbalanced classes (refer to Appendix B, Equations (B.1) and (B.2)). This metric offers insights into whether a method is more adept at

detecting one class over another, particularly given the unbalanced data distribution in our case study, where 64% of the data are labeled as US and 36% as DS.

### 2.3. Part c: Deployment design for real-time implementation

After training the models, the selection of methods for real-world application depends on several criteria, including balanced predictive performance, model complexity, and prediction time (discussed further in Section 5.3.c). The chosen methods can then be deployed to detect a label (1/0) for a new single voltage sag sample, as illustrated in Fig. 1c. When a voltage sag sample is recorded as $v_{abc}$ and $i_{abc}$ waveforms before the prediction phase by the selected methods, several preprocessing steps are required. These, among others, include resampling the number of samples of voltage sags per cycle to the specified $N$ for the methods, as illustrated in Fig. 1c. A majority voting process is performed on the predictions generated by each of the compact models out of the 5 models, due to the 5-fold process. Another idea could be retraining the selected model on the entire dataset $X$, while keeping the best hyper-parameters, so that there would be a single compact model to predict only one single output. The EL method also utilizes its inherent majority voting technique on the predictions obtained from SVM, RF, and KNN.

## 3. Generation of the dataset

Collecting voltage sag samples in real-world measurements presents several challenges, including: (a) Limited existing current waveform measurements recorded by PQMs, (b) Variability and unpredictability in the occurrence of voltage sags, (c) Potential impact on data quality due to measurement errors, (d) Costs associated with data collection, (d) Time and expertise required for labeling data as DS or US. A simulated case study is employed to address these challenges. This study, modeled using PSCAD/EMTDC, is based on real data from a reginal electrical power system in Mato Grosso, Brazil (Mohammadi et al., 2021) (depicted in simplified form in Fig. 3), to generate the dataset. PQMs are installed at 6 points with different topologies (Table 5). The basic features of the system are provided in Appendix A. Voltage sag samples are generated to consider various scenarios by changing the electrical loads (constant impedance, constant power, and induction motor) and power line impedances (the possibility of installing new lines) randomly. Power system faults are simulated for 5 to 10 cycles randomly under

**Table 5**
Configuration of voltage sag simulations to generate the training and testing datasets.

| Sag source | Parameter | Configuration | Count |
|---|---|---|---|
| Fault | Fault location | F1, F2, …, F15 | 15 |
| | Fault type | LLLE, LLL, $LE_a$, $LE_b$, $LE_c$, LLE $(LL)_{ab}$, LLE $(LL)_{bc}$, LLE $(LL)_{ca}$ | 11 |
| | Fault resistant | 0.001, 1, 10, 40 and 80 Ω | 5 |
| | PQM | PQM1 (at a 138 kV line), PQM2, PQM3, PQM5 (on the boundary between 230 and 138 kV systems), PQM6 (between the 138 kV and 13.8 kV systems) | 6 |
| | Total scenario number | | 15 × 11 × 5 × 6 = 4950 |
| TE | TE location | TE1, TE2, …, TE5 | 5 |
| | Transformer capacity | 20, 40, 60, 80, …, 140 MVA | 7 |
| | Total scenario number | | 5 × 7 × 6 = 210 |
| Total scenario number of faults and TEs | | | 5160 |

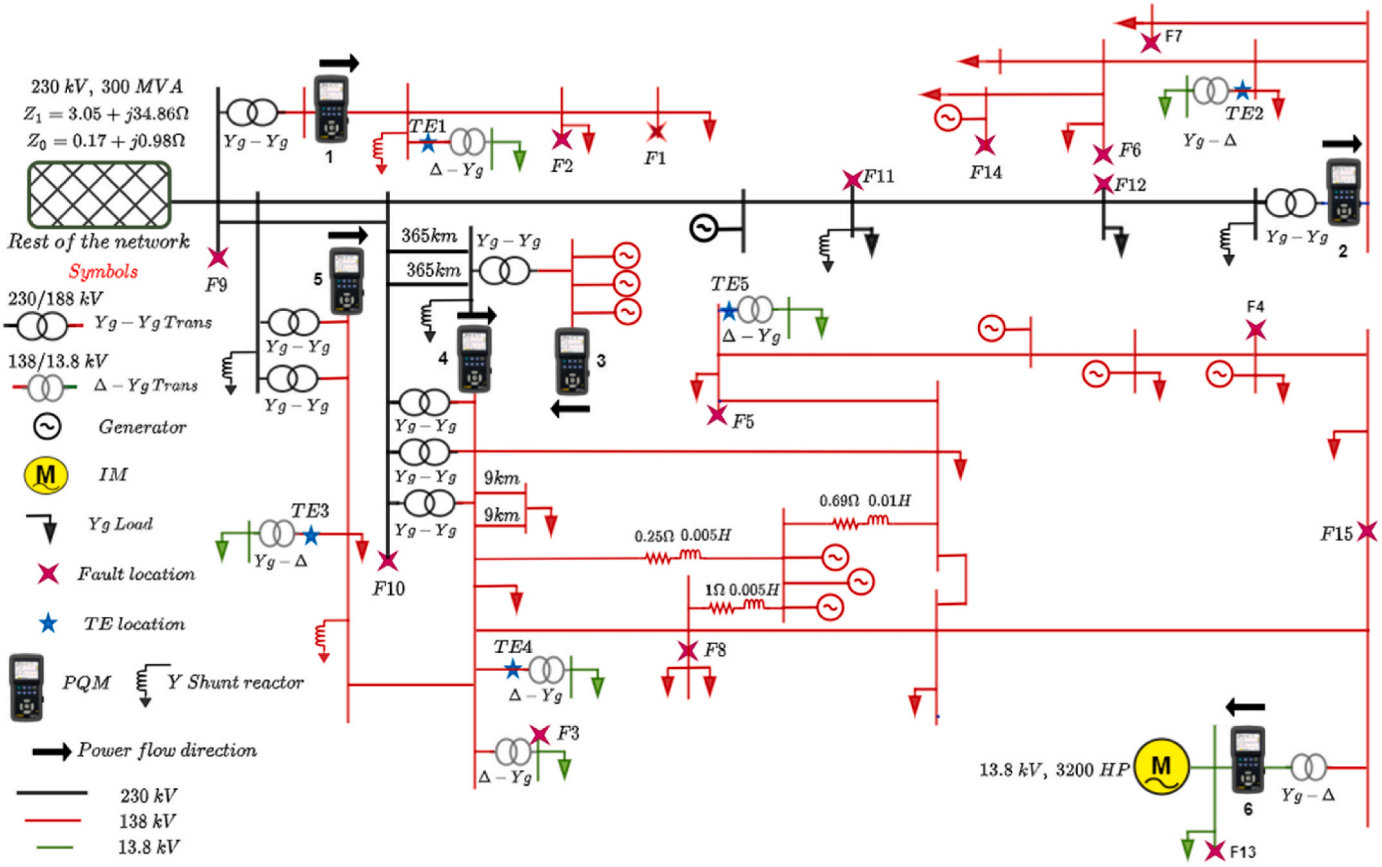


**Fig. 3.** Network for generating voltage sag samples due to short-circuit faults and TEs.

various configurations, as given in Table 5. TE events are also simulated by switching some transformers under the configurations specified in Table 5. However, rare events ((Tayjasanant et al., 2005) and many other works of literature) such as voltage sags due to simultaneous faults, two TEs, or a fault and TE at two DS and US sides, are not considered in this study. An overall of 5160 clean samples are generated, resulting in 4308 clean (noiseless) voltage sag data in voltage and current waveforms.

Real-word measurement errors are simulated by polluting the clean data with white Gaussian noises with signal-to-noise rates (SNRs) of 40 dB, 30 dB, 20 dB, and 15 dB, generating a total of $m = 21{,}540$ voltage sags samples. The samples are segmented into different categories, including Sym sags due to LLLE (LLL) faults, earth fault sags due to $LE_a$, $LE_b$, $LE_c$, $LLE_{ab}$, $LLE_{bc}$, $LLE_{ca}$, and LLLE faults, Asym sags due to $LE_a$, $LE_b$, $LE_c$, LLE $(LL)_{ab}$, LLE $(LL)_{bc}$, LLE $(LL)_{ca}$ faults, and sags due to TE samples. Table 6 summarizes the total number of input samples ($N_{total}$) for the matrix $X_{m\times n}$ and the categorized different parts. The $m$ voltage and current waveforms with a sampling frequency of 7.68 kHz, i.e., $N = 128$ samples/cycle of a network frequency of 60 Hz are later transferred to the four feature forms explained in Section 2.1, part II into four datasets for matrix $X_{m\times n}$. A DFT was used to compute the fundamental frequency components of the four feature input forms. The value of $f$ is 640, including $2 \times N$ before and $2 \times N$ during the voltage sag. Consequently, $n$ would be $6 \times 640 = 3840$ for the matrix $X_{m\times n}$ composed of feature form 1 and 640 for the matrix $X_{m\times n}$ composed of feature forms 2, 3, and 4.

The next step is labeling $m$ voltage sag samples. Voltage sags recorded in PQMs are labeled as DS (corresponding to label 1) if the sag source is in the forward side of the power flow direction (indicated by the black arrow on the top of the PQMs in Fig. 3), otherwise the sag is labeled as US (corresponding to label 0). Additionally, referring to Table 6, the size of the folds used in the cross-validation process applied to ML methods is $21540/5 = 4308$.

Two examples of DS and US voltage sags simulated in the system depicted in Fig. 3 are illustrated in Fig. 4. The time responses are presented here in per unit (PU), meaning that the three-phase voltage and current waveforms are scaled based on their maximum values before the occurrence of the voltage sag.

Concerning the DS Sym sag recorded in PQM2, the 3-phase voltages decrease and currents increase symmetrically, as shown in Fig. 4a and b. However, for the US Asym sag recorded in PQM6, the behavior of the 3-phase voltages and currents is influenced by the IM load and its transient nature during the voltage sag, as depicted in Fig. 4c and d. Feature form 1 (Fig. 4e) indicates a decrease in the RMS value of the voltages for both DS and US sags, which could potentially lead ML methods to make a mistake in distinguishing these two classes. Feature forms 2 to 4 exhibit stable behavior, with an increase in positive sequence current magnitude (Fig. 4f) and positive sequence real current component (Fig. 4g), as well as a decreasing trend for Feature form 4 (Fig. 4h) due to the DS sag. However, for the US sag, Feature form 4 shows a more stable behavior, as its trend was impacted less by the transient behavior of the IM. It's worth noting that analytical phasor-based methods for VSSL have previously demonstrated weak performance in the case of such US sags (Polajžer et al., 2015a, 2015b).

**Table 6**
The number of input samples $m$ used in matrix $X$.

| Dataset | All sags | Sym sags | Asym sags | Earth faults | TE | DS sags | US sags |
|---|---|---|---|---|---|---|---|
| **Total (mix of clean (noiseless) and noisy)** | 21540 | 3735 | 17015 | 13085 | 790 | 7740 | 13800 |
| **Clean (noiseless) or noisy** | 4308 | 747 | 3403 | 2617 | 158 | 1548 | 2760 |

Fig. 5 presents a two-dimensional visualization of all $m = 21{,}540$ generated sag samples, labeled as DS/US, from the four matrices $X$ comprising feature forms 1 to 4. The degree of separation between the sag samples is observed to increase from the dataset composed of feature forms 1 (Fig. 5a) to 4 (Fig. 5d), respectively. Therefore, the dataset composed of feature form 4 (Fig. 5d) appears to exhibit the best classification conditions. This underscores the effectiveness of the TSB Feature form 4, which simplifies the classification process. Section 4.1 will confirm this observation in more detail by employing k-means clustering.

## 4. Results

### 4.1. K-means clustering results to identify the most influential feature form

This section presents the results of k-means clustering depicted in Fig. 1a. The k-means clustering process grouped the data, comprising feature forms 1 to 4, into two clusters, as shown in the 2D scatter plots in Fig. 6. After assigning labels (1/0) to the clustered voltage sag samples and comparing them with the actual labels, the overall accuracies achieved using each feature form are as follows: 83.94% (Feature form 1), 88.89% (Feature form 2), 91.78% (Feature form 3), and 93.9% (Feature form 4). This demonstrates a consistent improvement in k-means' Acc across the utilization of feature forms 1 to 4, with the best performance observed in Feature form 4 (Fig. 6d). Notably, the number of wrong label predictions, denoted by yellow points, is remarkably fewer in Fig. 6d. These trends were also highlighted in Fig. 5, where the inherent class separation in Fig. 5d, representative of Feature form 4, aligns well with k-means' efficiency. The inherent class separation in Feature form 4 enables k-means to minimize the summation of distances of the voltage sag samples within a cluster to the cluster center, achieving the highest possible performance of approximately 94%, which is deemed acceptable for an unsupervised method. Consequently, Feature Form 4 is identified as the most influential feature form based on the outcome.

The exploration of k-means clustering, beyond identifying the most influential feature form, is elaborated further in Table 7. Firstly, the highest Acc was achieved for the clean data subset at 95.33%, maintaining this Acc level even with an SNR of 30 dB. Testing for voltage sags due to TE events has shown 96.96%, 100%, 100%, 100% Acc for the entire dataset, clean data, and datasets with 40 dB and 30 dB SNRs, respectively, further highlighting the appropriateness of Feature form 4. The F1-score for the DS class was consistently lower than that of the US class across all data subsets. This trend can be elucidated by considering the recall and precision values for the total data scenario for instance. Although the recall value for the DS class (98.8%) is higher than that for the US class (91.1%), the precision for the US class (99.3%) surpasses that of the DS class (86.2%) greatly. The substantially higher precision value for the US class resulted in an elevated F1-score for this class, surpassing that of the DS class. In conclusion, k-means clustering using Feature form 4 exhibited superior performance for the US class (the blue dots in Fig. 6d) in a balanced view of the data distribution between the two classes overall. A similar behavior is observed for different subsets of the dataset, with the highest F1-score observed for clean data in the US class. The results presented in Table 7 provide an opportunity for improving the application of k-means clustering as an independent unsupervised method for the VSSL problem. This enhancement will be discussed briefly in Section 5.8.

