

K-means 클러스터링 기반 소프트맥스 신경회로망 부분방전 패턴분류의 설계 : 분류기 구조의 비교연구 및 해석

Design of Partial Discharge Pattern Classifier of Softmax Neural Networks Based on K-means Clustering : Comparative Studies and Analysis of Classifier Architecture

정 병 진* · 오 성 권*
(Byeong-Jin Jeong · Sung-Kwun Oh)

Abstract - This paper concerns a design and learning method of softmax function neural networks based on K-means clustering. The partial discharge data information is preliminarily processed through simulation using an Epoxy Mica Coupling sensor and an internal Phase Resolved Partial Discharge Analysis algorithm. The obtained information is processed according to the characteristics of the pattern using a Motor Insulation Monitoring System program. At this time, the processed data are total 4 types that void discharge, corona discharge, surface discharge and slot discharge. The partial discharge data with high dimensional input variables are secondarily processed by principal component analysis method and reduced with keeping the characteristics of pattern as low dimensional input variables. And therefore, the pattern classifier processing speed exhibits improved effects. In addition, in the process of extracting the partial discharge data through the MIMS program, the magnitude of amplitude is divided into the maximum value and the average value, and two pattern characteristics are set and compared and analyzed. In the first half of the proposed partial discharge pattern classifier, the input and hidden layers are classified by using the K-means clustering method and the output of the hidden layer is obtained. In the latter part, the cross entropy error function is used for parameter learning between the hidden layer and the output layer. The final output layer is output as a normalized probability value between 0 and 1 using the softmax function. The advantage of using the softmax function is that it allows access and application of multiple class problems and stochastic interpretation. First of all, there is an advantage that one output value affects the remaining output value and its accompanying learning is accelerated. Also, to solve the overfitting problem, L2-normalization is applied. To prove the superiority of the proposed pattern classifier, we compare and analyze the classification rate with conventional radial basis function neural networks.

Key Words : EMC sensor, Principal component analysis, K-means clustering, Softmax function, Nonlinear least squares method, L2-normalization, Radial basis function neural networks

1. 서 론

전기는 우리가 살아가는데 밀접한 관계를 맺고 있는 꼭 필요한 에너지원이다. 과학문명의 발달로 생활에서 많은 풍족함과 편리함을 가져다주는 전기의 중요성과 필요성은 지금의 삶을 돌아봐도 이미 입증된 공공연한 사실이다. 하지만 이렇듯 언제나 우리에게 유익함을 주는 것은 아니다. 자칫하면 큰 사고로 이어져 물질적, 경제적 손실뿐만 아니라 생명까지도 위협할 수 있으므로 안전관리가 필요하다. 현재 산업현장과 전력기기의 안전관리 및 열화진단은 초음파 신호 검출, 가스누출 여부 검사 및 부분방전

신호 검출 등이 있다. 대부분 현장에서 직접 기기를 이용한 관측과 전문적인 지식의 숙련된 기술들을 요구하는 연구가 활성화되고 있지만, 실시간으로 진단을 시행하고 환경에 따라 급속도로 진행되는 열화를 예측하는 것은 어려운 시점이다. 그러나 최근 구글 딥마인드(Google DeepMind)의 인공지능 바둑 프로그램 '알파고(AlphaGo)'가 쟁점이 되면서 소프트웨어 산업이 다시 주목받기 시작했고, 인간의 사고와 유사하게 생각하고 사람이 할 수 있는 것들을 기계가 구현할 수 있도록 도와주는 심층신경망(Deep Neural Network)기술을 결합하여 활용하는 시대가 열렸다. 열화 진단 중에서 부분방전 측정기법은 여러 가지 열화 과정에서 발생하는 다양한 신호들의 일정한 패턴과 중요한 정보를 제공하는데 이러한 정보를 이용하여 AI(Artificial Intelligence)와 접목해서 실시간으로 정확한 진단이 가능하도록 연구를 진행하였다.

본 논문에서는 K-means 클러스터링 기반 소프트맥스 신경회로망 설계 및 학습방법을 제안한다. 모의실험을 통하여 획득한 데이터는 구성 방법에 그 특징들이 달라지고 분류기의 분류율에도 영향을 미칠 수 있으므로 중요하다. 이때 잡음을 감쇄시켜 주

* Corresponding Author : School of Electrical & Electronic Engineering, the University of Suwon, Korea

E-mail: ohsk@suwon.ac.kr

* School of Electrical & Electronic Engineering, the University of Suwon, Korea

E-mail: qudwls2@suwon.ac.kr

Received : November 13, 2017; Accepted : November 30, 2017

고 민감한 부분도 잘 감지해주는 EMC 센서를 사용했고 센서 내부에 탑재된 PRPDA를 이용하여 전처리한다. 또한, MIMS 프로그램을 통하여 펄스 파형 크기를 최댓값과 평균값으로 나눈 두 가지의 경우로 특징을 추출하여 데이터를 사용한다. 부분방전 데이터는 총 1,200개이고 4개의 부류로 구성되며 보이드, 코로나, 연면, 슬롯 방전으로 이루어져 있다. 제안된 분류기의 구조는 입력층, 은닉층, 출력층과 같이 세 개의 층으로 구성되며 입력층과 은닉층의 연결가중치는 K-means 클러스터링을 통해 학습된 멤버십 값을 사용한다. 비용함수는 변형된 크로스 엔트로피 오차 함수를 사용하고, 다중 클래스의 문제를 해결하기 위해 출력층에 소프트맥스 함수를 이용한다. 또한, 반복 학습으로 뉴턴법을 이용한 비선형 최소자승법을 통하여 비용함수를 최소화하는 근사값을 찾는다. 크로스 엔트로피 비용함수에 L2 정규화를 적용하여 과적합되는 문제를 해소하였다. 기존의 분류기와 제안된 분류기를 비교하여 성능 향상과 우수성을 입증하였다.

2. 부분방전과 EMC 센서 선행조사

본 장에서는 부분방전이 무엇인지 왜 발생하는지 부분방전을 통해 우리가 얻고자 하는 방향성이 무엇인지를 다룰 것이다. 또한, 부분 방전을 분석하기 위해 준비한 시료를 사용하여 모의실험을 진행하였고 실험을 통한 부분방전의 패턴과 그 특징에 대하여 서술한다. 마지막으로 각 부분방전을 분류하기 위한 패턴 분석 과정에 관하여 설명한다.

