

# 기계학습 기반 D-PMU 데이터를 활용한 전압 이벤트 분류 기법

김선홍, 김임준, 김태근, 윤성국  
승실대학교

## Machine Learning based Voltage Event Classification Using D-PMU Data

Seon-Hong Kim, Im-Joon Kim, Tae-Geun Kim, Sung-Guk Yoon  
Soongsil University

**Abstract** - 다량의 재생에너지의 보급 계획에 대비하기 위하여 전력계통의 고품질의 데이터의 수집 및 분석이 요구된다. 최근 배전계통에서의 고정밀 계측장치인 D-PMU (Distribution Phasor Measurement Unit)의 보급으로 인해 정밀한 데이터의 수집 및 분석이 가능해졌다. 본 논문에서는 D-PMU에서 획득한 배전계통 데이터를 활용한 기계학습 기반 전압 이벤트 분류 기법을 제안한다. 본 논문에서는 k-means 클러스터링, KNN (k-nearest neighbors), SVM (support vector machine)의 세 가지 머신러닝 기법을 사용하여 이벤트를 분류하였다.

### 1. 서 론

기후변화에 대응하여 우리나라는 2050년까지 재생에너지 비중 60~70% 달성을 목표로 하고 있다. 이를 달성하기 위하여 연간 재생에너지 보급 목표도 증가하고 있다. 대표적인 재생에너지원인 태양광과 풍력의 변동성과 불확실성은 전압을 불안정하게 하는 등 전력계통 운영의 어려움을 준다. 이러한 문제점을 해결하기 위한 필수적인 사전작업으로 전력계통 고품질 데이터의 수집 및 분석이 있다. 이에 최근 기계학습 기반 전력 데이터 분석 연구가 활발히 진행되고 있다[1].

배전계통에서도 재생에너지로 인한 문제가 동일하게 나타난다. 이를 해결하기 위해 GPS (Global Positioning System)로 시각 동기화를 맞춘 전압, 전류의 크기 및 위상과 주파수를 측정하는 단말 장치인 D-PMU가 보급되고 있는 추세이다. D-PMU는 계통 파라미터 실시간 추정, 배전계통 고장, 발전/부하 탈락과 같은 내부 이벤트 검출 및 위치추정을 통한 계통운영 신뢰도 향상 등 다양한 분야에 활용되고 있다[2].

본 논문에서는 D-PMU 데이터를 활용한 기계학습 기반 전압 이벤트 분류 기법을 제안한다. 본 논문에서 k-means 클러스터링, KNN, SVM의 세 가지 머신러닝 기법을 사용하여 배전계통의 전압 이벤트를 분류하였다. 승실대학교 캠퍼스 내에 설치된 D-PMU 데이터와 공개된 D-PMU 데이터를 본 연구에서 사용하였다.

### 2. 본 론

#### 2.1 전압 데이터 취득 및 생성

승실대학교 캠퍼스 내에 설치된 D-PMU의 2021년 5월부터 7월 사이의 전압 이벤트를 수집하였다. 이 기간 중 수집한 이벤트의 종류는 고조파(Harmonics), 급속전압변동(Rapid Voltage Change, RVC), 전압상승(Voltage Swell), 전압강하(Voltage Dip)이다. 그 중 전압상승과 전압강하 이벤트는 해당 기간 내에 발생하지 않거나 그 개수가 매우 적어서 공개된 D-PMU 데이터 세트를 기반으로 가우시안 잡음을 사용하여 데이터를 생성하였다[3]. 하나의 이벤트 데이터는 이상 전압 발생시간을 기준으로 100ms 단위 20개의 전압 데이터를 한 세트로 구성하였고 각 이벤트 당 총 100개의 데이터 세트를 사용하였다. 전압 데이터 중 선간전압의 경우 상전압의 값으로 표준화를 진행하여 모든 데이터를 상전압으로 사용하였다.

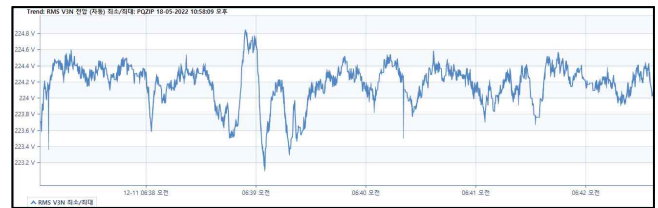
#### 2.2 전압 이벤트 특성 분석

##### 2.2.1 고조파(Harmonics)

고조파는 기본주파수에 대해 정수배를 갖는 전압 또는 전류로 정의된다. 그림 1과 같이 고조파를 포함한 파형은 포함하지 않는 기본파 파형에 비해 왜곡된 전압 파형으로 나타난다. 고조파가 심한 경우 변압기에서 철손 및 동손의 증가로 용량이 감소하고 정류기 및 각종 제어장치에서 제어신호의 위상차에 의한 오동작이 발생할 수 있다. 고조파로 인한 파형의 왜곡 정도를 나타내는 지표로 종합고조파왜율(Total Harmonic Distortion, THD)을 사용한다. THD는 다음과 같은 수식으로 나타낸다.

$$THD = \sqrt{\frac{\sum_{n=2}^{\infty} V_{n,rms}^2}{V_{1,rms}^2}} \quad (1)$$

$V_{n,rms}$ 는 기본주파수의 n배의 전압 RMS(Root Mean Square) 값이다. 배전계통에서 고조파로 인해 발생하는 여러 가지 문제를 없애기 위해 THD를 5% 이내로 유지하도록 하고 있다. 본 논문에서는 사용한 고조파 데이터는 THD값이 1.5 ~ 2.6%이다.