### 4.2. Insights for the results of the developed supervised methods to identify the most accurate method(s)

#### 4.2.1. Optimizing the hyperparameter values

For hyperparameter tuning, as proposed in Section 2.2. II, Bayesian optimization was employed with a maximum of 30 objective function evaluations for the traditional ML methods, although convergence was

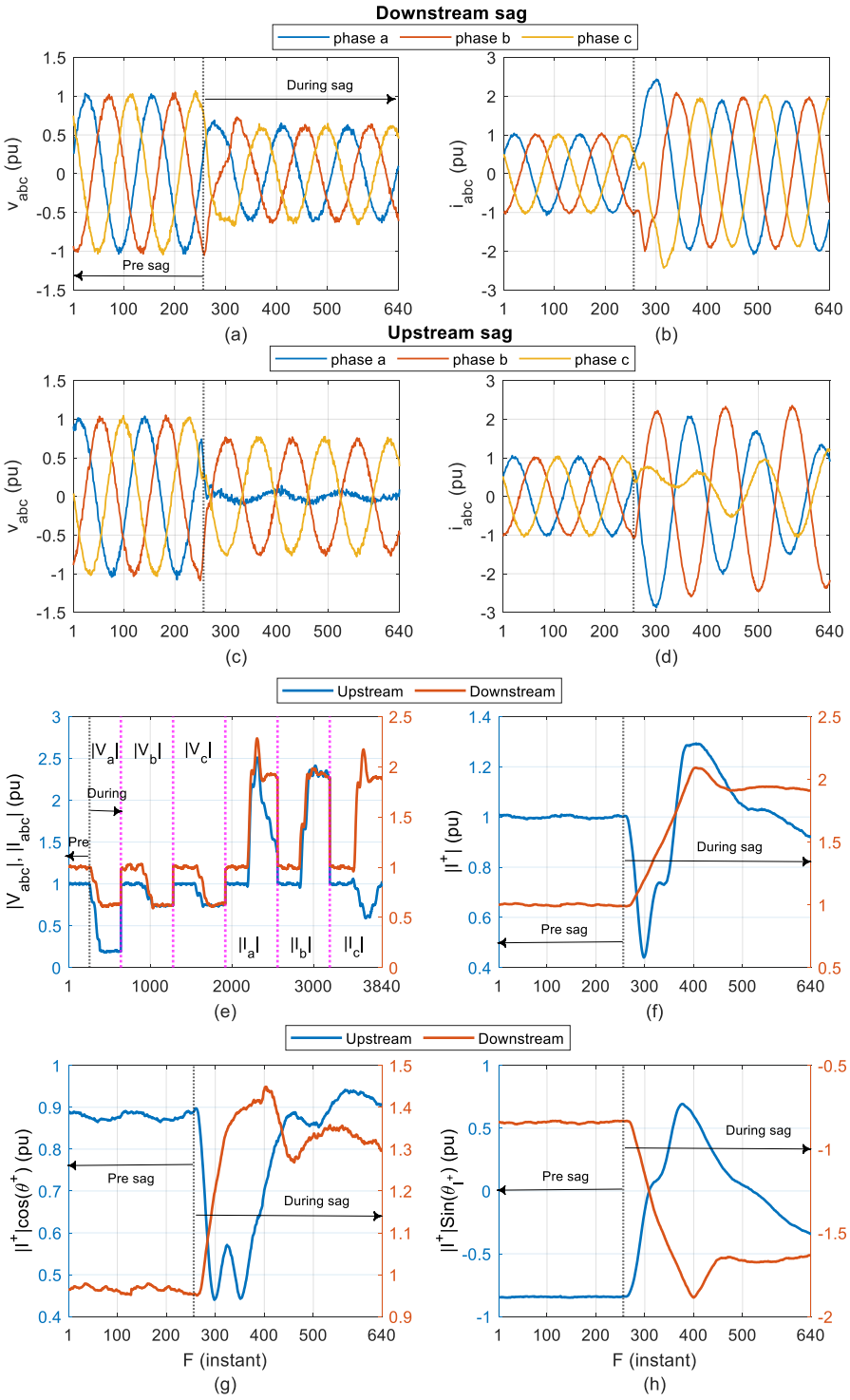


**Fig. 4.** Two examples of DS and US voltage sags polluted with an SNR of 30 dB. (a, b) 3-phase voltages and line current during a voltage sag recorded in PQM2, originating from a DS Sym high-resistance fault at F14; (c, d) 3-phase voltages and line currents during a voltage sag recorded in PQM6, caused by a US Asym low-resistance fault at F15; (e–g) Evolution of feature forms 1 to 4 in response to the voltage sags.

observed in earlier iterations in some cases. Table 8 provides the optimal parameters for these methods. A grid-search method was utilized for the 1D CNN model. The optimization parameters included the learning rate, batch size, architecture of the convolutional and dense layers, and dropout layer probability, tailored specifically for the commonly used Adam optimizer. The optimal architecture for the CNN was already detailed in Table 4, while the remaining hyperparameter values are outlined in Table 8. The maximum number of epochs was set to 50 after careful experimentation. Additionally, early stopping conditions were implemented to cease training if the validation loss stopped decreasing. This helped prevent the model from learning noise in the training data and avoided unnecessary time consumption, particularly in cases where the accuracy didn't improve.

### 4.2.2. *Classification results*

Table 9 presents the overall classification Acc of the six methods

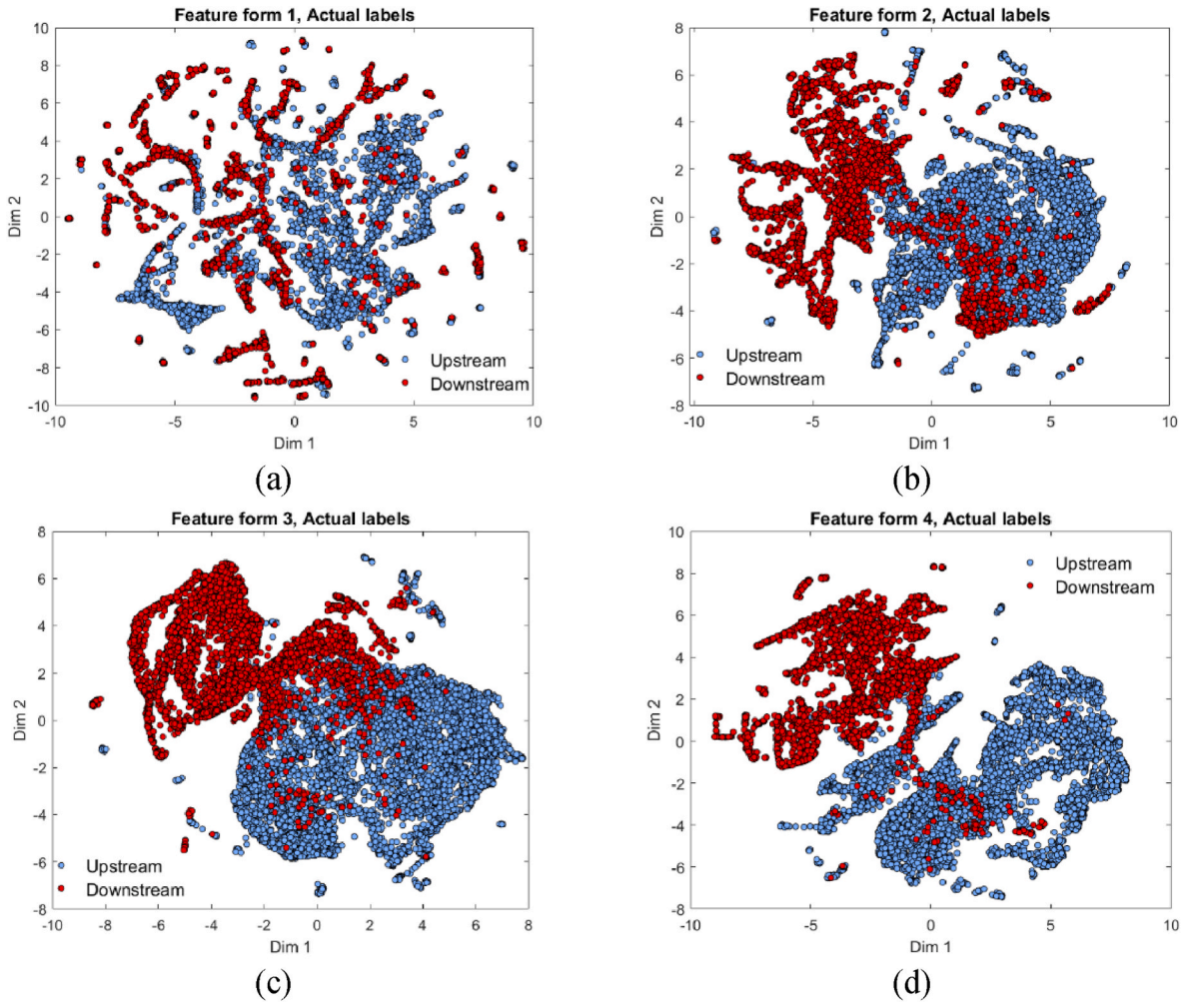


**Fig. 5.** 2D Visualization of *m* sag samples regarding four datasets *X* composed from (a–d) Feature form 1 to Form 4. Dim 1 and Dim 2 refer to the two principal components obtained from t-SNE. The sample's actual labels are marked with blue points (US) and red points (DS).

utilizing input Feature forms 1 and 4 across different data subsets. Within each column, the Acc values are arranged in ascending order of color intensity, ranging from the lowest (dark red) to the highest (dark green). Moreover, the entry marked with a ✓ represents the highest Acc, while the entry marked with ☑ is closest to the highest Acc. From the insights provided by Table 9, the following conclusions can be drawn.

1) The addition of noise to the clean dataset leads to a decrease in the methods' Acc, although with minimal impact on the CNN employing both Feature forms 1 and 4, highlighting its inherent resistance to noise.
2) Across different input Feature forms 1 or 4 and noise levels, the DT and CNN displayed the lowest and highest performance, respectively. Remarkably, KNN fed with feature form 1 achieved a 99.91% Acc, closely trailing the CNN with a 99.95% Acc utilizing Feature form 4 against 40 dB SNR.
3) Among the fundamental ML methods, the EL utilizing Feature form 1 achieved a notable Acc of 98.43% for the entire dataset, while SVM utilizing Feature form 4 exhibited a high Acc of 99.16%.
4) The superiority of Acc is evident when employing Feature form 4 compared to 1 across various methods, highlighting the effectiveness of Feature form 4 in our classification scenario. For example, SVM-Feature form 1, with an Acc of 97.8%, was improved to 99.16% with SVM-Feature form 4 for the entire dataset.
5) Exceptional Acc levels of 100% were attained by SVM, EL, and CNN when using Feature form 1. However, this trend was observed for all methods (except DT) when utilizing Feature form 4.
6) Overall, the most accurate TSB methods were identified as follows: CNN-Feature form 4 (99.37%), followed by SVM-Feature form 4 (99.16%), CNN-Feature form 1 (99.13%), and EL-Feature Form 4 (99.07%), RF-Feature form 4 (99%), all demonstrating 100% Acc against noiseless voltage sags.

The scatter plots in Figs. 7 and 8 depict the classification outcomes of the methods into two classes for the entire dataset, utilizing Feature forms 1 and 4, respectively. Notably, in Fig. 8f, representing the designed CNN method using Feature form 4, only a minimal number of erroneous predictions are denoted by the yellow data points, confirming this method as the most accurate one.

Emphasizing the significance of the DS and US classes, the confusion matrices of the methods are presented in Figs. 9 and 10. The diagonal elements of the confusion matrices represent the count of accurately predicted samples in each class (TP, TN), while the off-diagonal elements indicate incorrect classifications (FN, FP). The accuracies observed in Table 9 and Figs. 7 and 8 are also visualized in Figs. 9 and 10. The minimum FN and FP values are seen for CNN-Feature form 4 as 83 and 52. The reason that SVM-Feature form 4 is slightly more accurate than CNN-Feature form 1 is the lower FN value for SVM (94) compared to CNN (110).

To gain insight into the inaccuracies highlighted by the yellow points in Figs. 7 and 8, as well as the FN and FP values in Figs. 9 and 10,

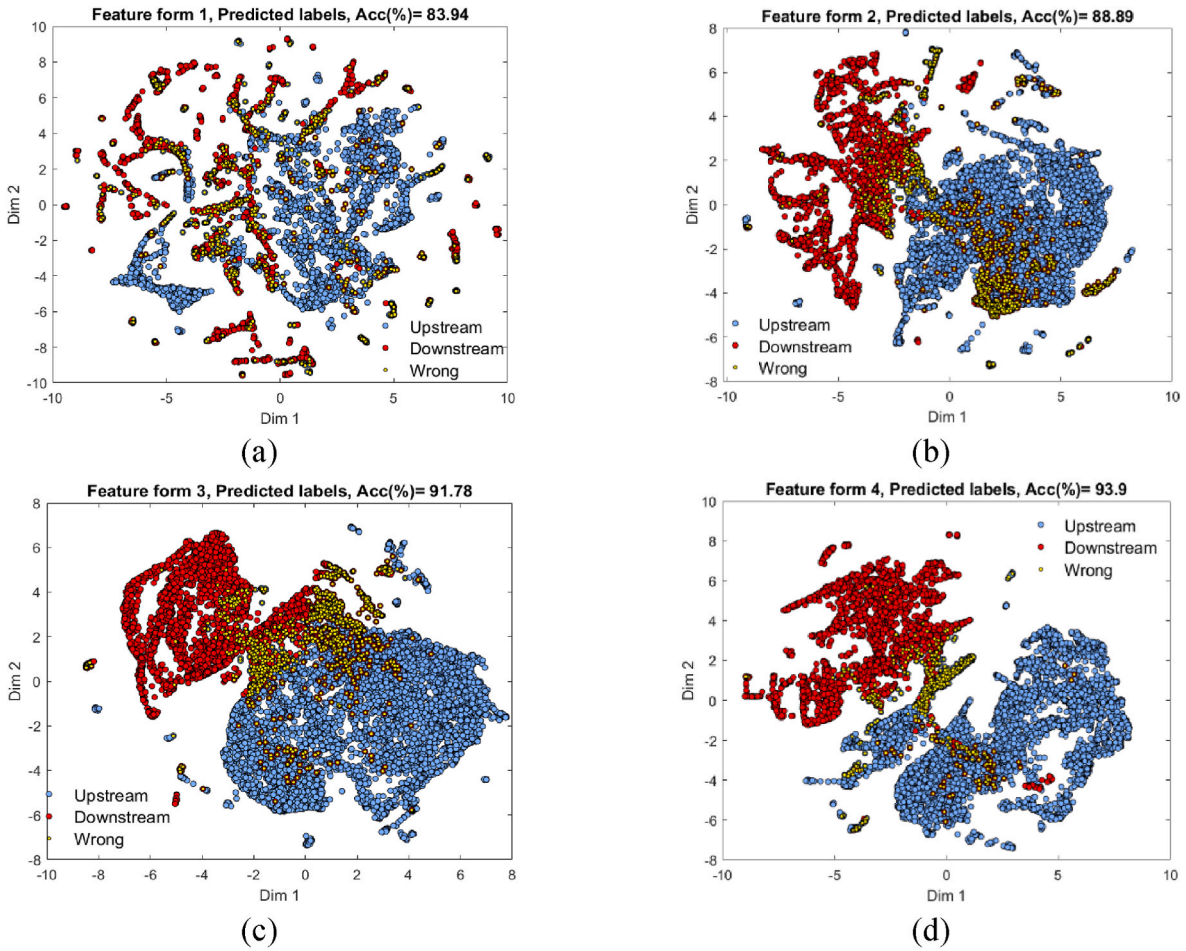


**Fig. 6.** K-means clustering scatter plot for four datasets $X$ composed from (a–d) Feature form 1 to form 4. Dim 1 and Dim 2 refer to the two principal components obtained from t-SNE. The sample's predicted labels are marked with blue points (US), red points (DS), and yellow points (Wrong). The Acc (%) for each dataset is given at the top of each Figure.

**Table 7**
Performance of k-means clustering utilizing Feature form 4 as the most influential Feature form.

| SNR | Acc. for all sags (%) | Acc. for TE sags (%) | Precision (%) | Recall (%) | F1-Score (%) | Class |
|---|---|---|---|---|---|---|
| Total (mix of clean and noisy) | 93.90 | 96.96 | 86.2<br>99.3 | 98.8<br>91.1 | 92.09<br>95.04 | DS<br>US |
| Clean | 95.33 | 100 | 88.5<br>100 | 100<br>92.7 | 93.9<br>96.22 | DS<br>US |
| 40 dB | 95.06 | 100 | 88.3<br>99.7 | 99.5<br>92.6 | 93.53<br>96 | DS<br>US |
| 30 dB | 94.85 | 100 | 87.9<br>99.6 | 99.4<br>92.3 | 93.27<br>95.83 | DS<br>US |
| 20 dB | 92.87 | 95.57 | 84.5<br>98.8 | 98.1<br>89.9 | 90.82<br>94.18 | DS<br>US |
| 15 dB | 91.39 | 89.24 | 82.2<br>98.2 | 97.1<br>88.2 | 89.01<br>92.92 | DS<br>US |

Table 10 illustrates eight random samples of voltage sags where most supervised TSB methods made incorrect predictions. This was due to mainly either the voltage sag being shallow or the high level of noise present. Except for samples 2 and 6, where the voltage sag magnitude was 0.75 and 0.8 pu, respectively, the remaining six samples had shallow sags with a magnitude around 0.9 pu. However, the CNN method, using Feature forms 4 (or 1), and the EL-Feature form 4 exhibited fewer incorrect detections for such samples.