2.1 부분방전 선행연구

전기장치에서 전극과 전극 사이를 완전히 교락(交絡)하지 않는 불안정한 절연과외 현상을 부분방전이라 한다. 이는 일종의 방전을 일컫거나 전계분포가 불평등한 절연물에 인가전압을 서서히 상승시키면 전계가 집중된 보이드나 크랙 등에서 부분적으로 발생한다. 또한, 복합 절연체 내의 보이드나 오염물질 등과 같은 결함도 방전을 일으킨다. 부분방전이 일어날 때 발열, 하전입자의 충돌, 분자와 이온에 의한 화학작용으로 주변의 고체 절연체에 손상을 입히게 된다. 부분방전의 종류는 고체 절연체 내부의 보이드나 공극에서 발생하는 내부방전, 기체 절연체에서 주로 첨점(尖點)에 발생하는 코로나와 절연체 표면에서 일어나는 표면방전 그리고 고체 절연체 내부에서 전기트리진전 시 발생하는 전기트리 방전 등으로 분류할 수 있다. 열화진전의 원인이 되는 부분방전의 발생 빈도가 잦아질수록 절연재료의 절연과외를 초래하게 된다. 부분방전이 일어나면서 나타나는 수많은 방전 수, 방전의 크기, 방전의 패턴 등과 같은 정보는 전기적 절연과외를 평가하는 데 매우 중요한 역할을 한다.

이처럼 유용한 정보를 담고 있는 부분방전은 기계적·전기적 스트레스 등 열화로 인한 이상을 조기에 발견하도록 도와준다. 비파괴검사 시험 등의 계측법을 이용하여 부분방전을 측정하고 결함의 위치를 찾으며 사전에 사고를 예방하는 데 활용된다. 위에서 언급한 정보 외에도 상시 설비된 장치를 이용하여 실시간 모니터링으로 On-line 진단도 가능하다는 측면에서 많은 장점과

가능성을 가지고 있다[1].

2.2 EMC(Epoxy Mica Coupling) 센서

스위치 기어, 변압기, 고전압 모터와 발전기의 선로 단말 등에 연결하여 전기 절연 및 부분방전의 여부를 감지하는 접촉식 센서를 가리켜 EMC라 한다. 8kV, 18kV, 28kV 전압 레벨로 구분되고, 80pF의 가장 일반적인 정전용량 값과 순수한 마이카 조각으로 구성되어있다. 마이카 섹션의 수는 전기적 스트레스를 감소시킬 수 있도록 가능한 한 커야 하며 신호 감쇄를 줄이면서 전기적 보호를 위해 권선에 근접하여 설치하거나 커패시터 위치에 설치되어야 한다. 또한, 설치 이후 반드시 접지시켜야 한다.

3. K-means 클러스터링 기반 소프트맥스 신경회로망 부분방전 패턴분류의 설계

3.1 주성분 분석법(Principle Component Analysis, PCA)

주성분 분석법은 대표적인 선형변환 특징 추출 방법이다. 고차원의 데이터를 저차원으로 축소하는 알고리즘으로 축소되기 전 실제 데이터 정보를 최대한 유지하는 것이 가장 중요하다. 축소된 데이터는 분류기의 입력데이터로 사용되어 연산의 양을 줄여줌으로 처리속도가 빨라진다. 이뿐만 아니라 데이터를 분석하는데 불필요한 신호나 잡음을 제거되는 효과를 기대할 수 있어 보다 신속하고 정확한 분석을 도와주는 전처리 알고리즘이다[3].

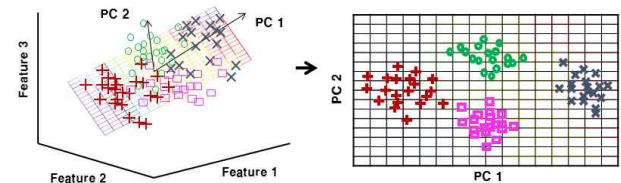


그림 1 주성분 분석법의 예

Fig. 1 Example of Principal Component Analysis

3.2 전반부 구조 동정 : K-means 클러스터링

전반부 구조 동정에서 데이터의 입력변수 선택과 멤버십 함수를 가우시안 함수, 삼각형 함수 등으로 결정된다. 본 논문은 비슷한 패턴, 속성 및 형태 등의 기준을 통하여 데이터를 분류하고 내부의 구조를 찾아내는 K-means 클러스터링 알고리즘을 사용하였다. 데이터와 각 클러스터와의 거리를 기준으로 근접한 정도를 측정하고 이를 바탕으로 특성별로 데이터를 분류한다. 이는 활성 함수의 적합도 개념과 같아 제안된 신경회로망의 은닉층 활성화 함수 형태를 표현할 수 있다. 아래 단계를 통해 K-means 클러스터링 전반부 동정이 수행된다[4].

Step 1: 클러스터의 개수를 $2 \leq c < n$ 로 결정하고, 소속행렬 U 을 $U^{(0)} \in M_c$ 으로 초기화

$$M_c = \left\{ U \mid u_{ik} \in \{0,1\}, \sum_{i=1}^c u_{ik} = 1, 0 < \sum_{k=1}^m u_{ik} < m \right\} \quad (1)$$

여기서 $u_{ik} (i = 1, 2, \dots, c; k = 1, 2, \dots, m)$ 는 소속행렬의 파라미터, m 은 데이터의 총수

Step 2: 각각의 클러스터에 대한 중심벡터 $\vec{\nu}_i$ 을 구함

$$\vec{\nu}_i = \{\nu_{i1}, \nu_{i2}, \dots, \nu_{ij}, \dots, \nu_{iN}\} \quad (2)$$

$$\nu_{ij} = \frac{\sum_{k=1}^m u_{ik} \cdot x_{kj}}{\sum_{k=1}^m u_{ik}} \quad (3)$$

여기서 i 는 i 번째 클러스터, j 는 j 번째 입력 변수($j = 1, \dots, N$), m 은 데이터 수

Step 3: 각각의 클러스터 중심과 데이터 거리를 계산하여 새로운 소속행렬 $U^{(r)}$ 을 생성

$$d_{ik} = d(\mathbf{x}_k - \vec{\nu}_i) = \|\mathbf{x}_k - \vec{\nu}_i\| = \left[\sum_{j=1}^N (x_{kj} - \nu_{ij})^2 \right]^{1/2} \quad (4)$$

$$u_{ik}^{(r+1)} = \begin{cases} 1 & d_{ik}^{(r)} = \min_j d_{jk}^{(r)} \text{ for all } j \in c \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

여기서 d_{ik} 는 k 번째 데이터 샘플 \mathbf{x}_k 와 i 번째 클러스터 중심 $\vec{\nu}_i$ 의 기하학적 거리

Step 4: 오차가 허용범위 안에 도달하면 종료하고, 그렇지 않으면 $r = r + 1$ 로 놓고 **Step 2**로 이동

$$\|U^{(r+1)} - U^{(r)}\| \leq \epsilon (\text{tolerance level}) \quad (6)$$

3.3 K-means 클러스터링 기반 RBFNNs 패턴분류기 구조

본 실험에서는 전력기기와 절연 시스템 등에서의 온도에 의한 열적 스트레스, 전기장에 의한 전기적 스트레스, 운전 조건에 따른 열화 스트레스, 진동 등에 의한 기계적 스트레스로 발생할 수 있는 다양한 종류의 부분 방전 중에서 보이드, 코로나, 표면, 슬롯 방전을 선별하여 모의실험을 진행하였다[1, 5].