<그림 1> 고조파가 포함된 RMS 전압 파형

##### 2.2.2 급속전압변동(Rapid Voltage Change, RVC)

급속전압변동은 불특정 기간에 전압이 변하는 것을 말한다. 급속전압변동은 전자 장비에 심각한 영향을 미치지 않지만 깜박임이 발생하여 사용자에게 불편함을 야기할 수 있다[4]. 본 논문에서 사용한 급속전압변동 데이터는 전압 변동 폭이 10% 이하인 이벤트로 정의하여 데이터를 추출하였다.

##### 2.2.3 전압상승(Voltage Swell) 및 전압강하(Voltage Dip)

정상 전압에서 10% 이상, 이하로 50ms 이상 유지하는 상황을 전압상승, 전압강하라고 정의한다. 전압상승 및 전압강하는 테스트의 결과를 신뢰할 수 없게 만들거나 생산물에 변동을 주고 장비의 정상적인 작동을 방해하는 문제를 야기할 수 있다[5].

#### 2.3 분류 기법

##### 2.3.1 K-Means 클러스터링

K-Means는 비지도학습의 일종으로 주어진 데이터를 k개의 군집으로 묶는 분할법의 한 종류이다. K-Means는 레이블이 달려 있지 않은 입력 데이터에 레이블을 달아주는 작업을 수행할 수 있다. K-Means는 n개의 데이터 집합  $\{x_1, x_2, \dots, x_n\}$ 이 주어졌을 때, n개의 데이터를 각 군집과의 분산을 최소화하는  $k(\leq n)$ 개의 집합  $S = \{S_1, S_2, \dots, S_k\}$ 으로 분할한다. 이를 수

식으로 표현하면 다음과 같다.

$$\arg \min_s \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 (\mu_i \text{는 집합 } S_i \text{의 중심점}) \quad (2)$$

### 2.3.2 KNN(K-Nearest Neighbors)

KNN은 클래스를 모르는 미지의 데이터가 주어졌을 때 k개의 가까운 인접 이웃 사이에서 다수결에 따라 클래스를 정하는 방법이다. KNN 분류기는 주어진 데이터  $x$ 와 테스트 데이터 사이의 거리를 계산하고  $x$ 의 가장 가까운 지점에 의해 클래스를 결정한다. n개의 데이터 집합  $c = \{x_1, x_2, \dots, x_n\}$ 에 테스트 데이터  $y$ 가 주어졌을 때,  $y$ 와 가장 가까운 거리에 있는 k개의 데이터를 선정한다. 그 후 k개 학습 데이터의 클래스들을 이용하여  $y$ 의 클래스는

$$\arg \max_c \left[ \sum_{i=1}^K \sigma \omega(l_i, c) \right] \quad (3)$$

로 나타낼 수 있다. 여기서  $\sigma\{\omega(l_i), c\}$ 는 다음과 수식과 같이 구할 수 있다.

$$\sigma(\omega(l_i), c) = \begin{cases} 1, & \omega(l_i) = c \\ 0, & \omega(l_i) \neq c \end{cases} \quad (4)$$

### 2.3.3 SVM(Support Vector Machine)

SVM은 데이터가 주어졌을 때 결정 경계를 찾아서 경계의 어느 쪽에 속하는 지 확인하는 방법이다. 데이터의 속성에 따라 결정 경계의 속성(차원)이 달라지며 2차원의 경우 직선의 형태, 3차원의 경우 평면의 형태를 이용해 분류한다. SVM의 오차를 최소화하기 위해서는 데이터 군과 결정 경계는 최대한 멀리 떨어지도록 해야 한다. 이때 마진은 최대가 된다. 마진의 길이는 두 개의 분류경계면을 통해 구할 수 있는데 두 개의 분류경계면은

$$w_0 + w^T x^+ = 1 \quad (4)$$

$$w_0 + w^T x^- = -1 \quad (5)$$

으로 나타낼 수 있다. 마진은 이 두 경계면 사이의 거리를 의미하므로 마진의 길이는 다음과 같이 나타낸다.