**Table 8**
Optimal hyperparameters of the supervised TSB methods.

| Method | Feature form 1 just for comparison goals | Feature form 4 as the most influential one |
|---|---|---|
| DT | MinLeafSize[1]:10 | MinLeafSize:10 |
| SVM | KernelFunction: Polynomial, $C^2$: 433.13, KernelScale: 42.216, $d^3$: 4 | KernelFunction: Polynomial, $C$: 895.54, KernelScale: 105.22, $d$: 2 |
| RF | Max_features[4]: sqrt, n_estimator[5]: 1000 | Max_features: sqrt, n_estimator: 100 |
| KNN | $K^6 = 3$, Distance: Correlation | $K = 7$, Distance: Euclidian |
| Designed 1D-CNN | Learning rate:$10^{-4}$, Batch size: 64 | Learning rate: $10^{-4}$, Batch size: 64 |

1: Minimum number of samples required to be at a leaf node, 2: Box constraints, 3: Degree of polynomial, 4: Maximum number of features when looking for the best split, 5: Number of trees in the forest, 6: Number of neighbors.

Table 11 provides a comprehensive performance comparison between the most-accurate identified methods for various types of voltage sags, considering different criteria such as recall, precision, and F1-score for both classes, which can be important due to the imbalance distribution of the two classes. As shown in this Table, among the five most-accurate identified methods, four of them utilized Feature form 4. However, the designed CNN, using Feature form 1 based on its ability to learn the input features as a NN, was placed in the third position in terms of overall Acc. Although all methods exhibit accuracy higher than 99%,

**Table 9**
Classification overall Acc (%) of the developed supervised methods fed by dataset *X* composed from Feature forms 1 and 4 for total sag data, clean data, and noisy data.

| | Method | Feature form 1 | | | | | | Proposed Feature form 4 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Total | Clean | 40 dB | 30 dB | 20 dB | 15 dB | Total | Clean | 40 dB | 30 dB | 20 dB | 15 dB |
| 1 | DT | 94.15 | 99.37 | 98.63 | 96.73 | 90.25 | 85.75 | 98.54 | 99.98☑ | 99.79☑ | 99.44 | 97.61 | 95.89 |
| 2 | Polynomial-SVM | 97.80 | 100.00✓ | 99.86☑ | 99.44 | 96.70 | 92.99 | 99.16☑ | 100.00✓ | 99.88☑ | 99.77☑ | 98.96 | 97.17 |
| 3 | RF | 97.81 | 99.91☑ | 99.72☑ | 99.30 | 97.01 | 93.11 | 99.00 | 100.00✓ | 99.77☑ | 99.72☑ | 98.49 | 97.01 |
| 4 | KNN | 98.09 | 99.98☑ | 99.91✓ | 99.61☑ | 97.35 | 93.59 | 98.90 | 100.00✓ | 99.79☑ | 99.63 | 98.19 | 96.87 |
| 5 | EL (SVM, RF, KNN) | 98.43 | 100.00✓ | 99.86☑ | 99.63☑ | 98.03 | 94.61 | 99.07 | 100.00✓ | 99.77☑ | 99.77☑ | 98.68 | 97.12 |
| 6 | Designed CNN | 99.13✓ | 100.00✓ | 99.86☑ | 99.68✓ | 98.77✓ | 97.35✓ | 99.37✓ | 100.00✓ | 99.95✓ | 99.91✓ | 99.61✓ | 97.40✓ |

Colors: Acc sorted from lowest (dark red) to highest (dark green) for each column; ✓: Best Acc for each column, ☑: Closest Acc to best one by tolerance of (– 0.21%)

the substantial data size ($m$ = 21,540) reveals notable differences in incorrect detections, which are crucial in real-time system protection applications. Specifically, CNN-Feature form 4 shows 135 wrong detections, whereas RF has 216. In regard to the F1-Score values, all the methods demonstrated higher values for the US class, a trend also observed with k-means clustering in Table 7. This suggests that discerning patterns within the input data corresponding to the US class (label 0) were relatively more straightforward for the methods. However, it is important to note that the higher values for recall were the contributing factor, rather than precision. Additionally, CNN-Feature form 4 outperformed other methods in the F1-score for both classes. The designed CNN-Feature form 4 surpassed almost all criteria compared to the other methods, except for the Acc against Sym voltage sags, which was very similar to CNN-Feature form 1. SVM-Feature form 4 showed slightly higher performance compared to CNN-Feature form 1. The EL-Feature form 4 had slightly higher Acc (98.86%) compared to CNN-Feature form 1 (98.61%) for TE cases.

Fig. 11 illustrates the noise-front Acc of the top three methods mentioned in Table 11 across various types of voltage sags. All three methods exhibited higher Acc for US sags compared to DS sags. Across all scenarios, the consistent trend emerges that employing Feature form 4 leads to better Acc, with exceptions observed only for the Sym 20 dB sags, as well as the Sym and Earth fault 15 dB sags. This demonstrates that the high Acc obtained for CNN-Feature form 1 (99.79%) for Sym sags, as mentioned in Table 11, comes from its good performance in high-noise conditions. However, this method showed the least performance in TE 15 dB cases, achieving an Acc of 95.57%. Meanwhile, this was enhanced by SVM and the designed CNN while utilizing Feature form 4, reaching 96.84% and 97.47%, respectively. Overall, the results demonstrate the robustness of the most accurate methods, specifically CNN-Feature form 4, even in scenarios involving voltage sags with substantial noise levels.

### 4.3. Testing the most accurate methods and k-means using measurements in areal electrical power system

This section aims to validate the performance of the most accurate supervised methods, along with k-means clustering, in addressing the VSSL problem through field test cases in a real electrical power system. As outlined in Table 12, we conducted 12 field tests to evaluate the methods. Among these, cases 1 to 6 involve measurements from naturally occurring voltage sags within the Slovenian electrical transmission network (Polajžer et al., 2015a) at different voltage levels recorded by protection relays. The other cases relate to a 20 kV Slovenian electrical distribution network (Fig. 12) with a loop topology (Polajžer et al., 2019), where voltage sags were induced intentionally by the network operator, while currents and voltages were recorded by over-current relays 1 to 4 (R1 to R4).

Based on the information provided in Fig. 1c regarding the deployment of methods in real-time applications, some pre-processing steps are necessary before prediction of the real-world samples. The first step involves resampling waveforms. It's worth noting that certain test cases were recorded at a higher number of samples per cycle than the agreed-upon $N = 128$ in this study, specifically cases 1 to 3. Consequently, these cases required down-sampling. Conversely, the remaining cases, except for test 6, had fewer samples per cycle, and thus needed to be up-sampled. The authors utilized the *resample* function in MATLAB to adjust the sampling rate to 128. The next step entails extracting two cycles prior to sag and three cycles during sag, integrating 128 samples per cycle for all 12 cases. The third step involved applying the DFT algorithm, and in the fourth step, Feature forms 1 and 4 are produced. The last preprocessing step involves normalizing the cases using the same normalization parameters saved during the training phase of the models.

As an illustration, cases 6 and 11 representing DS and US scenarios, respectively, are depicted in Fig. 13, which includes input current and voltage waveforms along with the generated Feature forms 1 and 4. The prediction results of all 12 field-testing cases, following the structure in Fig. 1c, are detailed in Table 12. In case 6, an evolving voltage sag is observed, where a TE event was followed by an LL fault at $t$ = 0 ms, subsequently tracked by an LLL fault.

As outlined in Table 13, the k-means clustering utilizing Feature form 4, as an unsupervised method, demonstrated commendable performance, predicting 10 out of the 12 cases correctly. This observation provides additional evidence of the true confirmation of Feature form 4's influential nature by this method. As anticipated, the designed CNN-Feature form 4 exhibits the best performance, predicting all 12 cases accurately. EL-Feature form 4 and CNN-Feature form 1 both follow closely with 10 correct predictions. SVM and RF-Feature form 4 were placed in the next tier with 9 and 8 accurate predictions, respectively. The outcomes of applying these methods to real-world field test cases affirmed that the methods, which displayed high Acc on the datasets used in this study, are also effective in real-world applications. Section 5.4 will provide further explanations regarding this.

## 5. Discussion

### 5.1. Comparison with previous methods

Table 14 provides a comprehensive comparison between the proposed methods and previous ones, both ML-based and analytical, tested within the same case study (Section 3). The comparison involves overall Acc, excluding 15 dB SNRs, which is considered a high-level noise, and highlights the types of features utilized. The evaluated features include TSB Feature forms 1 and 4 from this study, a subset of scalar features derived from phasor or instantaneous components, a discrete version of Feature forms 3 and 4 discussed in this paper, and a discrete phasor-based approach raised up in this study for VSSL.

Among the five highest Acc methods, the most accurate TSB ML methods identified in this paper are observed: CNN-Feature form 4 (99.87%), SVM-Polynomial-Feature form 4 (99.65%), CNN-Feature

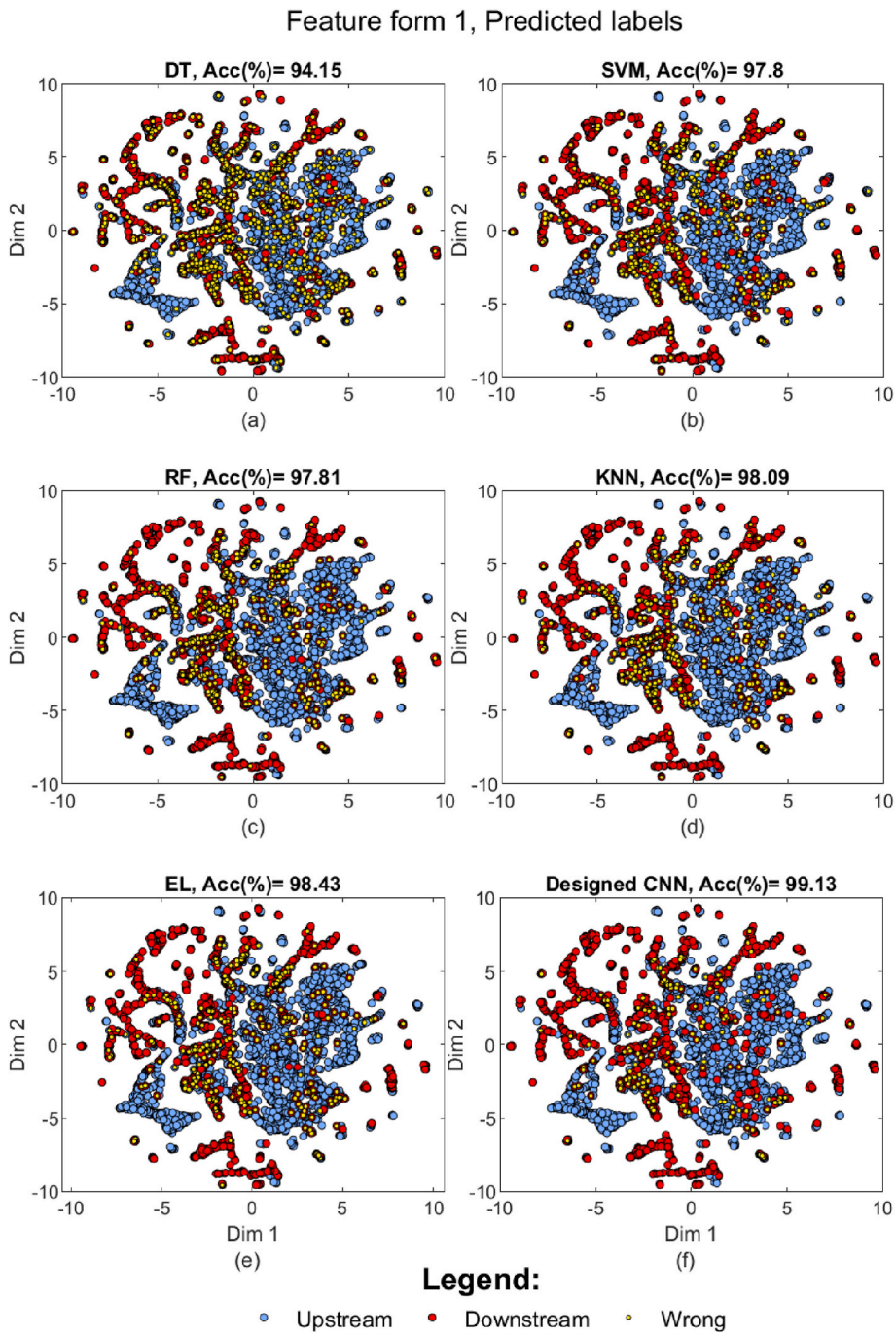


**Fig. 7.** Supervised methods' scatter plot for dataset *X* composed from Feature form 1; (a) DT; (b) SVM; (c) RF; (d) KNN; e (EL); (f) 1D-CNN. The samples' predicted labels are marked with blue points (US), red points (DS), and yellow points (Wrong). The overall Acc (%) for each method is given at the top of each Figure.

form 1 (99.58%), EL Feature form 4 (99.55%), and RF-Feature form 4 (99.49%). These results showcase the high performance of these identified methods, even when compared to previous methods in the literature. Looking at the methods placed in levels 6 to 13 in Table 14 shows that even the TSB ML methods utilizing Feature form 1 outperformed the ML methods employing a discrete subset of features, addressing the research gap identified in this paper.

The three least accurate methods include KNN using a subset of instantaneous-based features (Mohammadi et al., 2022b) (91.32%), the CBM analytical method employing the basic rule of changes in positive sequence current magnitude and angle (Moradi and Mohammadi, 2012; Mohammadi et al., 2021, 2022b) (90.55%), and the RCC analytical method using a discrete form of Feature form 3 (Mohammadi et al., 2017a, 2021, 2022b; Hamzah et al., 2004) (89.84%). This generally demonstrates that ML-based methods exhibit a satisfactory performance when compared to analytical approaches.

The authors of this study also evaluated the Acc of an analytical method while using a discrete type of Feature form 4 as, $\Delta(|I^{+}|\sin(\theta_{I^{+}})) < 0$ for DS voltage sags, yielding an Acc of 92.16% (27th method in Table 14). This highlights that Feature form 4 might not be suitable as an analytical rule. Instead, it proved to be a valuable TSB feature for the ML methods developed in this paper. Notably, the highest Acc was achieved by the designed CNN at 99.87%.