본 논문에서는 기존의 RBFNNs과는 달리 조건부에 K-means 클러스터링을 사용하고 연결가중치도 확장된 개념으로 이용한다. PCA를 이용하여 n 차원의 특징 데이터 $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ 가 입력층을 통해 은닉층에 입력되고 입력변수와 대응하는 연결가중치 $\nu_i = [\nu_{i1}, \nu_{i2}, \dots, \nu_{in}]$ 로 유클리드 거리를 계산한다. 식은 (7)과 같다.

$$d_i = \|\mathbf{x} - \nu_i\| \quad (7)$$

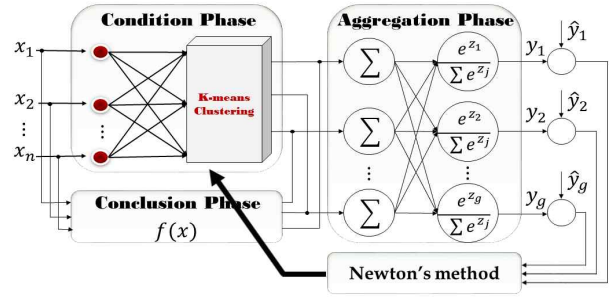


그림 2 제안된 소프트맥스 신경회로망 구조

Fig. 2 Architecture of proposed softmax neural networks

노드마다 전체 노드 수만큼의 $d_i (i = 1, 2, \dots, c)$ 가 생성된다. 입력층과 은닉층의 연결가중치 $v_{pi} (p = 1, 2, \dots, n)$ 는 K-means 클러스터링 목적함수의 중심점과 같은 의미를 가지며 소속행렬 $u_i(x)$ 는 가우시안 함수를 통한 은닉층 각 노드의 출력이 되고 식은 (8)과 같다.

$$u_i(x) = e^{-t} \quad (8)$$

$$t = \sum_{j=1}^c \frac{d_i^2 / d_j^2}{\sigma_j^2} \quad (9)$$

가우시안 함수를 이용하여 0과 1 사이의 수로 출력된다. 이는 입력변수와 은닉층 노드 사이의 적합도를 의미한다. 은닉층과 출력층의 연결가중치 $w_i^j (j = 1, 2, \dots, g)$ 는 1차 선형식을 사용하였고 식은 (10)과 같다.

$$w_i^j = f_i^j(x) = a_{i0}^j + \sum_{p=1}^n a_{ip}^j x_p \quad (10)$$

파라미터는 계수 a_{ip}^j 는 뉴턴법을 이용한 비선형 최소자승법을 통해 추정한다. 일반적으로 선형 최소자승법을 적용하여 오차제곱 합이 최소가 되는 파라미터 학습법이 불가능하다. 따라서 크로스 엔트로피 오차 함수를 이용하여 파라미터를 구한다. 여기서, 전반부 출력층의 z_j 는 은닉층 출력 $u_i(x)$ 와 연결가중치 w_i^j 에 의해 식 (11)로 정리된다.

$$z_j = \sum_{i=1}^c u_i w_i^j \quad (11)$$

후반부는 전반부 출력층에서 구한 각 클래스의 출력을 소프트맥스 함수로 정규화해준다. 이 출력된 값을 크로스 엔트로피 오차 함수에 적용하여 파라미터 계수를 추정한다.

3.4 후반부 파라미터 동정

가. 뉴턴법을 이용한 비선형 최소자승법

본 연구에서는 연결가중치의 학습 및 파라미터 계수를 추정하

기 위해 뉴턴법[Newton's method]을 사용한 비선형 최소자승법을 이용한다. 제안된 분류기의 비용함수는 실제 출력과 모델출력의 오차 제곱 합을 최소화하여 근사해를 찾는 일반적인 선형 최소자승법이 아닌 크로스 엔트로피 오차 함수(Cross-entropy Error function, CE)를 이용하였다. 다수의 클래스로 구성된 데이터를 분류하기 위해서는 변형된 CE를 사용해야 하며 다음 식 12와 같다[6].

$$CE = - \sum_{k=1}^N \sum_{j=1}^q t_{jk} \ln y_{jk} \quad (12)$$

여기서 t_{jk} ($j=1, 2, \dots, c; k=1, 2, \dots, N$)은 클래스의 출력을 의미하고 y_{jk} 은 패턴 분류기의 출력을 의미한다.

일반적으로 비선형 최소자승법은 반복되는 학습을 통해 근사해를 찾아가는 학습방법으로 가우스-뉴턴법이 많이 사용된다. 이 방법도 비용함수가 오차 제곱 합인 경우에만 가능하므로 제안된 분류기의 연결 가중치를 학습하기 위한 기준과는 다른 접근 방법이 필요하다. 대안으로 방정식 $f(x)=0$ 을 만족하는 해를 반복적으로 즉, 해의 추정값으로부터 반복적으로 참값을 구하는 뉴턴법을 이용하는 학습방법이다. 방정식 $f(x)$ 의 1차 미분방정식 $f'(x)$ 와 2차 미분방정식 $f''(x)$ 을 이용하여 $f'(x)=0$ 을 만족하는 최솟값을 찾는 것이다.

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)} \quad (13)$$

따라서 식 13을 응용하여 $CE'(A(r))$ 의 최솟값을 찾는 방법으로 접근하면 아래 식 14와 같다.

$$A(r+1) = A(r) - \frac{CE'(A(r))}{CE''(A(r))} \quad (14)$$

여기서 r 은 반복횟수를 의미하고, 1차 선형식의 파라미터 계수는 $A = [a_{10}^1, \dots, a_{d0}^1, a_{c1}^1, \dots, a_{cn}^1, \dots, a_{10}^g, \dots, a_{d0}^g, a_{c1}^g, \dots, a_{cn}^g]^T$ 이다.