$$\frac{2}{\|w\|} \quad (6)$$

### 2.4 실험 방법

본 논문에서는 각각의 기법에 3D 모델을 사용하였고 이를 위해 3차원 공간의 좌표축  $x$ 는 표준편차,  $y$ 는 전압 평균치에서 최대전압과 최소전압까지의 변위 절댓값 중 더 큰 값,  $z$ 는 전압의 평균치에서 최대, 최소전압까지의 변위의 절댓값 중 더 큰 값의 부호로 설정하였다. 이를 다음과 같은 수식으로 표현할 수 있다.

$$x = \sqrt{\frac{\sum (V_i - \bar{V})^2}{n-1}} \quad (7)$$

$$y = \max(|V_i - V_{avg}|) \quad (8)$$

$$z = \begin{cases} 1, & V_{yi} - V_{avg} > 0 \\ 0, & V_{yi} - V_{avg} < 0 \end{cases} \quad (9)$$

이때  $V_i$ 는 표본집단 안에 속해 있는 전압,  $\bar{V}$ 는 표본집단의 평균 전압,  $V_{avg}$ 는 모집단의 평균 전압,  $V_{yi}$ 는 기준  $y$ 를 만족하는  $V_i$ 를 말한다. 학습 및 시험 데이터를 KNN은 100개의 테스트 데이터와 300개의 훈련 데이터로 학습을 진행하였고 SVM은 테스트 데이터 80개, 학습 데이터 320개로 학습을 진행하였다.

### 2.5 실험 결과

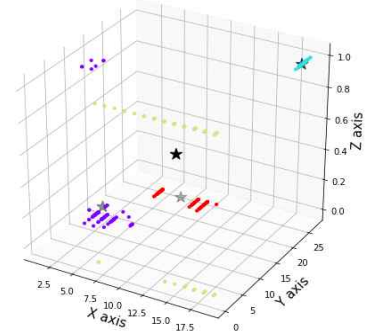
<표 1>은 K-Means 클러스터링에서 4개 군집을 만들어 분류를 진행했을 때의 결과이다. k-means 클러스터링은 이벤트 한 개를 제외하고 각 이벤트를 같은 군집으로 분류한다. <그림 2>는 K-Means 클러스터링 결과를 3차원 그래프로 시각화한 것이다. <표 2>는 KNN과 SVM의 학습 및 시험 데이터 정확도이다. KNN의 경우, 학습과 시험 정확도는 각각 99.65%, 100%로 나오는 것을 확인하였다. SVM의 경우, 학습과 시험 정확도는 각각 98.75%, 96.25%로 나오는 것을 확인하였다.

<표 1> K-Means 클러스터링 결과

	A	B	C	D
고조파	0	0	100	0
급속전압변동	100	0	0	0
전압상승	0	1	0	99
전압강하	0	100	0	0

<표 2> KNN, SVM 정확도(%)

	KNN	SVM
학습	99.65	98.75
시험	100	96.25



<그림 2> K-Means 클러스터링 시각화

## 3. 결 론

본 논문에서는 D-PMU 데이터를 활용하여 k-means 클러스터링, KNN, SVM의 세 가지 머신러닝 기법으로 전압 이벤트 분류를 진행하였다. 숭실대학교 캠퍼스 내에 설치된 D-PMU를 실 데이터와 공개 데이터 세트를 활용한 데이터를 통해 고조파, 급속전압변동, 전압상승, 전압강하의 4가지 이벤트 데이터 세트를 확보하였다. 전압의 표준편차, 변위의 절댓값의 최대값, 변위의 부호를 활용하여 학습한 모델의 실험 결과 K-Means, KNN, SVM은 각각의 분류 정확도가 99.75%, 100%, 96.25%로 나오는 것을 확인하였다. 본 연구에서 사용한 데이터로는 KNN이 가장 정확한 분류 성능을 보였다. 후속 연구로 머신러닝을 통해 계통에서 일어난 전압 이벤트들을 감지하고 이를 토대로 계통에서 일어나고 있는 문제점을 역추적하는 연구를 진행할 예정이다.

### 감사의 글

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임  
(No. 2020R1F1A1075137)

### [참 고 문 헌]

[1] 이승민, 이우진, “기계학습을 이용한 태양광 발전량 예측 및 결합 검출 시스템 개발”, 컴퓨터 및 통신시스템, 5.10, pp. 353-360, 2016  
 [2] 윤상윤, 홍지송, “배전계통에서의 PMU 활용 기술 동향”, 전기의 세계, 69(6), pp.29-33, 2020  
 [3] T. J. Swenson, E. Vrettos, J. Müller and C. Gehbauer, “Open  $\mu$ PMU Event Dataset: Detection and Characterization at LBNL Campus”, 2019 IEEE Power & Energy Society General Meeting (PESGM), pp. 1-5, 2019  
 [4] D. Macii and D. Petri, “On the detection of Rapid Voltage Change(RVC) events for power quality monitoring”, 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), pp.1-6, 2017  
 [5] N. Edomah, “Effects of voltage sags, swell and other disturbances on electrical equipment and their economic implications”, CIREC 2009 - 20th International Conference and Exhibition on Electricity Distribution - Part 1, pp.1-4, 2009