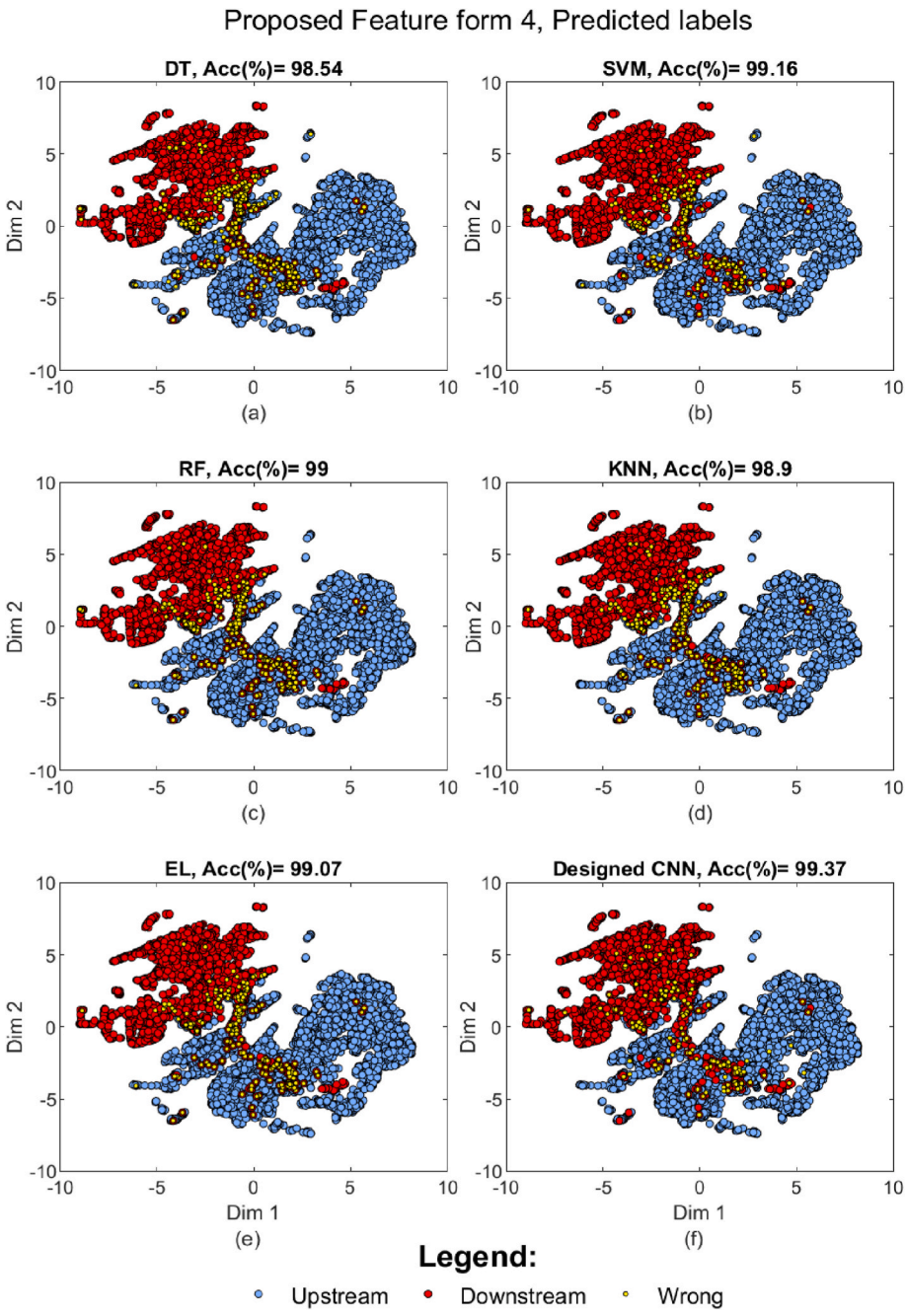


**Fig. 8.** Supervised methods' scatter plot for dataset *X* composed from Feature form 4; (a) DT; (b) SVM; (c) RF; (d) KNN; e (EL); (f) 1D-CNN. The samples' predicted labels are marked with blue points (US), red points (DS), and yellow points (Wrong). The overall Acc (%) for each method is given at the top of each Figure.

### *5.2. Performance of the most accurate TSB methods in terms of prediction time*

Similar to various types of studies, the TSB methods developed in this paper were initially trained offline. Once their parameters are fine-tuned, they will be put into practice for online real-world applications (Fig. 1c). This way, the training phase of these methods will not affect their speed during real-time use. Instead, what will matter is the time taken for predictions and preparing the input feature forms. Unlike our previous study (Mohammadi et al., 2022b), which used a subset of scalar features that could be time-consuming, this work utilizes TSB feature forms containing magnitudes and angles of currents or voltages. This type of information is highly likely to be available directly from PQMs. The only remaining aspect is the evolution in the form of Feature form 4 (the most influential form).

Consequently, we have calculated the prediction time for the most accurate methods, and the results are detailed in Table 15. The computation time was determined using an Apple M1 with 8 GB RAM memory. The computed time was derived from the prediction average time of an actual voltage sag sample, which was detailed in Table 12. It is reported for each of the 5 folds/models per each method. The time for each fold was almost the same, and the longest time among the folds is recorded as the method's prediction time. The designed CNN-Feature form 1 took more time compared to those utilizing Feature form 4. The reason for this lies in the size difference between Feature forms 1 and 4, and the higher complexity of the CNN-Feature form 1 architecture. Specifically, Feature form 1 boasts

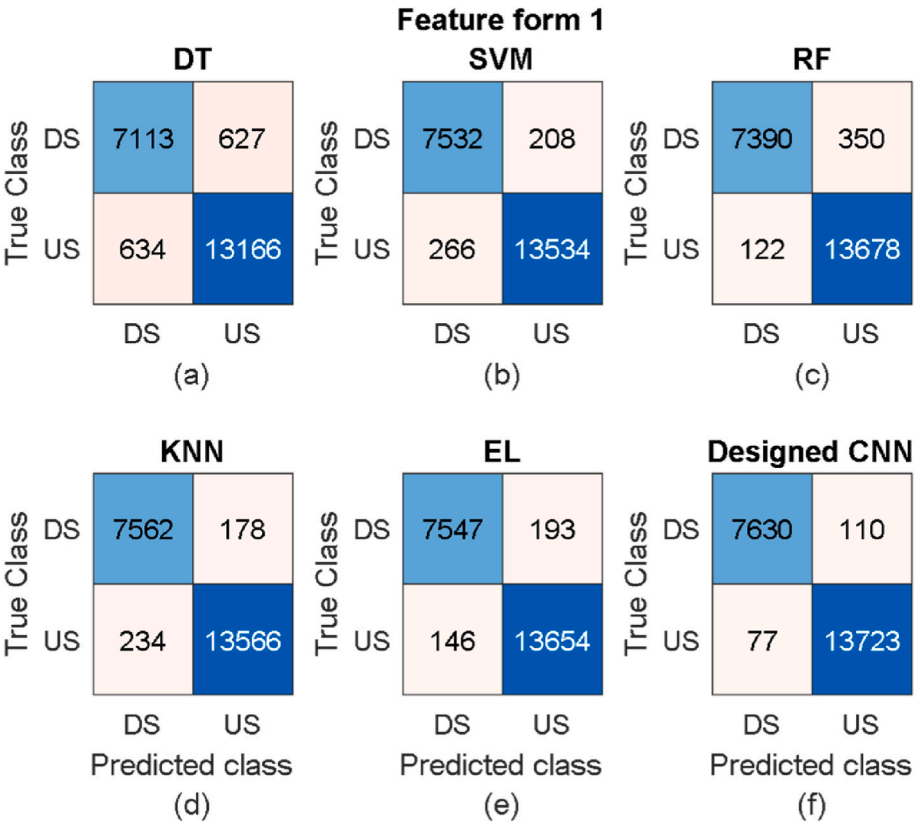


**Fig. 9.** Confusion matrix of the methods-Feature from 1 (total sag data).

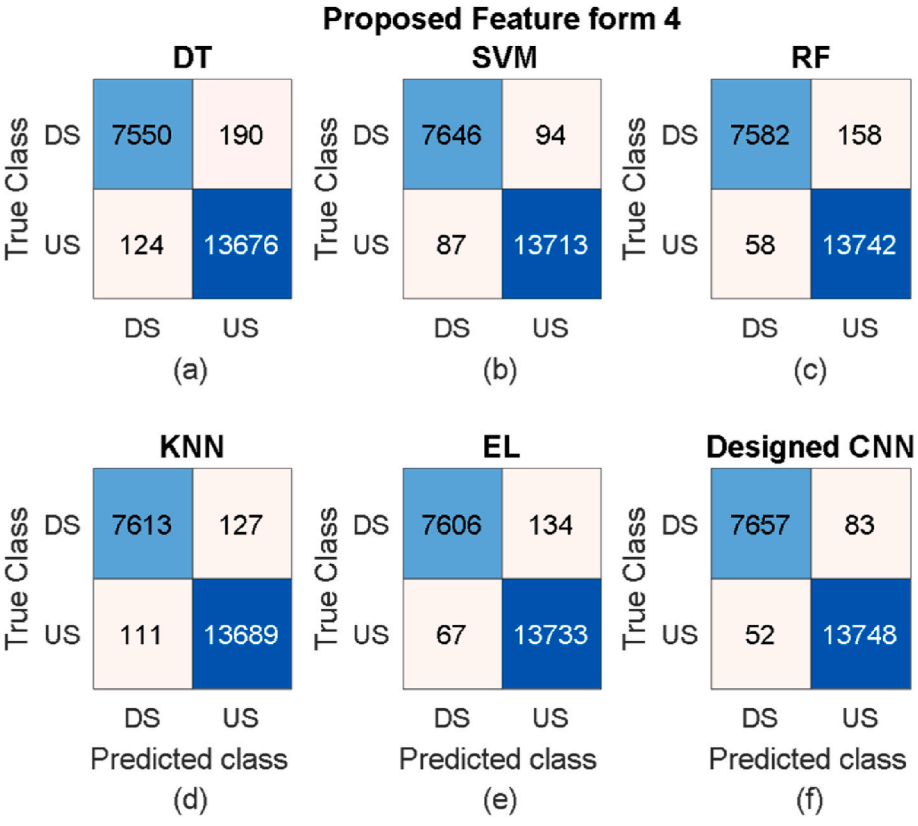


**Fig. 10.** Confusion matrix of the methods-Feature from 4 (total sag data).

**Table 10**
The prediction of supervised methods against 8 random voltage-sag samples, in which most of the methods were wrong at least once. Underlines refer to wrong predictions.

| No. | Sag source-PQM | SNR (dB) | Actual labels | Feature form 1 | | | | | | Proposed Feature form 4 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | DT | SVM | KNN | RF | EL | Designed CNN | DT | SVM | KNN | RF | EL | Designed CNN |
| 1 | F5, $LE_b$, $40^{\Omega}$ - PQM1 | 20 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 2 | F10, $LLE_{ca}$, $40^{\Omega}$ - PQM1 | 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 3 | F13, $LL_{bc}$, $1^{\Omega}$ - PQM5 | 30 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 |
| 4 | F13, $LLE_{ab}$, $1^{\Omega}$ - PQM4 | 15 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| 5 | F13, LLLE, $1^{\Omega}$ - PQM1 | 15 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 6 | F14, $LLE_{bc}$, $10^{\Omega}$ - PQM1 | 20 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 7 | TE3, $120^{MVA}$ - PQM2 | 20 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 8 | TE3, $120^{MVA}$ -PQM2 | 15 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 |

**Table 11**
A comparison between the most accurate supervised methods for total sag data against different kinds of sags.

| Method | Wrong detection out of *m* | Acc for all sags (%) | Acc for Sym sags (%) | Acc for Asym sags (%) | Acc for Earth faults (%) | Acc for TE (%) | DS sags | | | US sags | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Recall (%) | Precision (%) | F1-Score (%) | Recall (%) | Precision (%) | F1-Score (%) |
| Designed CNN-Feature form 4 | 135 | 99.37 | 99.73 | 99.29 | 99.28 | 99.37 | 98.93 | 99.93 | 99.13 | 99.62 | 99.4 | 99.51 |
| SVM-Feature form 4 | 181 | 99.16 | 99.71 | 99.04 | 99.07 | 99.24 | 98.79 | 98.87 | 98.83 | 99.37 | 99.32 | 99.34 |
| Designed CNN-Feature form 1 | 187 | 99.13 | 99.79 | 99.01 | 99.07 | 98.61 | 98.58 | 99.9 | 98.79 | 99.44 | 99.2 | 99.32 |
| EL-Feature form 4 | 201 | 99.07 | 99.73 | 98.93 | 98.9 | 98.86 | 98.27 | 99.13 | 98.7 | 99.51 | 99.03 | 99.27 |
| RF-Feature form 4 | 216 | 99 | 99.71 | 98.9 | 98.88 | 97.85 | 97.96 | 99.24 | 98.6 | 99.58 | 98.86 | 99.22 |

Colors: values are sorted from highest (dark green) to lowest (dark yellow) for each column.

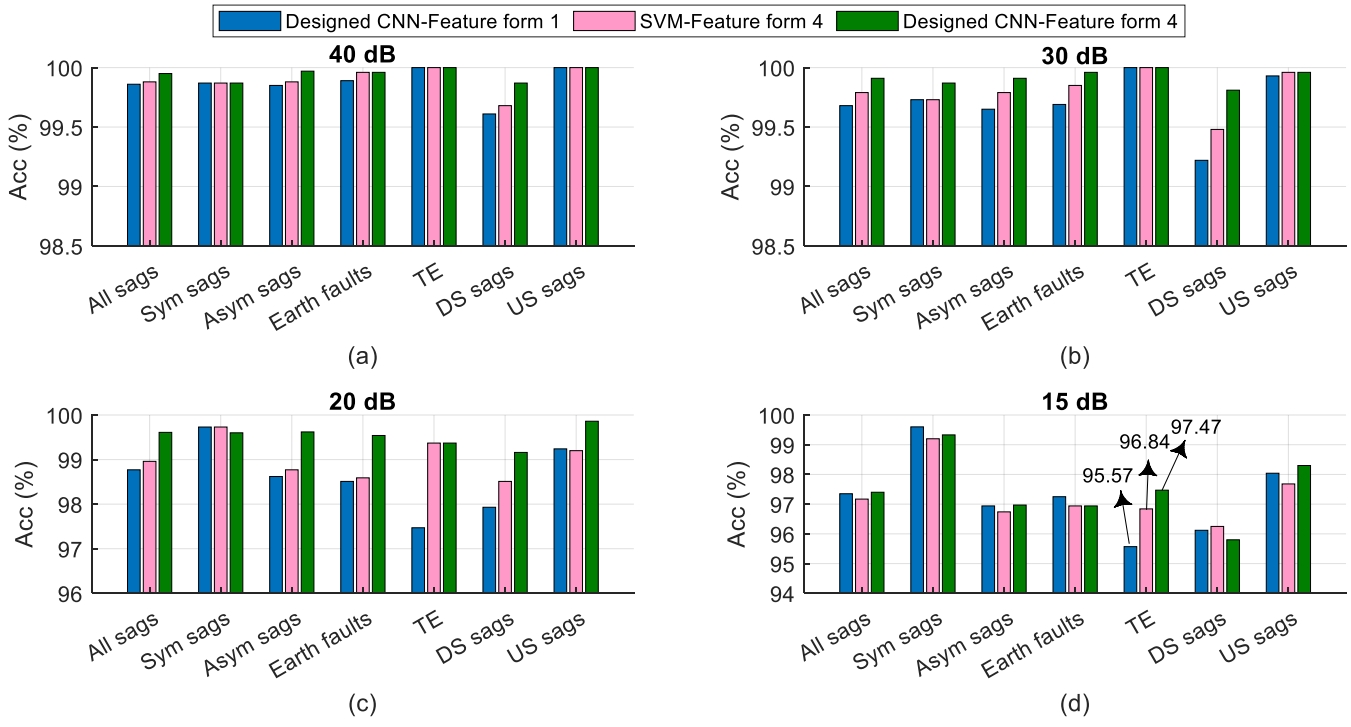


**Fig. 11.** Acc (%) of the three top methods against different kinds of sags for noisy data (a) 40 dB data; (b) 30 dB data; (c) 20 dB data; (d) 15 dB data.

a size of 3,840, which is six times larger than that of Feature form 4. SVM, RF, EL, and the designed CNN, all using Feature form 4, are sorted from minimum prediction time to the highest before CNN-Feature form 1.

However, it is worth noting that all prediction times fell within the range of 0.05 cycles to 1.41 cycles (at 50 Hz), indicating a swift process. Specifically, CNN-Feature form 4, which attained the highest Acc, required 28.17 ms to predict a single sample. This equates to approximately 1.41 cycles at 50 Hz or 1.69 cycles at 60 Hz. This rapid prediction time underscores the feasibility of implementing this high-performance method in real-time practical scenarios, such as serving as a directional function within protection relays positioned across distribution and transmission systems.