연결가중치의 초기 파라미터 계수는 일정 범위 내에서 무작위값으로 설정하고 반복 학습을 거쳐 최적의 파라미터로 추정된다. 전체 파라미터 계수 A 를 추정하기 위해서 편미분 방정식을 이용한다. 학습하는 클래스 이외의 나머지 클래스는 상수항으로 간주하여 미분 시 소거되기 때문에 실질적으로 파라미터 계수를 각 클래스에 대응하는 노드별 출력을 추정할 수 있다. 그러므로 j 번째 출력 노드와 연결된 연결가중치의 파라미터 계수 $A_j = [a_{10}^j, \dots, a_{d0}^j, a_{c1}^j, \dots, a_{cn}^j]^T$ 을 구하기 위해 행렬로 표현하면 다음 식 15와 같다.

$$A_j(r+1) = A_j(r) - (X^T R_j X)^{-1} X Y_j \quad (15)$$

위 식을 $(X^T R_j X)^{-1}$ 로 묶어서 정리하면 선형 최소자승법과

비슷한 형태로 정리된다.

$$\begin{aligned} A_j(r+1) &= (X^T R_j X)^{-1} \{ (X^T R_j X) A_j(r) - X^T Y_j \} \\ &= (X^T R_j X)^{-1} X^T R_j \{ X A_j(r) - R_j^{-1} Y_j \} \\ &= (X^T R_j X)^{-1} X^T R_j Z_j (\because Z_j = X A_j(r) - R_j^{-1} Y_j) \end{aligned} \quad (16)$$

여기서,

$$Y_j = \begin{bmatrix} y_{j1} \\ y_{j2} \\ \vdots \\ y_{jN} \end{bmatrix}, \quad R_j = \begin{bmatrix} y_{j1}(1-y_{j1}) & 0 & \dots & 0 \\ 0 & y_{j2}(1-y_{j2}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & y_{jN}(1-y_{jN}) \end{bmatrix},$$

$$A_j(r) = [a_{10}^j(r), \dots, a_{d0}^j(r), a_{c1}^j(r), \dots, a_{1n}^j(r), \dots, a_{cn}^j(r)]^T,$$

$$X^T = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1N} \\ \vdots & \vdots & & \vdots \\ u_{c1} & u_{c2} & \dots & u_{cN} \\ u_{11}x_{11} & u_{12}x_{12} & \dots & u_{1N}x_{1N} \\ \vdots & \vdots & & \vdots \\ u_{c1}x_{11} & u_{c1}x_{12} & \dots & u_{cN}x_{1N} \\ \vdots & \vdots & & \vdots \\ u_{11}x_{n1} & u_{11}x_{n2} & \dots & u_{1N}x_{nN} \\ \vdots & \vdots & & \vdots \\ u_{c1}x_{n1} & u_{c1}x_{n2} & \dots & u_{cN}x_{nN} \end{bmatrix}$$

$A_j(r)$ 는 후반부 다항식 파라미터 계수, Y_j 는 출력 데이터, X^T 는 K-means 클러스터링의 소속행렬이 곱해진 입력데이터의 행렬로 x 는 입력행렬, u 는 소속행렬을 의미한다. 여기서 N 은 데이터의 수이며, n 는 입력변수의 수, c 는 클러스터 수가 된다. 위와 같은 행렬식을 이용하여 각 클래스에 대응하는 출력층 노드의 연결가중치를 학습하며 비용함수 $CE(r)$ 와 $CE(r+1)$ 값이 같거나 차이가 일정 값 이하일 때 학습을 종료한다.

나. 소프트맥스 함수

후반부 출력층에 소프트맥스 함수를 사용함으로써 확률값이 출력되고 전체 합계가 1이 된다. 이 출력값은 CE에 적용하여 실제 클래스의 출력과 학습되는 클래스 사이의 출력값을 가지고 0으로 수렴할수록 패턴 분류율이 증가하는 반비례 관계가 형성되게 한다. 소프트맥스 함수의 최종 출력은 다음 식 17과 같다.

$$y_j = \frac{e^{z_j}}{\sum_{q=1}^g e^{z_q}} \quad (17)$$

여기서 z_j 는 은닉층 출력 $u_i(x)$ 과 연결가중치 w_i^j 로 이루어진 전반부 출력값이다.

다. L2 정규화 방법에 의한 효과

과적합은 모델의 구조 및 데이터의 특성에 따라 발생할 수 있

다. 머신러닝에서 흔히 일어나는 문제이며 주로 모델을 설계할 때, 한정적인 데이터를 가지고 학습하면 데이터의 패턴에 너무 특화되어 학습된 나머지 테스트 분류율이 떨어지는 현상을 말한다. 모집단에 가까운 학습 데이터로 구성되면 과적합을 줄일 수 있겠지만 새로운 자료를 수집하고 분석하는데 있어서는 많은 시간과 비용이 들기 때문에 불가능에 가깝다. 과적합을 바라보는 다른 해석으로 학습 시에 오차 함수의 값이 지역 극소점에 갇힌 상황이라고 정의하기도 한다. 과적합은 연결가중치의 수가 클수록 발생할 가능성이 큰데 단순히 연결가중치의 수를 낮추는 접근은 바람직하지 않은 방법이다. 이유는 연결가중치는 데이터의 특성과 분류기 구조에 따라 변하기 때문에 설계하는 과정에서 해결하기란 쉽지 않다. 본 논문에서는 이러한 문제를 해결하기 위한 해결책으로 크로스 엔트로피 오차 함수에 L2 정규화를 사용하여 페널티 항을 추가했다. 그러면서 연결가중치의 편차를 줄이고 동시에 테스트 데이터의 일반화 성능을 향상하며 과적합 문제를 개선하였다. 아래 식 18은 크로스 엔트로피 함수에 L2 페널티 항을 추가한 새로운 비용함수이다[7, 8].

$$CEL = - \sum_{k=1}^N \sum_{j=1}^g t_{jk} \ln y_{jk} + \lambda \sum_{j=1}^g \sum_{i=1}^c \sum_{p=0}^n (a_{ip}^j)^2 \quad (18)$$

여기서 λ 는 정규화 매개변수로 0.01에서 0.00001 정도의 범위 내에서 선택한다. 위의 식을 편미분방정식을 통해 행렬로 표현하면 다음 식 19와 같다.

$$A^j(r+1) = A^j(r) - (X_j^T R_j X_j + \lambda I)^{-1} (X_j^T Y_j + \lambda A^j(r)) \quad (19)$$

여기서 $(X_j^T R_j X_j + \lambda I)^{-1}$ 로 묶어서 정리하면 선형 최소자승법과 유사한 형태로 정리된다.

$$\begin{aligned} A^j(r+1) &= (X_j^T R_j X_j + \lambda I)^{-1} \{ (X_j^T R_j X_j + \lambda I) A^j(r) - (X_j^T Y_j + \lambda A^j(r)) \} \\ &= (X_j^T R_j X_j + \lambda I)^{-1} X_j^T R_j Z_j \quad (\because Z_j = X_j A^j(r) - R_j^{-1} Y_j) \end{aligned} \quad (20)$$

4. 시뮬레이션 및 결과 고찰

제안된 K-means 클러스터링 기반 소프트맥스 신경회로망 부분방전 패턴분류의 설계와 비교연구 및 해석을 위하여 두 가지 데이터를 사용하였다. 데이터는 센서에 탑재되어있는 PRPD의 분석을 통하여 데이터가 저장되고 MIMS 프로그램을 통하여 데이터를 추출한다. 패턴의 특징은 크게 두 가지로 나뉜다. 진폭의 크기(Amplitude, AMP)와 초당 진동수(Pulse Per Second, PPS)가 있는데, 여기서 중점된 AMP의 정보를 최댓값과 평균값으로 설정하여 뽑은 데이터를 가지고 실험하였다. 본 연구에서 선택한 부분방전의 종류는 보이드, 코로나, 표면, 슬롯 방전이고 각각 250개, 435개, 270개, 245개로 총 1,200개의 데이터로 구성된다[2].

4.1 부분방전 모의실험

가. 부분방전 선정 및 모의결합

본 연구에서는 현장에서 주로 일어나는 부분방전을 근거하여

4가지 모의 결함을 선정하여 실험을 진행했다.

1) 보이드(void)

외부의 자극으로 도체와 절연물 사이에 유전율이 낮은 불순물이 생기거나 노후화로 절연체 내에 공극이 존재하게 된다. 이것이 원인이 되어 한쪽으로 전계가 집중되었다가 일정 값 이상이 되면 전리가 일어나 절연 파괴가 일어난다.

2) 코로나(corona)

기체나 도체의 가장자리 외(外) 불평등한 전계에 의해 잘 발생하는 부분방전이다. 공기분자로부터 계속해서 충돌이 일어나 이온화가 되고, 동시에 강한 산화력을 갖는 오존을 발생시켜 열화를 일으킨다.

3) 연면(surface)

일반적으로 표면 방전이라 칭하고 전극 끝에 공기와 고체 절연물이 조합되어 있는 경우에 공기의 절연내력은 고체 절연내력보다 낮고, 유전율도 공기보다 낮으므로 전압이 높아진다. 이때, 어느 전압에 이르렀을 때 전극 끝의 공기 부분에서 절연물 표면을 따라 불꽃과 함께 방전이 발생한다.

4) 슬롯 방전(Slot discharge)

권선 표면의 코일, 웨지와 슬롯의 상대적인 움직임으로 마모되어 반도체 전 층을 손상해 국부적으로 권선의 접지상태가 나빠진다. 장기간 열화 과정에서 트리잉(treeing)이 생기고 결과적으로 결함 부분에 지속적인 방전이 일어나 심각한 훼손을 초래한다.

나. 모의실험을 통한 데이터 구축 및 시스템 구성

본 실험에서는 슬롯에 권선을 감아서 고압 전류를 흘려주고 이때 발생하는 부분방전의 패턴을 분석하며 데이터를 구축하였다. 센서의 PRPDA로 정돈된 정보는 MIMS 프로그램을 통하여 다듬어지고 이를 분류기로 보내어 분석한다. 아래 그림 3은 실험을 진행한 개념도이다.

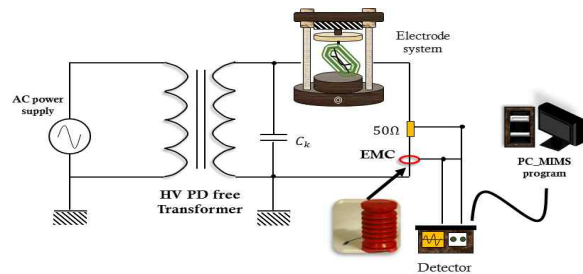


그림 3 EMC에 의해 측정된 시스템 개념도

Fig 3 Concept Map of system measured by EMC

다. PRPDA를 통한 데이터 전처리

본 논문에서는 공학 분야에서 많이 사용하고 일반화된 전처리 알고리즘 PRPDA가 사용된다. 실험을 통해 15초 동안 진동수가

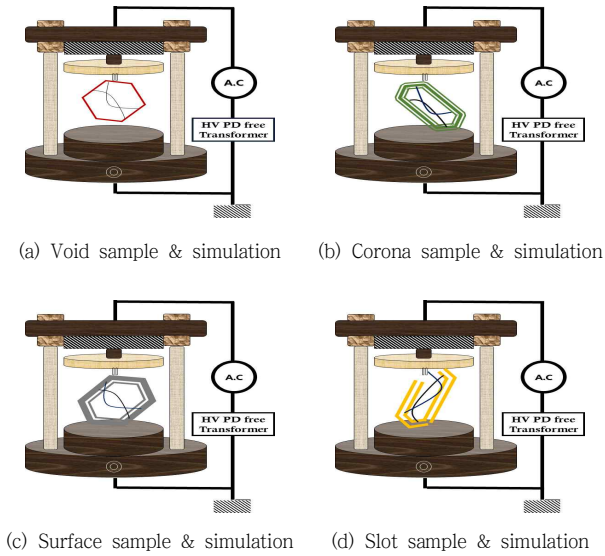


그림 4 각각의 방전 시료 및 모의실험

Fig. 4 Each discharge sample and imitating experiment

60Hz 고전압을 인가하면서 발생하는 부분방전 신호가 센서에 입력되는데, 방전 크기 $q(\phi)$ 와 방전 수 $n(\phi)$ 의 정보가 중첩되어 쌓인다.

4.2 AMP를 최댓값과 평균값으로 설정하여 데이터 구성

센서로부터 받은 부분방전 신호의 전체 위상이 360°이며 2° 간격으로 AMP와 PPS 정보가 저장된다. 이 정보를 이용하여 MIMS 프로그램을 통하여 데이터가 추출된다. 아래 식 21은 하나의 위상 구간에서 측정된 방전 q_i 의 총 합이다.

$$q_s = \sum q_i \quad (21)$$

여기서 q_s 는 총 방전량의 합을 의미한다. 하나의 위상구간에서 저장된 q_s 를 가지고 최대 방전크기(q_{max})와 평균 방전크기(q_{avg})를 구한다.

$$q_{max} = MAX(q_s) \quad (22)$$

$$q_{avg} = AVERAGE(q_s) \quad (23)$$

식 (22)와 (23)을 통해 구한 부분방전은 아래 그림 5와 같이 패턴과 특징을 나타내어 비교해보았다.