**Table 12**
Field-testing cases in Slovenian electrical power systems.

| Test no. | Actual labels | Sag source | Measurement device-voltage level (kV) | Sampling frequency (kHz) | Sample per cycle | Sag duration (s) | Sag magnitude (pu) |
|---|---|---|---|---|---|---|---|
| 1 | 1 | $LE_a$ | 20 | 10 | 200 | 0.34 | 0.22 |
| 2 | 1 | $LL_{bc}$ | 20 | 10 | 200 | 0.9 | 0.16 |
| 3 | 1 | LLL | 20 | 10 | 200 | 0.9 | 0.15 |
| 4 | 1 | $LE_a$ | 400 | 1 | 20 | 0.65 | 0.3 |
| 5 | 0 | $LE_c$ | 220 | 1 | 20 | 0.35 | 0.06 |
| 6* | 1 | TE + LL-ab + LLL | 20 | 6.4 | 128 | 0.04 + 0.04+0.03 | 0.02 |
| 7 | 1 | $LE_a$ | R1-20 | 1 | 20 | 0.09 | 0.1 |
| 8 | 1 | $LL_{ac}$ | R1-20 | 1 | 20 | 0.09 | 0.64 |
| 9 | 0 | $LE_a$ | R2-20 | 1 | 20 | 0.09 | 0.15 |
| 10 | 0 | $LL_{ac}$ | R2-20 | 1 | 20 | 0.09 | 0.47 |
| 11* | 0 | $LE_c$ | R3-20 | 1 | 20 | 0.09 | 0.05 |
| 12 | 0 | $LE_c$ | R4-20 | 1 | 20 | 0.09 | 0.06 |

Cases 1 to 6: Measurements recorded due to natural sags which occurred in networks, Case 7: a TE developed by fault, Cases 7 to 12: Measurements recorded due to sags applied by the operators in the looped network shown in Fig. 12, Cases 7 and 9: the substation transformer is grounded by a Petersen coil, Cases 8 and 12: substation transformer is grounded by 80-Ω resistance, *: Cases shown in Fig. 13.

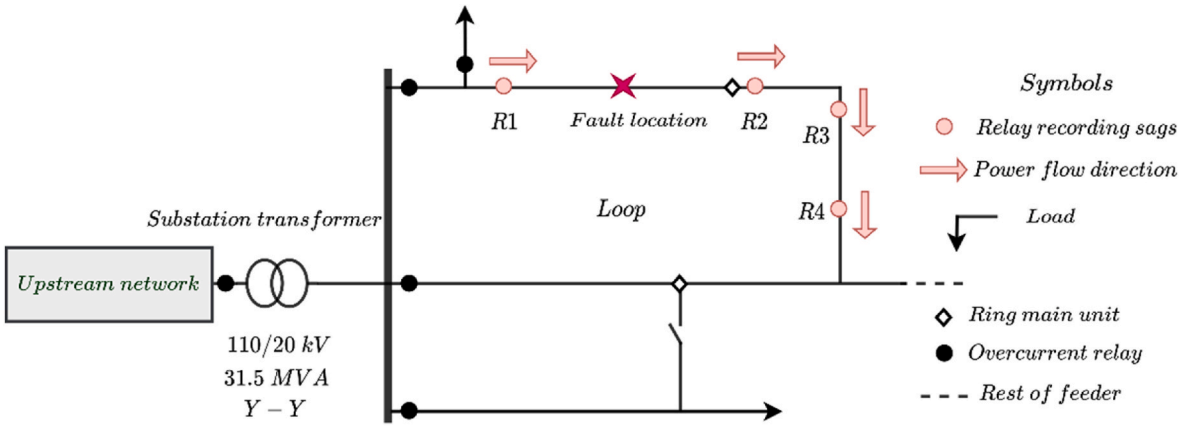


**Fig. 12.** Part of the 20 kV Slovenian electrical distribution network.

### 5.3. Original findings, strengths, challenges, and selection of methods

#### 5.3.1. Original findings and strengths

Overall, the research contributes to an enhanced approach for VSSL, a critical field in electrical power quality and system protection within electrical engineering. A primary objective of this research has been to identify the most suitable Feature form for VSSL assisted by ML. The identification of Feature form 4 as the most prominent addresses the challenge of selecting an optimal set of input features impacting traditional ML methods. In a comparative study, for instance, combining SVM with Feature form 4 yielded a high accuracy of 99.16%. Furthermore, the selection of Feature form 4 was guided by certain physical concepts distinct from conventional feature engineering.

Compared to Feature form 1, Feature form 4 exhibited lower complexity, reduced signal variation, and a shorter length, approximately 17% of Feature form 1's length. These characteristics enable the models, particularly CNNs, to learn meaningful features with simpler architectures and track the shape of the Feature form more effectively over a sequence of time-series. Additionally, it requires less effort from the CNN during the training phase, facilitating its utilization in real-time applications.

However, the designed CNN demonstrated impressive accuracy when utilizing both Feature forms 1 and 4, leveraging its ability to learn feature representations from input data. Nonetheless, the combination of CNN with Feature form 4 exhibited true detection of 49 additional voltage sag samples, which is particularly valuable in online applications, such as operation as part of a protection device (as explained in Section 1.3), compared to Feature form 1, with using a smaller dataset in terms of feature instances.

Initially utilizing k-means clustering as a simple clustering method demonstrated its effectiveness in distinguishing between two classes, particularly when utilizing Feature form 4. Surprisingly, even in scenarios with high levels of noise (15 dB and 30 dB), its performance surpassed that of the supervised DT method when using Feature form 1. Furthermore, k-means with Feature form 4 exhibited a performance on par with (EL with Feature form 4, CNN with Feature form 1), or superior to (RF and SVM with Feature form 4), some of the most accurate supervised methods when tested on real-world voltage sags (refer to Table 13). The capabilities of k-means as an unsupervised method without relying on input labels bring the idea to support it by adding some feature size reduction methods to enhance its success, as discussed in Section 5.9.

#### 5.3.2. Challenges or issues

1. **Long training time**: Both traditional ML methods and the 1D CNN exhibited lengthy training durations, with CNN being much more time-consuming. For instance, SVM-Feature form 4 required 43.48 min for training. This extended training period could pose a challenge in scenarios necessitating online training to adjust parameters based on new data. However, since system parameters change gradually over time, as discussed in Section 5.5, implementing techniques like transfer learning may prove effective.
2. **Trust in the methods**: Despite achieving accuracy above 99% using traditional ML methods in this study for VSSL, reported in Tables 11

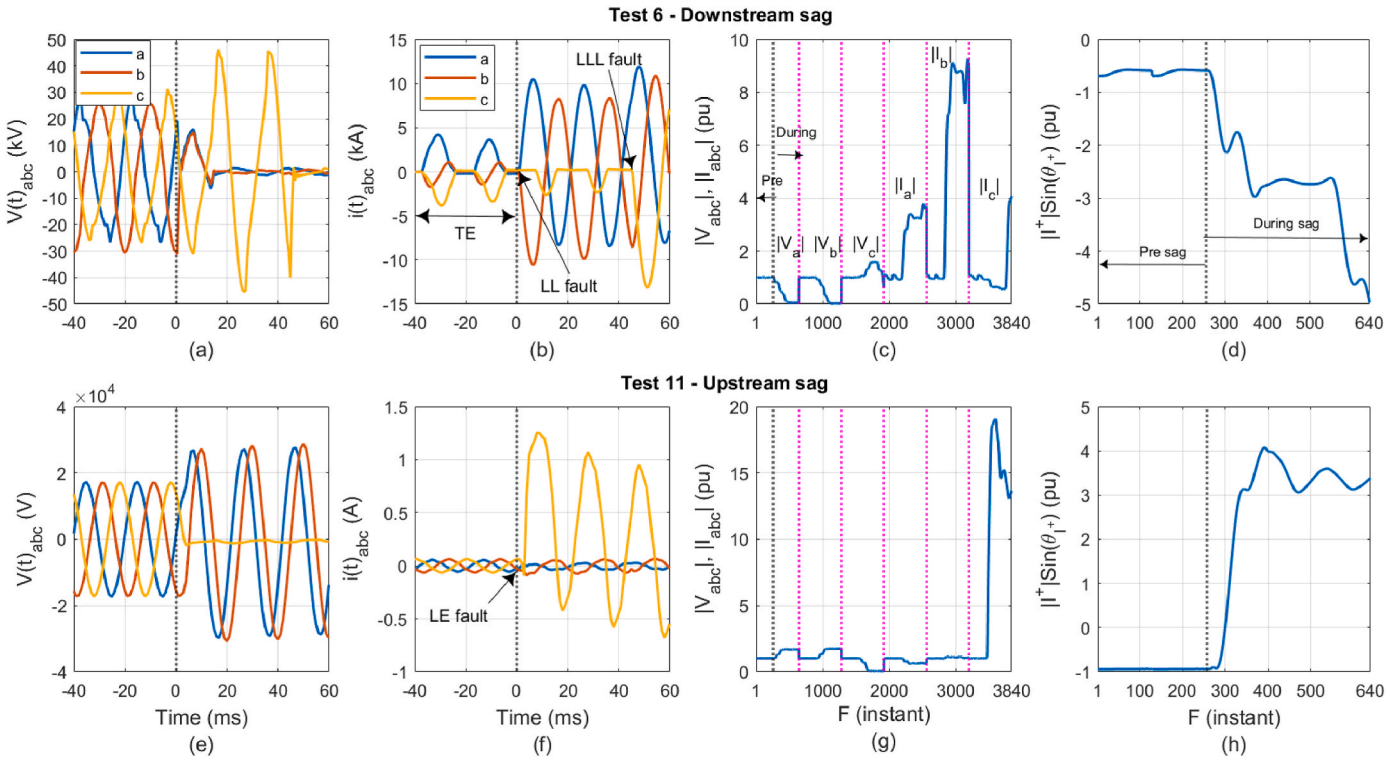


**Fig. 13.** Field testing 6 (a–d) and 11 (e–h), explained in Table 12. (a, e) Three phase voltages; (b, f) Line currents; (c, g) evolution of Feature form 1; (d, h) Evolution of Feature form 4. To retain the true scale of values, voltages and currents are displayed here in their original units. However, the Feature forms are considered in per unit (PU) for consistency.

**Table 13**
Field-testing prediction results for the most accurate methods and k-means. The blue-colored numbers refer to correct predictions.

| Test no. | Actual labels | Designed CNN-Feature form 1 | k-means-Feature form 4 | RF-Feature form 4 | EL-Feature form 4 | SVM-Feature form 4 | Designed CNN-Feature form 4 |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| 2 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 3 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 4 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| 5 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 6* | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 7 | 1 | 1 | 1 | 0 | 1 | 0 | 1 |
| 8 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| 9 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11* | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 12 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| True detection out of 12 | | 10 | 10 | 8 | 10 | 9 | 12 |

*: Cases shown in Fig 12.

and in online applications such as operating as part of a protection device in electrical power systems, there remains doubt and challenge for electrical engineers to trust and implement such intelligent methods rather than analytical methods based on electrical rules.

3. **Generalization issues**: Despite efforts to reduce the overfitting risk, such as normalization and employing cross-validation and preparing a sufficiently large number of data samples, traditional ML methods and even the 1D CNN with high accuracy might struggle to generalize well to unseen data. However, Table 13 showed the good performance of the most effective ML methods against new unseen voltage sag samples.
4. **Complexity and interpretability**: Despite achieving very high accuracy for SVM and 1D CNN with Feature form 4, they could be more complex than other investigated methods regarding interpretability and providing insights into the decision-making process to understand why certain predictions are made.
5. **Feature engineering**: The necessity of manual feature engineering for traditional ML methods remains a challenge and weakness. However, the employed TSB Feature forms in this study, specially

**Table 14**
A comparison of the overall Acc of discussed supervised methods in this study with the previous ones.

| No. | Method | Used feature types | Publication | Acc (%) |
|---|---|---|---|---|
| 1 | Designed CNN | TSB, form 4 | Current study | 99.87 |
| 2 | Polynomial-SVM | TSB, form 4 | Current study | 99.65 |
| 3 | Designed CNN | TSB, form 1 | Current study | 99.58 |
| 4 | EL (SVM[1], RF, KNN) | TSB, form 4 | Current study | 99.55 |
| 5 | RF | TSB, form 4 | Current study | 99.49 |
| 6 | RF | A discrete subset of phasor and instantaneous-based ones | [64] | 99.48 |
| 7 | RF | A discrete subset of phasor-based ones | [64] | 99.46 |
| 8 | KNN ($K = 7$) | TSB, form 4 | Current study | 99.4 |
| 9 | EL (SVM[1], RF, KNN) | TSB, form 1 | Current study | 99.38 |
| 10 | KNN ($K = 3$) | TSB, form 1 | Current study | 99.21 |
| 11 | DT | TSB, form 4 | Current study | 99.2 |
| 12 | Polynomial-SVM | TSB, form 1 | Current study | 99 |
| 13 | RF | TSB, form 1 | Current study | 98.98 |
| 14 | EL (SVM[2], LR, DT, RF, KNN) | A discrete subset of phasor-based ones | [64] | 98.85 |
| 15 | KNN ($K = 5$) | A discrete subset of phasor-based ones | [64] | 98.73 |
| 16 | EL (SVM[2], LR, DT, RF, KNN) | A discrete subset of phasor and instantaneous-based ones | [64] | 98.71 |
| 17 | DT | A discrete subset of phasor and instantaneous-based ones | [64] | 98.59 |
| 18 | DT | A discrete subset of phasor and instantaneous-based ones | [64] | 98.53 |
| 19 | LR | A discrete subset of phasor and instantaneous-based ones | [64] | 98.45 |
| 20 | KNN ($K = 2$) | A discrete subset of phasor and instantaneous-based ones | [64] | 98.41 |
| 21 | LR | A discrete subset of phasor-based ones | [64] | 98.3 |
| 22 | RF | A discrete subset of instantaneous-based ones | [64] | 97.43 |
| 23 | DT | TSB, form 1 | Current study | 96.24 |
| 24 | LR | A discrete subset of instantaneous-based ones | [64] | 95.56 |
| 25 | DT | A discrete subset of instantaneous-based ones | [64] | 95.17 |
| 26 | EL (SVM[2], LR, DT, RF, KNN) | A discrete subset of instantaneous-based ones | [64] | 93.77 |
| 27 | Analytical method as $\Delta^3(\lvert I^+\rvert \sin(\theta_{I^+})) < 0$ for DS sags else US | A discrete kind of form 4 | Current study | 92.16 |
| 28 | KNN ($K = 2$) | A discrete subset of instantaneous-based ones | [64] | 91.32 |
| 29 | CBM analytical method as $\Delta(\lvert I^+\rvert) > 0$ & $\Delta(\theta_{I^+}) < 0$ for DS sags else US | A discrete phasor-based one | [37,62,64] | 90.55 |
| 30 | RCC analytical method as $\Delta(\lvert I^+\rvert \cos(\theta^+)) > 0$ for DS sags, else US | A discrete kind of form 3 | [22,26,62,64] | 89.84 |

1: Polynomial kernel, 2: Polynomial, Sigmoid, Linear, RBF, 3: $\Delta = (.)_{during-dip} - (.)_{pre-dip}$ , Colors: accuracies sorted from highest (dark green) to lowest (dark red).

**Table 15**
Average prediction time of the most accurate TSB methods for a real-world voltage sag sample, discussed in Section 4.3.

| Method | fold 1 | fold 2 | fold 3 | fold 4 | fold 5 | max time (msec) | Cycle (60 Hz) | Cycle (50 Hz) |
|---|---|---|---|---|---|---|---|---|
| Designed CNN-Feature form 4 | 27.94 | 27.53 | 27.05 | 28.17 | 27.94 | 28.17 | 1.69 | 1.41 |
| SVM-Feature form 4 | 0.77 | 0.91 | 0.77 | 0.76 | 0.76 | 0.91 | 0.05 | 0.05 |
| Designed CNN-Feature form 1 | 30.14 | 28.64 | 28.60 | 28.60 | 28.60 | 30.14 | 1.81 | 1.51 |
| EL-Feature form 4 | 10.36 | 10.65 | 10.88 | 11.32 | 11.55 | 11.55 | 0.69 | 0.58 |
| RF-Feature form 4 | 3.37 | 3.12 | 3.14 | 3.13 | 3.21 | 3.37 | 0.2 | 0.17 |

Colors: values are sorted from the lowest (dark green) to the highest (dark yellow) for the last three columns.

form 4, helped the methods achieve high accuracy. This is attributed to the relatively simpler tracking of the shape of the Feature form across a sequence of time-series data.