Negative(0° ~ 90°)구간과 Positive(180° ~ 270°)구간을 포개어 분석한 분포도이다. 외관상으로 최댓값 분포도와 평균값 분포도의 다른 특징을 보이고 특히 AMP의 크기 폭이 확연한 차이를 나타낸다. 먼저 최댓값 패턴 보이드의 경우, 파형은 Negative구간과 Positive구간 모두 균일하게 밀집되어 있고 특정 구간을 넘어서는 급격하게 떨어지는 모양을 보인다. 반면 평균값

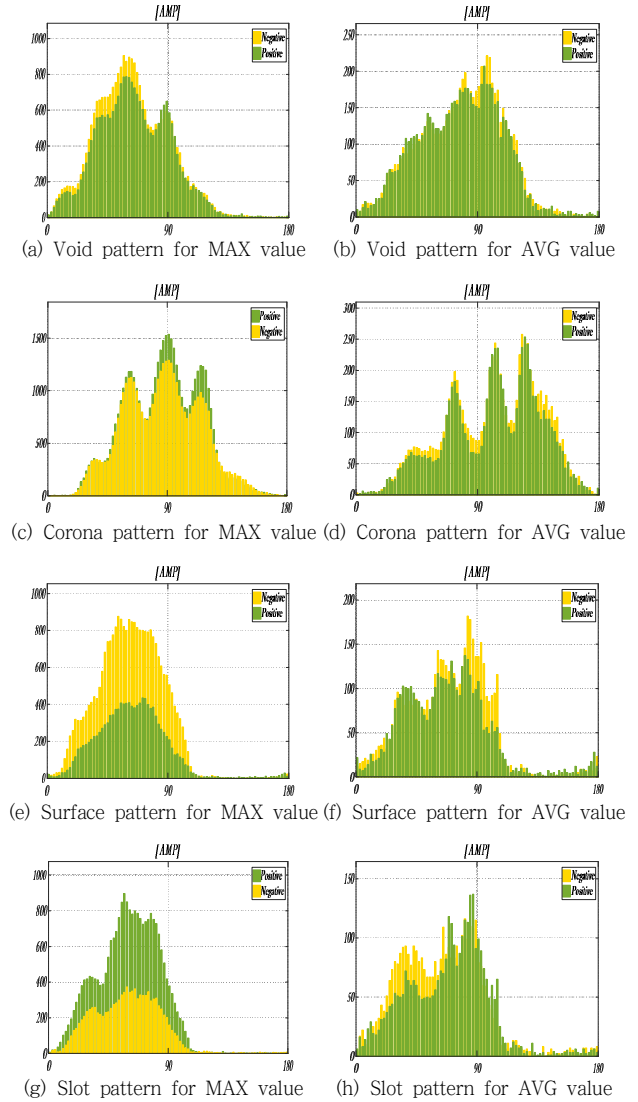


그림 5 각 부분방전 최댓값과 평균값 패턴 특징 비교

Fig. 5 Pattern feature comparison of each partial discharge of maximum and average value

은 Negative와 Positive 구간을 벗어났고 상대적으로 완만한 산모양을 이루는 것을 확인할 수 있다. 코로나의 최댓값 패턴은 균등한 산자 모양을 나타내지만, 평균값 패턴은 우측으로 기울어진 산자를 이루고 있다. 최댓값 패턴에서 표면방전은 Positive구간보다 Negative구간이 더 큰 분포도를 보이지만 슬롯 방전에서는 정반대의 양상을 보인다. 하지만 평균값 패턴에서는 표면과 슬롯 두 파형의 특징이 비등하여 육안으로는 구분하기 힘들음을 보였다. 오히려 보이드 패턴의 모양과 비슷한 모습을 보인다.

4.2 입력 데이터 구성에 대한 실험 설정 및 결과

제안하는 패턴 분류기의 비교 모델로 K-means 클러스터링을 이용한 방사형 기저 함수 기반 신경회로망(Radial Basis

Function-based Neural Networks, RBFNNs)을 사용하였다. 두 분류기의 가장 큰 차이가 비교 모델의 비유함수로 오차제곱 합을 사용하는 것이다. 제안된 패턴 분류기와 비교 모델의 구조 및 학습방법은 다음과 같다. 두 분류기의 은닉층은 같은 구조지만, 비교 모델은 출력층에 소프트맥스 함수를 사용하지 않고 은닉층 출력과 연결가중치의 선형결합을 직접 분류기의 출력으로 사용하며 선형 최소자승법을 통해 파라미터를 학습하는 차이를 보인다. 최종적으로 과적합으로 생길 수 있는 성능저하를 방지하기 위해 두 분류기 모두 L2 정규화 방법을 사용하여 개선 효과가 있는지 확인하였다.

표 1 비교 모델의 구조 및 학습 방법

Table 1 Structure & learning method of comparative model

Models	RBFNNs	Proposed classifier
Layers		
Learning method	<i>LLSM</i>	<i>NLSM</i>
Cost function	$SSE = \sum_{k=1}^N \sum_{j=1}^c (t_{jk} - y_{jk})^2$	$CE = - \sum_{k=1}^N \sum_{j=1}^c t_{jk} \ln y_{jk}$
L2 Regularization	Used	Used
Output later - output	$z_j = \sum_{i=1}^c u_i w_i^j$	$y_j = \frac{e^{z_j}}{\sum_{q=1}^c e^{z_q}}$

LLSM: Linear least squares method, *NLSM*: Nonlinear least squares method, *SSE*: Sum of squared error, *CE*: Cross entropy error function

표 2 부분방전 실험 조건

Table 2 Experimental condition of partial discharge data

Parameters		Values			
PD data	Class No.	1 (void)	2 (corona)	3 (surface)	4 (slot)
	No. of inputs	360	360	360	360
	No. of data	250	435	270	245
Total number of data		1200			
Ratio of training & test		4:1			
Number of clusters		3, 5			
Learning method		Nonlinear using Newton's method			

본 연구에서는 부분방전 데이터로 보이드 250개, 코로나 435개, 표면 방전 270개, 슬롯 방전 245개로 구성되어있는 총 1,200개의 데이터를 사용한다. 학습과 시험의 비율은 4:1로 각각 학습 데이터 960개 시험데이터 240개이다. 클러스터 수는 [3 5]로 통상 많은 실험을 통한 분류율을 잘 나타냈던 값으로 선정하였고, 각 데이터의 입력변수는 360개이지만 빠르고 정확한 학습을 위하여 주성분 분석법을 통해 [9 13] 입력변수를 축소하여 실험을 진행하였다. 각 특징 데이터의 분류율을 아래 표와 같다.

표 3 최댓값 데이터의 패턴 분류율 비교

Table 3 Comparison of pattern classification of Max data

Model		RBFNNs				Proposed Classifier				
R	D	λ	<i>CR_TR</i>		<i>CR_TE</i>		<i>CR_TR</i>		<i>CR_TE</i>	
			Mean	STD	Mean	STD	Mean	STD	Mean	STD
3	9	0	93.81	0.99	92.41	2.30	99.91	0.04	97.41	0.61
		0.01	93.52	0.90	91.75	1.39	99.43	0.05	98.83	0.74
		0.001	93.50	0.86	91.75	1.39	99.83	0.05	98.75	0.51
	13	0	96.35	0.70	95.75	1.07	100.0	0.00	97.00	1.07
		0.01	96.35	0.70	95.75	1.07	99.85	0.05	98.75	0.77
		0.001	96.35	0.70	95.75	1.07	99.95	0.05	99.41	0.22
5	9	0	97.29	0.79	96.08	1.65	100.0	0.00	95.16	0.69
		0.01	97.31	0.79	96.08	1.65	99.56	0.08	99.00	0.47
		0.001	97.29	0.79	96.08	1.65	99.87	0.04	99.00	0.22
	13	0	98.39	0.13	97.08	0.65	100.0	0.00	97.25	1.90
		0.01	98.39	0.13	97.08	0.65	99.93	0.05	98.66	0.54
		0.001	98.39	0.13	97.08	0.65	100.0	0.00	99.41	0.69

R: Number of rules, D: Number of reduced dimensions using PCA, λ : regularization parameter, *CR_TR*: Classification rate of training data, *CR_TE*: Classification rate of test data

표 4 평균값 데이터의 패턴 분류율 비교

Table 4 Comparison of pattern classification of Avg data

Model		RBFNNs				Proposed Classifier				
R	D	λ	<i>CR_TR</i>		<i>CR_TE</i>		<i>CR_TR</i>		<i>CR_TE</i>	
			Mean	STD	Mean	STD	Mean	STD	Mean	STD
3	9	0	88.91	1.55	87.16	2.75	95.95	0.69	90.50	2.30
		0.01	89.75	1.26	88.50	1.90	95.33	0.58	92.75	1.33
		0.001	89.75	1.26	88.50	1.90	95.83	0.77	91.25	1.14
	13	0	92.31	1.28	90.83	1.06	98.70	0.72	90.91	2.45
		0.01	92.31	1.28	90.83	1.06	97.58	0.25	92.41	1.62
		0.001	92.31	1.28	90.83	1.06	98.85	0.32	93.41	1.39
5	9	0	92.12	0.89	89.33	1.54	97.41	1.06	89.75	2.31
		0.01	92.12	0.89	89.33	1.54	95.93	0.23	92.75	1.83
		0.001	92.12	0.89	89.33	1.54	97.50	0.67	93.08	2.31
	13	0	94.39	0.94	92.00	1.59	99.77	0.13	89.33	1.68
		0.01	94.39	0.94	92.00	1.59	98.52	0.24	95.00	1.69
		0.001	94.39	0.94	92.00	1.59	99.22	0.35	93.58	0.95

실험조건은 동일하게 클러스터 수를 3, 5로 고정해주고 주성분 분석법을 통하여 입력변수를 9, 13차원으로 축소하였다. L2 정규화를 적용해 범위를 [0.01 0.001]로 주었다. 그 결과 전반적으로 λ 를 추가하였을 때 개선된 분류율을 보여준다.

실험결과 제안된 분류기가 기존의 RBFNNs의 분류율보다 확연하게 개선된 모습을 보여준다. 또 L2 정규화를 적용하고 λ 값을 추가하였을 때, 더 높은 분류율을 보여준다. 두 데이터 모두 λ 값이 0.01일 때가 가장 높은 분류율을 보이며 최댓값 데이터가 평균값 데이터의 분류율보다 나은 모습을 확인하였다.

표 5 최댓값 데이터에 대한 각 fcv의 분류율

Table 5 Classification rate of each fcv for Max data

Model			Proposed Classifier		
R	D	λ	5-fcv	CR_TR(CE)	CR_TB(CE)
3	13	0	1st	100.0 (0.0032)	96.66 (0.1239)
			2nd	100.0 (0.0046)	98.33 (0.0558)
			3rd	100.0 (0.0037)	95.41 (0.2188)
			4th	100.0 (0.0040)	97.08 (0.0933)
			5th	100.0 (0.0043)	97.50 (0.0386)
			Mean±STD	100.0±0.00 (0.0040±0.0005)	97.00±1.07 (0.1061±0.0712)
	0.001	1st	100.0 (0.0325)	99.58 (0.0442)	
		2nd	99.89 (0.0299)	99.58 (0.0664)	
		3rd	100.0 (0.0294)	99.58 (0.0331)	
		4th	99.89 (0.0295)	99.16 (0.0563)	
		5th	100.0 (0.0279)	99.16 (0.0586)	
Mean±STD	99.95±0.05 (0.0299±0.0016)	99.41±0.22 (0.0517±0.0131)			

표 6 평균값 데이터에 대한 각 fcv의 분류율

Table 6 Classification rate of each fcv for Avg data

Model			Proposed Classifier		
R	D	λ	5-fcv	CR_TR(CE)	CR_TB(CE)
5	13	0	1st	99.68 (0.0614)	88.75 (0.7821)
			2nd	99.58 (0.0879)	90.41 (0.4137)
			3rd	99.79 (0.0521)	86.66 (0.8323)
			4th	99.89 (0.0364)	90.00 (0.8382)
			5th	99.89 (0.0143)	90.83 (1.2486)
			Mean±STD	99.77±0.13 (0.0504±0.0275)	89.33±1.68 (0.8230±0.2961)
	0.01	1st	98.64 (0.1407)	95.00 (0.1759)	
		2nd	98.64 (0.1358)	97.08 (0.1619)	
		3rd	98.75 (0.1233)	95.83 (0.2255)	
		4th	98.43 (0.1277)	94.58 (0.2375)	
		5th	98.12 (0.1297)	92.50 (0.2308)	
Mean±STD	98.52±0.24 (0.1315±0.0068)	95.00±1.69 (0.2063±0.0347)			

표 5와 6은 5-fold cross validation을 적용하여 fold마다 성능의 변화를 나타낸 것이다. 특히 L2 정규화를 적용하지 않았을 때, 평균값 데이터의 학습 분류율이 100% 가까이 나오지만 실험 데이터의 경우 fcv의 세 번째에서 86.66%로 가장 낮은 분류율을 나타낸다. 이와 같은 현상이 과적합이다. 제안된 학습방법을 통해 각 세대의 비용함수와 패턴 분류율 수렴과정을 그림 6과 7로 확인한다. 그래프를 통해서도 L2 정규화 방법을 사용 여부의 수렴 과정 차이가 확연히 있음을 보인다. 두 데이터 모두 학습 데이터의 비용함수는 세대가 지날수록 0으로 수렴하는 모습을 보이지만 일정 세대 이후부터 실험 데이터가 증가하다 발산된다. 패턴 분류율도 학습 데이터는 100%로 수렴하지만, 비용함수의 실험 데이터가 증가하는 세대 이후로 실험 데이터의 패턴 분류율이 떨어

지는 반비례 양상을 보인다. 그러나 L2 정규화 방법을 적용하면 비용함수가 같은 범위로 수렴 하고 패턴 분류율도 함께 개량된 것을 확인하였다.