6. **Data requirement for DL models**: While the amount of data used in this study was sufficient for the designed 1D CNN, employing real-time voltage sag samples to train such DL models always requires a large amount of labeled data, presenting a challenge.
7. **Computational expense**: The designed 1D CNN could be computationally expensive to train compared to traditional ML methods. We also observed this issue while training our large dataset, and found that 8 GB RAM was not enough, necessitating a powerful GPU.
8. **Hyperparameter tuning**: Both traditional ML methods and the 1D CNN were trained with an optimal selection of hyperparameters, as reported in Table 8. However, how to tune the hyperparameters can affect the methods' accuracy directly.

#### 5.3.3. *Selection of methods*

The most accurate methods obtained from this study, as reported in Table 11, all demonstrated accuracy levels above 99%. However, it is

imperative to consider not only accuracy but also the complexity and real-time efficiency of each model. In our evaluation, we compared the 1D CNN-Feature form 4 with other models, assessing metrics such as model size, inference time, and computational resources required. Our findings revealed that while the CNN model achieved high accuracy, it typically incurred longer prediction times compared to traditional ML methods (as indicated in Table 15). Additionally, due to its neural network architecture, the CNN model could exhibit higher complexity and computational expense than traditional ML methods (as discussed in the previous Section). Achieving a balance between model selection and real-time requirements is essential in practical applications. Strategies such as employing model optimization techniques or considering trade-offs between accuracy and efficiency are crucial to meet the specific needs of the application while ensuring real-time performance. These factors should be weighed carefully when making decisions about model selection for deployment in real-world scenarios (see Fig. 1c). An obvious choice would be CNN-Feature form 4, given its higher accuracy, lower model complexity, and faster prediction time compared to CNN-Feature form 1.

### 5.4. Influence of real-world voltage sags

Utilizing real-world voltage sag samples, exemplified by the 12 test cases outlined in Table 12, for dataset generation is undoubtedly beneficial. Due to the challenges which emerged, explained in Section 3, the authors undertook the well-known case study shown in Fig. 3, an actual Brazilian electrical network (60 Hz) with updated parameters. This case study holds a respected position in the existing body of knowledge as employed in the studies such as (Moradi et al., 2012; Moradi and Mohammadi, 2012, 2013; Mohammadi et al., 2017b, 2022a), and the dataset obtained from the real system is highly trustworthy and dependable (see Appendix A for more information). The methods deliberated upon in this paper were trained and tested using this dataset. Moreover, putting the most-accurate identified methods to the test with 12 unseen real-world voltage sags derived from the Slovenian electrical power systems (50 Hz) revealed that these methods indeed exhibit tangible efficacy in real-world scenarios. Furthermore, the results shown in Table 13 highlight the advantages of using a wide range of real-world voltage sags from different places and re-training the methods thoroughly to get robust performance for the methods when dealing with real voltage sags. However, assigning labels to the data could still be a challenging task. In general, in practical scenarios, real-world voltage sags, encompassing both voltage and current waveforms, are susceptible to noise, measurement inaccuracies, and indeterminate variations attributed to varying power system fault conditions. Consequently, the Feature forms under scrutiny in this study may encounter some degree of influence. Recognizing this potential concern, this paper placed substantial emphasis on accounting for the effects of noise, specifically at 40 dB, 30 dB, 20 dB, and 15 dB.

### 5.5. Generalization and scalability of the developed methods across different electrical power systems or configurations

Typically, to create scalable models that can handle new and unfamiliar situations effectively, training data needs to encompass all the complexities of a problem, including various scenarios and potential parameter changes. However, in this study, while forming the datasets, efforts were made to encompass a wide range of potential electrical power system scenarios. These included changes in electrical load sizes, variations in power line impedances, different power system typologies (such as the placement of PQMs at various points), and various combinations of power system faults and TEs. Furthermore, the Feature form that was identified as the most influential in this study, Feature form 4, involves both the magnitude and angle components of the current. By normalizing the data both across rows and columns, any changes in the parameters of the case study are anticipated to impact the evolution of this Feature form uniformly. Consequently, any adjustments made to the system's parameters are not foreseen to have a substantial impact on the models' performance when encountering newly generated voltage sags resulting from parameter changes within the electrical power system.

One concern regarding the developed TSB methods is how well they work when tested on a different electrical power system. Usually, the models need to be re-trained in order to get a similar Acc. However, the models related to the most accurate methods, might not need to be changed. Besides, to avoid re-training, we can use the active learning and fine-tuning of pre-trained models using certain new samples from the new different electrical power system. Transfer learning (Li et al., 2022) is one way to do this. A study in (Li et al., 2023c) similarly employed transfer learning for the VSSL problem by initially adjusting model parameters using simulation data, followed by fine-tuning the pre-trained model using a limited set of measurement data.

### 5.6. Choosing the cycles before and during a voltage sag

The selection of two cycles before the voltage sag event was based on previous research (Khosravi et al., 2007, 2008) and the fact that the learning methods can detect occurred changes effectively in the time series composed of Feature forms 1 to 4 caused by voltage sags. However, even one, or maybe a half-cycle before the voltage sag, could be sufficient for the analysis. Regarding the selection of three cycles during the voltage sag, previous studies (Mohammadi et al., 2022a) suggested that the transient period, occurring half or one cycle after the voltage sag, provides valuable information regarding the voltage sag source direction. However, the steady-state period after the voltage sag depends on the power system fault period and the TE duration, which may vary. In the case of extracting voltage sag information from protection relays, where power system faults are the voltage sag sources, typically, these devices offer a pre-fault period of a few cycles as part of their fault capture mechanism. Additionally, many protection devices do not operate immediately upon detecting the first fault cycle. Instead, they may delay action until the third cycle from the fault initiation, ensuring sufficient waveform capture under fault conditions (Ye et al., 2023). Therefore, in this study, three cycles during the voltage sag were considered, and the results have proven to be satisfactory.

### 5.7. Choosing the number of samples(instances) per cycle

In the framework experiment discussed in Section 2, the number of samples per cycle for the recorded voltage sag examples was set to $N = 128$ samples/cycle, consistent with previous works (Khosravi et al., 2007, 2008; Meléndez et al., 2008). The effectiveness of the most accurate methods was validated in field-testing (Section 4.3), wherein $N$ was varied across different values and ultimately standardized at 128. This sampling rate is considered acceptable in modern protective relays (Saleh et al., 2023), and falls within other acceptable ranges. Notably, as mentioned in (Ye et al., 2023), WMUs also operate at higher values, such as 256 samples per cycle. However, real-world scenarios, as explored in Section 4.3, may involve different values for $N$. When $N$ is higher or lower than 128, down sampling or up sampling is necessary to extract the required voltage sag cycles, as illustrated in Fig. 1c. Moreover, the methods developed in this study, particularly those deemed most accurate, exhibited adaptability to diverse $N$ values, including 64 and 32 samples. To illustrate, the SVM-Feature form 4 underwent retraining with these alternative $N$ values, leading to a reduction in the value of $n = f$ down to 320 and 160 during the preparation of Feature form 4. Remarkably, the results demonstrated consistent accuracy, with an overall Acc of 99.16% for $N = 64$ and nearly identical Acc of 99.13% for $N = 32$. In summary, while the choice of $N = 128$ in this study is considered prudent, it is not obligatory. The methods developed herein showcase versatility, accommodating different $N$ ranges readily, providing flexibility for practical applications.

### 5.8. Computing the components of the Feature forms 1 and 4

As mentioned in Section 2.1, part II, a one-cycle running window DFT was employed on the fundamental component of voltage and current signals, to calculate the primary components for all four Feature forms. These primary components encompass specifically the RMS values for Feature form 1; and the magnitude and angle of the current phasor (positive sequence) for Feature form 4. Protection relays, fault recorders or some of the PQMs, are designed to generate RMS and positive-sequence values at a high sampling frequency. In the case of WMUs, they produce raw waveforms and so a DFT need to be applied afterwards. Consequently, producing these TSB features is straightforward and even already available as recorded measurements. It is worth noticing that alternative phasor extraction methods, such as signal least-square (LSQ) (Lobos et al., 2001), could also be implemented. This variation might lead to different observations of transients within the initial cycle following voltage sags. Nevertheless, in practical scenarios within the PQMs, only a specific method will likely be adopted. Nonetheless, as part of future work, introducing a subset of voltage sag data involving diverse phasor extraction methods to the primary dataset could enhance the robustness of the most identified accurate methods in this study.

### 5.9. Observation of high accuracy for the developed supervised ML/DL methods

As depicted in Table 11, the most accurate methods achieved an Acc above 99%. This remarkable Acc is attributed to the effectiveness of Feature form 4, evidenced by the k-means clustering achieving approximately 94% accuracy, as illustrated in Fig. 6d. Consequently, it is believed that the non-complex nature of Feature form 4 has facilitated the VSSL problem for the ML methods to learn, as other types of input features applied to ML exhibited high accuracy, but remained below the 99% (see Table 14). Initially, the dataset used in this study consisted of a clean subset comprising 4308 voltage sag samples (see Table 6), which was later expanded by incorporating various SNRs to yield a total dataset size of $m = 5 \times 4308 = 21{,}540$ voltage sag samples. The objective was to develop a comprehensive model robust against measurement errors, suitable for real-world implementation (see Fig. 1c). Thus, validation accuracies were reported during the cross-validation of the models. Furthermore, the performance of the most accurate methods was tested on a few unseen real-world voltage sag samples, yielding high performance as reported in Table 13.

However, it is well-understood in ML that achieving high Acc, especially above 99%, can sometimes indicate overfitting, particularly when such performance is not reflected totally on unseen data. Several strategies were implemented to mitigate the risk of overfitting.

a. the implementation of 5-fold cross-validation for both the ML and DL methods, ensuring the validation of models against multiple diverse splits of data (see Section 2.2. II).
b. the incorporation of a dropout layer within the 1D CNN architecture to penalize overly complex models, thereby reducing overfitting (see Table 4).
c. the inclusion of an early stop mechanism in the 1D CNN to halt training if the validation loss ceased to decrease, preventing the model from learning noise in the training data (see Section 4.2.a).
d. the monitoring learning curves for both ML and DL methods to track how the training and validation error (loss) change during the training phase. No significant disparity was observed between the training error and the validation error, suggesting no overfitting; the Figures are not shown here.
e. the high number of dataset samples ($m = 21{,}540$, see Table 6) also serves as an indication that may help the ML methods prevent overfitting.

### 5.10. Initialization of the experimental framework in this study

To initialize the framework of experience and activate the different included models as shown in Fig. 1, the following steps were necessary.

a. Before beginning k-means shown in Fig. 1a: collecting voltage sag samples as $v_{abc}$ and $i_{abc}$ waveforms as well as labels, extracting values $2 \times N$ before and $3 \times N$ during voltage sag, applying DFT, formulating Feature forms 1 to 4, creating input matrix $X$ and normalization.
b. Before beginning the traditional ML and DL models show in Fig. 1b: normalized matrix $X$ and labels.
c. Before beginning the best-selected models shown in Fig. 1c: recorded voltage sag sample as $v_{abc}$ and $i_{abc}$ waveforms, preprocessing steps such as resampling $N$, extracting values $2 \times N$ before and $3 \times N$ during voltage sag, applying DFT, formulating Feature forms 1 and 4, normalization with the same parameters in the training phase.

From another perspective, for the CNN model, the initiation of the method requires the specification of initial model parameters, the learning rate, batch size, and dropout value. We employed the Glorot uniform initializer for weight initialization and determine the initial learning rate and batch size based on preliminary experiments and literature benchmarks. To assess the impact of these initial conditions on our algorithm's outcome, we conducted a sensitivity analysis, varying these parameters and evaluating the effect on model performance. Our analysis revealed that, while initial conditions can influence early training phases, the use of adaptive optimization, coupled with methodical hyperparameter tuning and cross-validation, ensures the model's ultimate performance is stable and generalizable across different data subsets.

### 5.11. Future works

Future research directions or suggestions stemming from this study can be summarized as follows.

- incorporating synthetic noises into voltage sag datasets to simulate the real waveform behavior of a voltage sag accurately when preparing a dataset.
- exploring additional classes (labels) such as short circuit types (i.e., LG, LL, etc.) added to the DS class or voltage sag source types (i.e., power system fault, TE, IM loading, etc.) added to both DS/US classes, transitioning from binary to multi-class classification.
- extending the developed methods coupled with the input feature forms beyond VSSL, such as applying them to protect the radial and/or loop topology of electrical distribution systems using multi classes with labels designed as not-operation of an over current relay (class 1), operation of relay as primary (class 2), operation as back-up (class 3), and operation as back-up of back-up (class 4).
- investigating other DL methods for VSSL such as LSTM (as discussed in Section 2.2.f), transfer learning (as discussed in Sections 5.3.b and 5.5).
- considering alternative input Feature forms for VSSL, such as sequences of impedance magnitude and/or angle instances or raw current waveforms.
- assessing the effects of selecting a higher number of extracted voltage sag cycles, such as four or five.
- drawing lessons learned from this research, among others, including (a) the importance of comprehensive datasets involving clean and noisy samples, making more trust to the methods performance, (b) the effectiveness of time-series input Feature forms for the DL methods to extract relevant features, (c) the significance of proper Feature forms for the traditional ML methods, (d) recent publications on VSSL offer an opportunity to delve into the field and uncover methods for pinpointing the source location of various power-quality

disturbances, especially in systems with a high penetration of RERs, where identifying the origin of disturbances can be challenging.

- ➢ exploring the applicability of Feature forms and developed methods within modern electrical power systems, including RERs and EVs.
- ➢ improving the robustness performance of k-means clustering by refining the unsupervised method utilizing Feature form 4 and incorporating feature reduction techniques like kernel principal component analysis (KPCA) (Mohammadi et al., 2022c, 2022d) applied to real-world recorded waveforms (without labeled data), which could also be influenced by climate change factors (Mohammadi et al., 2023). This approach involves extracting the most critical features from the initial set of 640 instances in the evolution of Feature form 4. Subsequently, retraining the k-means clustering might bolster its effectiveness further.
- ➢ exploring the economic implications of implementing ML-based methods in industrial settings to enhance overall performance and reduce production losses of large factories or microgrids, particularly within the context of the Internet of Things (IoT). This would require sufficiently robust data recorded by electrical sensors within the framework of the IoT. For example, an approach proposed in (Kliestik et al., 2023) involves utilizing predictive analytics coupled with artificial intelligence to enhance industrial performance within the IoT context.

## 6. Conclusion

The paper introduced three distinct input time-sample-based Feature forms tailored for machine learning (ML) and deep learning methods aimed at precise voltage sag source localizations (VSSL). Employing the straightforward concept of k-means clustering to differentiate between downstream and upstream classes of sag sources, we assessed the effectiveness of these Feature forms. Through extensive voltage sag simulations involving various noise levels within a regional electrical power network, k-means identified Feature Form 4 as the most prominent, achieving the highest accuracy among the tested feature forms. Subsequently, we developed traditional ML methods, including a specially designed one-dimensional convolutional neural network (1D-CNN). Applying these methods to the proposed Feature Form 4 and comparing them with the already existing Feature Form 1 through the extensive simulations, revealed that Feature Form 4 enhanced the performance of traditional ML methods consistently, and facilitated the creation of a CNN model with less complexity than using Feature form 1, achieving accuracy levels exceeding 99%. Furthermore, when Feature form 4 was utilized as an analytical rule for VSSL, it achieved a lower accuracy of approximately 92%, indicating its unsuitability as an analytical rule but highlighting its value as a time-sample-based Feature form for supervised learning methods.

Validation through real-world field measurements confirmed the satisfactory performance of k-means as an unsupervised method for identifying the most influential Feature form, along with the exceptional accuracy of the designed CNN, both utilizing Feature form 4. Notably, the prediction time for the combination of the designed CNN and Feature form 4 was notably short, less than 30 ms, making it applicable in real-time scenarios of VSSL. It is crucial to emphasize that the implementation of the most prominent Feature form coupled with the highly effective developed methods did not require specific optimization algorithms, complex mathematical techniques, or mandatory thresholds. This approach ensures straightforward implementation with low complexity and computational cost. Based on the analysis conducted in this study, identifying the best methods for real-time deployment requires a trade-off between accuracy, prediction speed, and model complexity. For instance, traditional ML methods offer less complexity than CNNs, with faster prediction times but slightly lower accuracy.

The findings of this study have practical applications in real-time scenarios, serving as a directional function within protection relays in electrical transmission, distribution, and microgrid systems for both transient and permanent faults, as well as addressing voltage sags from upstream sources. Electrical network operators, large factory owners, and producers of renewable energy parks can benefit from these findings by implementing preventive maintenance measures, reducing equipment downtime and damage in industry and electrical power systems, mitigating financial losses, and facilitating the assignment of power-quality penalties to responsible parties.

## CRediT authorship contribution statement

**Younes Mohammadi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Boštjan Polajžer:** Writing – review & editing, Investigation, Conceptualization. **Roberto Chouhy Leborgne:** Writing – review & editing, Investigation. **Davood Khodadad:** Writing – review & editing, Supervision, Funding acquisition.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgement

The authors would like to thank the funding support from the Kempe Foundation (*Kempestiftelserna*) for Grant number JCK22-0025, Sweden. This work has been also supported by the ARIS under Project P2-0115, Slovenia.

## Appendix A. Extra information regarding the case study

The “basic” features of the simulated case study depicted in Fig. 3 involved extensive distances between the load centers and power generation plants and involved 6619 km of 230 kV and 138 kV overhead transmission and sub-transmission lines (modeled as Bergeron with 3 conductors), with minimum 9 km and maximum 365 km length, as marked in Fig. 3. The primary generation units comprised natural gas facilities, hydroelectric plants, and thermoelectric plants. They were modeled by a 3-phase voltage source serried with an impedance with a fix control. The system incorporates 93 transformers with a total installed capacity of 2076 MVA. The winding type of the transformers was Y–Y grounded, or Δ-Y grounded, as shown in Fig. 3. The grounding is as solid; however, the neutral grounding resistance may influence the sags and the VSSL methods, which presents a potential area for future investigations. Additionally, there is a generation park composed of various concessionaires, self-producers, and independent producers, contributing to a total generation capacity of approximately 1643 MVA. Loads having active and reactive powers are modeled as constant power and impedance. Shunt reactors are modeled as 3-phase inductive load with a Y configuration, installed on 138 and 230 kV lines to enhance the system stability. Regarding the PQMs utilized in the network, PQM1 is installed at the boundary of a radial network with a single source and constant impedance/power loads. PQM2 is situated at the border of a radial network with two sources, owing to the presence of a 15 MVA DG on the DS side.

PQM3, PQM4 and PQM5 are installed at the boundary of an interconnected network with two-source, whereas PQM6 is situated at the boundary of a radial network with a single-source and a large load composed of an induction motor load (13.8 kV, 3200 HP) (Mohammadi et al., 2021).

## Appendix B. Recall, precision, and F1-score

Equations B.1 and B.2 elucidate the F1-score calculation for both DS and US classes. True positive (TP) and true negative (TN) represent the count of samples classified accurately for classes 1 and 0, respectively (i.e., TP and TF represent the number of samples with label 1 or 0 which are predicted correctly as 1 or 0, respectively). False negative (FN) and false positive (FP) correspond to the count of samples classified incorrectly as class 0 and 1, respectively. The F1 scores per class indicate the model's balanced precision and recall (sensitivity) capability for each specific class.

$$\text{F1}-\text{Score (DS)}(\%)=2\times\frac{Recal(DS)\times Precision(DS)}{Recal(DS)+Precision(DS)}\times 100, Recal(DS)=\frac{TP}{TP+FN}, Precision(DS)=\frac{TP}{TP+FP} \tag{B.1}$$

$$\text{F1}-\text{Score (US)}(\%)=2\times\frac{Recal(US)\times Precision(US)}{Recal(US)+Precision(US)}\times 100, Recal(US)=\frac{TN}{TN+FP}, Precision(US)=\frac{TN}{TN+FN} \tag{B.2}$$

## References


Ahmadi-Gorjayi, F., Mohsenian-Rad, H., 2023. Data-driven models for sub-cycle dynamic response of inverter-based resources using WMU measurements. IEEE Trans. Smart Grid 14, 4125–4128. https://doi.org/10.1109/TSG.2023.3280367.

Ahn, S.-J., Won, D.-J., Chung, I.-Y., Moon, S.-I., 2008. A new approach to determine the direction and cause of voltage sag. J Electr Eng Technol 3, 300–307. https://doi.org/10.5370/jeet.2008.3.3.300.

Aljarrah, R., Karimi, M., Azizipanah-Abarghooee, R., Salem, Q., Alnaser, S., 2024. Voltage dip propagation in renewable-rich power systems utilizing grid-forming converters. IET Renew. Power Gener. 1–11. https://doi.org/10.1049/rpg2.12939.

Aristi, I.A., Holbøll, J., Sørensen, T., Nielsen, A.H., Holmstrøm, O., Sørensen, P.E., 2009. Voltage dip caused by the sequential energization of wind turbine transformers. In: 2009 Eur. Wind Energy Conf. Exhib..

Bagnall, A., Lines, J., Bostrom, A., Large, J., Keogh, E., 2017. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Min. Knowl. Discov. 31, 606–660. https://doi.org/10.1007/s10618-016-0483-9.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogrammetry Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Bhujade, R., Maharjan, S., Khambadkone, A.M., Srinivasan, D., 2023. Economic analysis of annual load loss due to voltage sags in industrial distribution networks with distributed PVs. Sol. Energy 252, 363–372. https://doi.org/10.1016/j.solener.2023.01.041.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 2017. Classification and Regression Trees. Routledge. https://doi.org/10.1201/9781315139470.

Castello, P., Muscas, C., Pegoraro, P.A., Sulis, S., Rens, J., Van Zyl, J., 2023. A practical solution for locating the source of voltage dips in HV/MV interconnected grids. IEEE Open Access J Power Energy 10, 406–414. https://doi.org/10.1109/OAJPE.2023.3268499.

Cebrian, J.C., Giacomini, Jr J., Rossi, A.L.D., Trindade, F.C.L., Morales-Paredes, H.K., 2024. Methodology to estimate the financial impacts of the integration of PV generators in distribution systems on voltage sags and energy losses. IEEE Trans. Power Deliv. 1–13. https://doi.org/10.1109/TPWRD.2024.3358999.

Cen, S., Kim, D.O., Lim, C.G., 2023. A fused CNN-LSTM model using FFT with application to real-time power quality disturbances recognition. Energy Sci. Eng. 2267–2280. https://doi.org/10.1002/ese3.1450.

Chen, X., Li, D., Zhang, J., Ning, K., 2020. Research on location and recognition method of voltage sag disturbance. IOP Conf. Ser. Earth Environ. Sci. 619, 12027 https://doi.org/10.1088/1755-1315/619/1/012027.

Dehong, L., Jiekang, W., Zhen, L., 1990. Identification method of voltage sag source based on PSO algorithm with correlation analysis of multiple measures. In: 2023 IEEE 6th Int Electr Energy Conf, pp. 1990–1995. https://doi.org/10.1109/CIEEC58067.2023.10167207.

Deng, Y., Liu, X., Jia, R., Huang, Q., Xiao, G., Wang, P., 2021. Sag source location and type recognition via attention-based independently recurrent neural network. J Mod Power Syst Clean Energy 9, 1018–1031. https://doi.org/10.35833/MPCE.2020.000528.

Depally, S., Samyuktha, P., 2023. Detection of voltage sags and compensation in single phase power systems. In: Szymanski, J.R., Chanda, C.K., Mondal, P.K., Khan, K.A. (Eds.), Lect. Notes Electr. Eng. Springer Nature Singapore, Singapore, pp. 45–52. https://doi.org/10.1007/978-981-99-3691-5_4.

Dhara, S., Shrivastav, A.K., Sadhu, P.K., 2024. Power quality enhancement of microgrid using fuzzy logic controlled inverter and SFCL. Microsyst. Technol. https://doi.org/10.1007/s00542-023-05597-5.

Fatima, K., Shareef, H., Costa, F.B., Bajwa, A.A., Wong, L.A., 2024. Machine learning for power outage prediction during hurricanes: an extensive review. Eng. Appl. Artif. Intell. 133, 108056 https://doi.org/10.1016/j.engappai.2024.108056.

Gasquet, C., Witomski, P., 1999. Fourier Analysis and Applications : Filtering, Numerical Computation, Wavelets. Springer, New York (N.Y.).

Hamzah, N., Mohamed, A., Hussain, A., 2004. A new approach to locate the voltage sag source using real current component. Elec. Power Syst. Res. 72, 113–123. https://doi.org/10.1016/j.epsr.2004.03.010.

Huchche, V., Patne, N., 2023. Analytical assessment of torque and stator currents of an induction motor due to voltage sags. Int. J. Electr. Comput. Eng. 13, 3613–3621. https://doi.org/10.11591/ijece.v13i4.pp3613-3621.

IEEE Guide for voltage sag indices, 2014. IEEE Std 1–59. https://doi.org/10.1109/IEEESTD.2014.6842577.

Ismail, F., Jamaludin, J., 2023. Enhanced energy delivery for solar PV distributed generators at voltage sags. IEEE Access 11, 139688–139705. https://doi.org/10.1109/ACCESS.2023.3338361.

Jing, Y., Ma, Q., 2022. Locating the source of voltage sags based on CNN-Disturbance power method. Proc - 2022 Glob Conf Robot Artif Intell Inf Technol GCRAIT 632–635. https://doi.org/10.1109/GCRAIT55928.2022.00138.

Junjian, C., Jiekang, W., Zhen, L., 2023a. Voltage sag source identification method based on extreme learning. In: Proc - IEEE 6th Int Electr Energy Conf, pp. 1984–1989. https://doi.org/10.1109/CIEEC58067.2023.10166480.

Junjian, C., Jiekang, W., Zhen, L., 2023b. Prediction and classification of voltage sag trend based on support vector machine with parameter optimization. In: 2023 IEEE 6th Int. Electr. Energy Conf. IEEE, pp. 116–121. https://doi.org/10.1109/CIEEC58067.2023.10166514.

Kai, D., Wei, L., Jianfeng, S., Xianyong, X., Ying, W., 2021. Convolutional neural network for voltage sag source azimuth recognition in electrical Internet of Things. Wireless Commun. Mobile Comput. 2021, 6656564 https://doi.org/10.1155/2021/6656564.

Kanokbannakorn, W., Saengsuwan, T., Sirisukprasert, S., 2011. Unbalanced voltage sag source location identification based on superimposed quantities and negative sequence. In: 8th Electr. Eng. Electron. Comput. Telecommun. Inf. Technol. Assoc. Thail. - Conf. 2011, pp. 617–620. https://doi.org/10.1109/ECTICON.2011.5947915.

Khosravi, A., Melendez, J., Colomer, J., Sanchez, J., 2007. A hybrid method for sag source location in power network. In: 2007 9th Int Conf Electr Power Qual Util EPQU. https://doi.org/10.1109/EPQU.2007.4424087.

Khosravi, A., Melendez, J., Zapateiro, M., Colomer, J., 2008. Classification of Voltage Sags Based on Multiway Principal Component Analysis and Case Based Reasoning. vol. 41. IFAC. https://doi.org/10.3182/20080706-5-kr-1001.00932.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D.J., 2021. 1D convolutional neural networks and applications: a survey. Mech. Syst. Signal Process. 151, 107398 https://doi.org/10.1016/j.ymssp.2020.107398.

Kliestik, T., Nica, E., Durana, P., Popescu, G.H., 2023. Artificial intelligence-based predictive maintenance, time-sensitive networking, and big data-driven algorithmic decision-making in the economics of Industrial Internet of Things. Oeconomia Copernicana 14, 1097–1138. https://doi.org/10.24136/oc.2023.033.

Kong, W., Dong, X., Chen, Z., 2008. Voltage sag source location based on instantaneous energy detection. Elec. Power Syst. Res. 78, 1889–1898. https://doi.org/10.1016/j.epsr.2008.03.016.

Li, C., Tayjasanant, T., Xu, W., Liu, X., 2003. Method for voltage-sag-source detection by investigating slope of the system trajectory. Gener Transm Distrib IEE Proceedings-150, 367–372. https://doi.org/10.1049/ip-gtd:20030214.

Li, H., Ma, Z., Weng, Y., 2022. A transfer learning framework for power system. IEEE Trans. Power Syst. 37, 4424–4435. https://doi.org/10.1109/TPWRS.2022.3153445.

Li, T., Wu, Z., Liu, Y., Jia, R., 2023a. Voltage sag source location based on multi-layer perceptron and transfer learning. Front. Energy Res. 11 https://doi.org/10.3389/fenrg.2023.1237239.

Li, G., Yu, Z., Zheng, B., Qi, B., Su, Z., Wang, D., 2023b. Voltage sag source location based on the random forest. J Phys Conf Ser 2584, 012144. https://doi.org/10.1088/1742-6596/2584/1/012144.

Li, T., Wu, Z., Liu, Y., Jia, R., 2023c. Voltage sag source location based on multi-layer perceptron and transfer learning. Front. Energy Res. 11, 1–12. https://doi.org/10.3389/fenrg.2023.1237239.

Ling Ai, W., Shareef, H., 2013. A single monitor method for voltage sag source location using hilbert huang transform. Res. J. Appl. Sci. Eng. Technol. 5, 192–202. https://doi.org/10.1109/ASSCC.2012.6523296.

Liu, C., Zhao, J., Wu, W., Lu, Z., Zhang, R., 2020. Voltage sag source location based on comprehensive criterion and neural network method. {IOP} conf ser. Earth Environ Sci 512, 12131. https://doi.org/10.1088/1755-1315/512/1/012131.

Liu, Y., Zhang, J., Jia, H., Yuan, L., Zhou, M., 2022. Identification of voltage sag sources in the electrified railway power supply system based on CNNs. Wireless Commun. Mobile Comput. 2022 https://doi.org/10.1155/2022/4602187.

Liu, J., Wang, X., Xie, F., Wu, S., Li, D., 2023. Condition monitoring of wind turbines with the implementation of spatio-temporal graph neural network. Eng. Appl. Artif. Intell. 121 https://doi.org/10.1016/j.engappai.2023.106000.

Lobos, T., Kozina, T., Koglin, H.-J., 2001. Power system harmonics estimation using linear least squares method and SVD. IEE Proc. Generat. Transm. Distrib. 148, 567–572. https://doi.org/10.1049/ip-gtd:20010563.

Lu, X., Ye, X., Cheng, Y., 2024. An overlapping minimization-based over-sampling algorithm for binary imbalanced classification. Eng. Appl. Artif. Intell. 133, 108107 https://doi.org/10.1016/j.engappai.2024.108107.

Meléndez, J., Berjaga, X., Herraiz, S., Barrera, V., Sánchez, J., Castro, M., 2008. Classification of sags according to their origin based on the waveform similarity. In: 2008 IEEE/PES Transm Distrib Conf Expo Lat Am T D-LA. https://doi.org/10.1109/TDC-LA.2008.4641819.

Miraftabzadeh, S.M., Colombo, C.G., Longo, M., Foiadelli, F., 2023. K-means and alternative clustering methods in modern power systems. IEEE Access 11, 119596–119633. https://doi.org/10.1109/ACCESS.2023.3327640.

Moghaddam, M.J.H., 2021. Power quality improvement in the distribution network using optimization of the hybrid distributed generation system. Inst Sustain Ind Liveable Cities Coll Eng Sci Victoria Univ.

Mohammadi, Y., Leborgne, R.C., 2020a. Improved DR and CBM methods for finding relative location of voltage sag source at the PCC of distributed energy resources. Int. J. Electr. Power Energy Syst. 117, 105664 https://doi.org/10.1016/j.ijepes.2019.105664.

Mohammadi, Y., Leborgne, R.C., 2020b. A new approach for voltage sag source relative location in active distribution systems with the presence of inverter-based distributed generations. Elec. Power Syst. Res. 182 https://doi.org/10.1016/j.epsr.2020.106222.

Mohammadi, Y., Moradi, M.H., Chouhy Leborgne, R., 2017a. Locating the source of voltage sags: full review, introduction of generalized methods and numerical simulations. Renew. Sustain. Energy Rev. 77, 821–844. https://doi.org/10.1016/J.RSER.2017.04.017.

Mohammadi, Y., Moradi, M.H., Chouhy Leborgne, R., 2017b. Employing instantaneous positive sequence symmetrical components for voltage sag source relative location. Elec. Power Syst. Res. 151 https://doi.org/10.1016/j.epsr.2017.05.030.

Mohammadi, Y., Moradi, M.H., Chouhy Leborgne, R., 2017c. A novel method for voltage-sag source location using a robust machine learning approach. Elec. Power Syst. Res. 145, 122–136. https://doi.org/10.1016/j.epsr.2016.12.028.

Mohammadi, Y., Polajžer, B., Leborgne, R.C., Khodadad, D., 2024. Quantifying power system frequency quality and extracting typical patterns within short time scales below one hour. Sustain. Energy Grids Netw., 101359 https://doi.org/10.1016/j.segan.2024.101359.

Mohammadi, Y., Salarpour, A., Chouhy Leborgne, R., 2021. Comprehensive strategy for classification of voltage sags source location using optimal feature selection applied to support vector machine and ensemble techniques. Int. J. Electr. Power Energy Syst. 124, 106363 https://doi.org/10.1016/j.ijepes.2020.106363.

Mohammadi, Y., Leborgne, R.C., Polajžer, B., 2022a. Modified methods for voltage-sag source detection using transient periods. Elec. Power Syst. Res. 207, 107857 https://doi.org/10.1016/j.epsr.2022.107857.

Mohammadi, Y., Miraftabzadeh, S.M., Bollen, M.H.J., Longo, M., 2022b. Voltage-sag source detection: Developing supervised methods and proposing a new unsupervised learning. Sustain Energy, Grids Networks 32, 100855. https://doi.org/10.1016/J.SEGAN.2022.100855.

Mohammadi, Y., Mahdi Miraftabzadeh, S., Bollen, M.H.J., Longo, M., 2022c. Seeking patterns in rms voltage variations at the sub-10-minute scale from multiple locations via unsupervised learning and patterns' post-processing. Int. J. Electr. Power Energy Syst. 143 https://doi.org/10.1016/j.ijepes.2022.108516.

Mohammadi, Y., Miraftabzadeh, S.M., Bollen, M.H.J., Longo, M., 2022d. An unsupervised learning schema for seeking patterns in rms voltage variations at the sub-10-minute time scale. Sustain Energy, Grids Networks 31, 100773. https://doi.org/10.1016/j.segan.2022.100773.

Mohammadi, Y., Palstev, A., Polajžer, B., Miraftabzadeh, S.M., Khodadad, D., 2023. Investigating winter temperatures in Sweden and Norway: potential relationships with climatic indices and effects on electrical power and energy systems. Energies 16. https://doi.org/10.3390/en16145575.

Mohseni, M., Islam, S.M., Masoum, M.A.S., 2011. Impacts of symmetrical and asymmetrical voltage sags on DFIG-based wind turbines considering phase-angle jump, voltage recovery, and sag parameters. IEEE Trans. Power Electron. 26, 1587–1598. https://doi.org/10.1109/TPEL.2010.2087771.

Mohsenian-Rad, H., Xu, W., 2023. Synchro-waveforms: a window to the future of power systems data analytics. IEEE Power Energy Mag. 21, 68–77. https://doi.org/10.1109/MPE.2023.3288583.

Molla, E.M., Kuo, C.C., 2020. Voltage sag enhancement of grid connected hybrid PV-wind power system using battery and SMES based dynamic voltage restorer. IEEE Access 8, 130003–130013. https://doi.org/10.1109/ACCESS.2020.3009420.

Moradi, M.H., Mohammadi, Y., 2012. Voltage sag source location: a review with introduction of a new method. Int. J. Electr. Power Energy Syst. 43 https://doi.org/10.1016/j.ijepes.2012.04.041.

Moradi, M.H., Mohammadi, Y., 2013. A new current-based method for voltage sag source location using directional overcurrent relay information. Int Trans Electr Energy Syst 23. https://doi.org/10.1002/etep.659.

Moradi, M.H., Mohammadi, Y., Hoseyni Tayyebi, M., 2012. A novel method to locate the voltage sag source: a case study in the Brazilian power network (Mato Grosso). Przeglad Elektrotechniczny 88, 112–115.

Morais, J., Pires, Y., Cardoso, C., Klautau, A., 2010. A framework for evaluating automatic classification of underlying causes of disturbances and its application to short-circuit faults. IEEE Trans. Power Deliv. 25, 2083–2094. https://doi.org/10.1109/TPWRD.2010.2052932.

Nair, V., Hinton, G., 2010. Rectified linear units improve restricted Boltzmann machines vinod nair. Proc. ICML 27, 807–814.

Núñez, V.B., Frigola, J.M., Jaramillo, S.H., Losada, J.S., 2010. Evaluation of fault relative location algorithms using voltage sag data collected at 25-kV substations. Eur. Trans. Electr. Power 20, 34–51. https://doi.org/10.1002/etep.393.

Olshen, R.A., Rajaratnam, B., 2010. Successive normalization of rectangular arrays. Ann. Stat. 38, 1638–1664. https://doi.org/10.1214/09-AOS743.

Parsons, A.C., Grady, W.M., Powers, E.J., Soward, J.C., 2000. A direction finder for power quality disturbances based upon disturbance power and energy. IEEE Trans. Power Deliv. 15, 1081–1086. https://doi.org/10.1109/61.871378.

Passos, F.O., de Carvalho Filho, J.M., Leborgne, R.C., da Silveira, P.M., Ribeiro, P.F., 2015. An alternative approach to locating voltage sag source side at the point of common coupling based on power-flow information. J Control Autom Electr Syst 26, 579–587. https://doi.org/10.1007/s40313-015-0199-x.

Paul, D., Goswami, A.K., Kumar, S., Jain, S., Pandey, A., 2020. Propagation of voltage sag considering different winding connections: impact on the healthiness of transformers. IEEE Trans. Ind. Appl. 56, 6186–6196. https://doi.org/10.1109/TIA.2020.3016618.

Peng, J., Li, H., Wang, Z., Ghassemi, F., Jarman, P., 2013. Stochastic assessment of voltage dips caused by transformer energisation. IET Gener. Transm. Distrib. 7, 1383–1390. https://doi.org/10.1049/iet-gtd.2013.0091.

Polajžer, B., Štumberger, G., Seme, S., Dolinar, D., 2007. Impact of asymmetrical disturbance events on voltage sag source detection. Renew Energy Power Qual J 1, 111–114. https://doi.org/10.24084/repqj05.227.

Polajzer, B., Stumberger, G., Seme, S., Dolinar, D., 2008. Generalization of methods for voltage sag source detection using vector space approach. In: 2008 IEEE Ind. Appl. Soc. Annu. Meet, pp. 1–8. https://doi.org/10.1109/08IAS.2008.370.

Polajžer, B., Štumberger, G., Dolinar, D., 2015a. Detection of voltage sag sources based on the angle and norm changes in the instantaneous current vector written in Clarke's components. Int. J. Electr. Power Energy Syst. 64, 967–976. https://doi.org/10.1016/j.ijepes.2014.08.011.

Polajžer, B., Štumberger, G., Dolinar, D., 2015b. Instantaneous positive-sequence current applied for detecting voltage sag sources. IET Gener. Transm. Distrib. 9, 319–327. https://doi.org/10.1049/iet-gtd.2014.0483.

Polajžer, B., Pintarič, M., Rošer, M., Štumberger, G., 2019. Protection of MV closed-loop distribution networks with Bi-directional overcurrent relays and GOOSE communications. IEEE Access 7, 165884–165896. https://doi.org/10.1109/ACCESS.2019.2952934.

Pradhan, A.K., Routray, A., 2005. Applying distance relay for voltage sag source detection. IEEE Trans. Power Deliv. 20, 529–531. https://doi.org/10.1109/TPWRD.2004.839741.

Pradhan, A.K., Routray, A., Madhan Gudipalli, S., 2007. Fault direction estimation in radial distribution system using phase change in sequence current. IEEE Trans. Power Deliv. 22, 2065–2071. https://doi.org/10.1109/TPWRD.2007.905340.

Rao, C.S.V.P., Pandian, A., Reddy, C.R., Gulzar, M.M., Khalid, M., 2024. A novel hybrid RERNN-SCSO technique-based unified power quality conditioner of microgrid in an EV charging station. Arabian J. Sci. Eng. https://doi.org/10.1007/s13369-024-08765-5.

Saadat, A., Hooshmand, R.A., Kiyoumarsi, A., Tadayon, M., 2022. Voltage sag source location in distribution networks with DGs using cosine similarity. IEEE Trans. Instrum. Meas. 71, 1–10. https://doi.org/10.1109/TIM.2022.3218516.

Saleh, S.A., Ozkop, E., Al-Durra, A., Hill, T., Meng, J., Valdes, M.E., 2023. On the assessment of sampling rate impacts on responses of digital protective relays. IEEE Trans. Ind. Appl. 1–11. https://doi.org/10.1109/TIA.2023.3287477.

Shao, Z., Peng, J., Kang, J., 2010. Locating voltage sag source with impedance measurement. In: 2010 Int. Conf. Power Syst. Technol., pp. 1–6. https://doi.org/10.1109/POWERCON.2010.5666557.

Shareef, H., Mohamed, A., Ibrahim, A.A., 2013. Identification of voltage sag source location using S and TT transformed disturbance power. J Cent South Univ 20, 83–97. https://doi.org/10.1007/s11771-013-1463-5.

Solak, B., Sikorski, T., 2019. Analysis of voltage dip source location methods. 2019 Mod. Electr. Power Syst. 1–6. https://doi.org/10.1109/MEPS46793.2019.9394971.

Sun, H., Yi, H., Yang, G., Zhuo, F., Hu, A., 2019. Voltage sag source identification based on few-shot learning. IEEE Access 7, 164398–164406. https://doi.org/10.1109/ACCESS.2019.2953226.

Sundarakani, B., Rajamani, H.-S., Madmoune, A., 2024. Sustainability study of electric vehicles performance in the UAE: moderated by blockchain. Benchmarking An Int J 31, 199–219. https://doi.org/10.1108/BIJ-10-2021-0624.

Tan, M gang, Zhang, C., Zhang, R., Chen, B., 2023. Disturbance extracted methods for auxiliary power quality monitor-based voltage sag localization in distribution network. Int. J. Electr. Power Energy Syst. 145, 108675 https://doi.org/10.1016/j.ijepes.2022.108675.

Tayjasanant, T., Li, C., Xu, W., 2005. A resistance sign-based method for voltage sag source detection. IEEE Trans. Power Deliv. 20, 2544–2551. https://doi.org/10.1109/TPWRD.2005.852341.

Veizaga, M., Delpha, C., Diallo, D., Bercu, S., Bertin, L., 2023. Classification of voltage sags causes in industrial power networks using multivariate time-series. IET Gener. Transm. Distrib. 17, 1568–1584. https://doi.org/10.1049/gtd2.12765.

Wang, Y., Xiao, Z., Cao, G., 2022. A convolutional neural network method based on Adam optimizer with power-exponential learning rate for bearing fault diagnosis. J Vibroengineering 24, 666–678. https://doi.org/10.21595/jve.2022.22271.

Wang, R., Zhuang, Z., Tao, H., Paszke, W., Stojanovic, V., 2023. Q-learning based fault estimation and fault tolerant iterative learning control for MIMO systems. ISA Trans. 142, 123–135. https://doi.org/10.1016/j.isatra.2023.07.043.

Wu, L., Zhang, Y., Hao, X., Chen, W., 2020. Research on a location method for complex voltage sag sources based on random matrix theory. Math. Probl Eng. 2020, 7870461 https://doi.org/10.1155/2020/7870461.

Wu, L., He, L., Chen, W., Hao, X., 2023. Recognition method of voltage sag sources based on RMT-CNN model. IAENG Int. J. Appl. Math. 53.

Xu, D., Li, Y., Zhang, B., Liu, S., Tang, Y., Yan, H., et al., 2022. Upstream and downstream location of voltage sag source based on sparrow search algorithm and RBF neural network. In: Proc - 2022 Power Syst Green Energy Conf PSGEC 2022, pp. 760–765. https://doi.org/10.1109/PSGEC54663.2022.9881060.

Yalman, Y., Uyanık, T., Tan, A., Bayındır, K.Ç., Terriche, Y., Su, C.L., et al., 2022a. Implementation of voltage sag relative location and fault type identification algorithm using real-time distribution system data. Mathematics 10. https://doi.org/10.3390/math10193537.

Yalman, Y., Uyanık, T., Atlı, İ., Tan, A., Bayındır, K.C., Karal, Ö., et al., 2022b. Prediction of voltage sag relative location with data-driven algorithms in distribution grid. Energies 15. https://doi.org/10.3390/en15186641.

Ye, Z.J., Izadi, M., Farajollahi, M., Mohsenian-Rad, H., 2023. A remedy to losing time synchronization at D-PMUs, H-PMUs, and WMUs in event location identification in power distribution systems. IEEE Trans. Smart Grid 1. https://doi.org/10.1109/TSG.2023.3277853.

Yi, T., Jie, H., Hao, L., Lei, W., 2016. Method for voltage sag source location based on the internal resistance sign in a single-port network. IET Gener. Transm. Distrib. 10, 1720–1727. https://doi.org/10.1049/iet-gtd.2015.1329.

Yilin, Y., Yonghai, X., 2011. Research of method for voltage sag source detection in power distribution network. In: 2011 6th IEEE Conf. Ind. Electron. Appl, pp. 485–488. https://doi.org/10.1109/ICIEA.2011.5975634.

Zeyu, W., Wei, T., Jiekang, W., Zhen, L., 2023. Identification method for complex voltage sag source using fuzzy grey relational analysis. In: 2023 IEEE 6th Int. Electr. Energy Conf. IEEE, pp. 1972–1977. https://doi.org/10.1109/CIEEC58067.2023.10166406.

Zhang, S., Li, X., Zong, M., Zhu, X., Wang, R., 2017. Efficient kNN classification with different numbers of nearest neighbors. IEEE Transact. Neural Networks Learn. Syst. 1–12. https://doi.org/10.1109/TNNLS.2017.2673241.

Zhang, L., Jiang, C., Chai, Z., He, Y., 2024. Adversarial attack and training for deep neural network based power quality disturbance classification. Eng. Appl. Artif. Intell. 127, 107245 https://doi.org/10.1016/j.engappai.2023.107245.

Zhu, Z., Zhang, Y., Deng, Z., Wang, M., 2023. A novel voltage sag detection method for analyzing charging quality of electric vehicle. Comput. Electr. Eng. 112, 108991 https://doi.org/10.1016/j.compeleceng.2023.108991.

Zhuang, Z., Tao, H., Chen, Y., Stojanovic, V., Paszke, W., 2023. An optimal iterative learning control approach for linear systems with nonuniform trial lengths under input constraints. IEEE Trans Syst Man, Cybern Syst 53, 3461–3473. https://doi.org/10.1109/TSMC.2022.3225381.