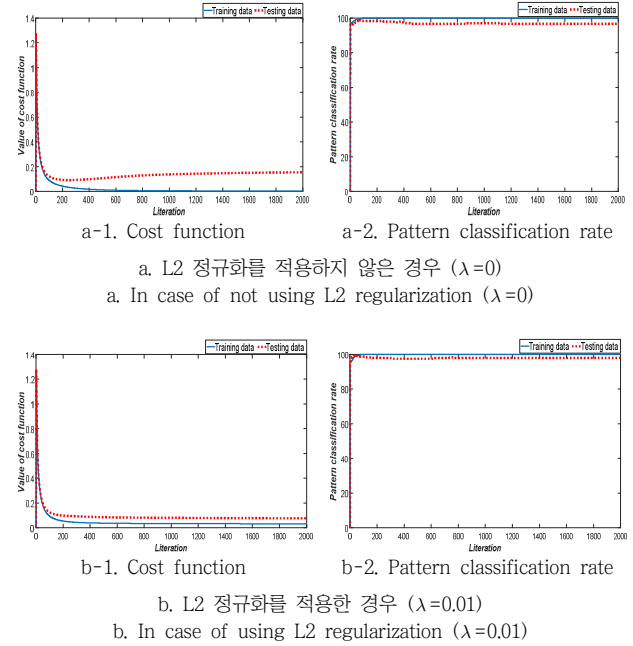


그림 6 비용함수와 분류율 비교(최댓값)

Fig 6 Comparison of cost function & recognition rate(Max)

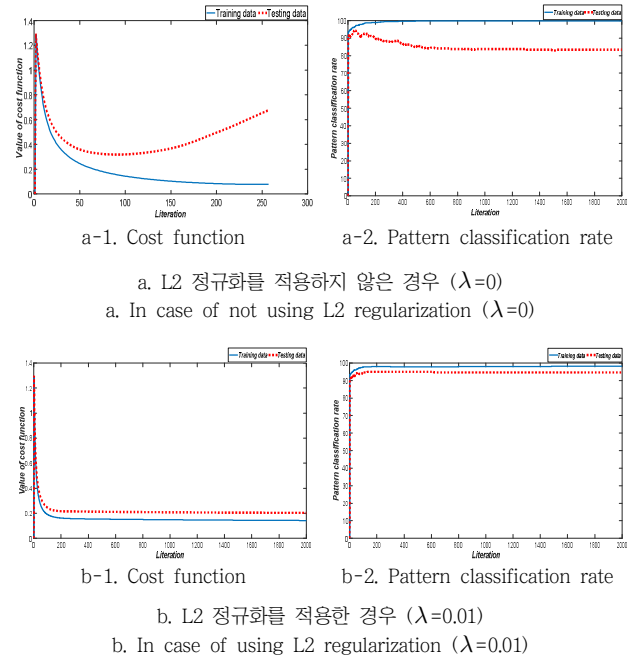


그림 7 비용함수와 분류율 비교(평균값)

Fig. 7 Comparison of cost function & recognition rate (Avg)

표 7 각 데이터별 모델에 따른 선별된 분류율 비교

Table 7 Comparison of classification rate selected according to Models of each dataset

Models	T	R	D	CR_TR		CR_TE	
				Mean	STD	Mean	STD
Neural Network[9]	Max	110	X	95.23	1.62	93.33	1.72
	Avg	25	X	88.56	3.55	84.75	4.19
Support Vector Machine[10]	Max	X	40	100.0	0.0	96.25	1.06
	Avg	X	60	97.77	0.15	93.75	0.65
RBFNNs without PSO[11]	Max	4	50	99.50	0.13	98.91	0.75
	Avg	4	60	99.00	0.28	95.66	1.12
RBFNNs with PSO[11]	Max	3	75	99.86	0.07	98.95	0.88
	Avg	3	60	99.20	0.29	96.31	1.87
Proposed Models	Max	3	13	99.95	0.05	99.41	0.22
	Avg	5	13	98.52	0.24	95.00	1.69

T : Type of Data

표 7은 다른 기존의 다양한 분류기와 비교 정리한 것이다. 먼저 Back Propagation을 이용하여 학습하는 순수 신경회로망과 널리 알려진 Support Vector Machine 분류기, 그리고 최적화 알고리즘 Particle Swarm Optimization를 이용한 RBFNNs 분류기를 사용하였다. 최적화 알고리즘을 적용한 모델의 전반부 구조 동정은 K-means 클러스터링이 아닌 Fuzzy C-Means 클러스터링을 적용하였다. 표 7은 각 분류기를 실험을 통하여 데이터별로 가장 우수한 성능을 보이는 파라미터 값을 사용하여 나타난 것이다.

5. 결 론

본 연구에서 부분방전 신호와 K-means 클러스터링 기반 소프트웨어 신경회로망 패턴분류기를 이용하여 부분방전의 패턴을 비교·분석하고 전력기기 열화를 진단한다. 분류기의 구조는 입력층, 은닉층, 출력층이며 비용함수로 크로스 엔트로피 함수를 사용하였고 연결 가중치는 뉴턴법을 이용한 비선형 최소자승법을 통하여 업데이트하였다. 또한, 과적합을 줄이기 위하여 L2 정규화를 적용하여 성능 개선을 입증하였다. 데이터는 두 가지 특징 데이터를 사용하여 비교하였고, 제안된 분류기에서 모두 높은 분류율을 보였다. 두 가지 데이터 중 최댓값 데이터가 더 높은 분류율을 보이는데, 이는 패턴의 특징을 통해서 검증하였다. 다른 기존의 분류기와 비교하여 제안된 분류기가 평균값 데이터에서는 분류율이 1% 정도 뒤지지만, 최댓값 데이터 부분에서는 1% 향상되었음을 확인하였다. 뉴턴법을 이용한 비선형 최소자승법을 통해 더욱 정밀한 학습이 이루어지고 소프트웨어 함수를 통하여 속도 면에서 더 빠른 계산 작업을 나타냄을 보이며 제안하는 분류기의 우수성을 입증하였다.

감사의 글

This search was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (NRF-2017R1D1A1B03032333) and by R&D Support Project of Innovative Enterprise Technology Development Program for Small and Medium-sized Businesses in 2017

References

- [1] S. M. Hassan Hosseini, S. M. Hosseini Bafghi, "Comparison of High Frequency Detailed Generator Models for Partial Discharge Localization", Journal of Electrical Engineering & Technology 10, 1752-1758, 2015.7.
- [2] Lianwei Bao, Jian Li, Jing Zhang, Tianyan Jiang, "Partial Discharge Process and Characteristics of Oil-Paper Insulation under Pulsating DC Voltage", Journal of Electrical Engineering & Technology 11, 437-445, 2016. 03.
- [3] Furqan Asghar, Muhammad Talha, Se-Yoon Kim, SungHo Kim, "Hotelling T2 Index Based PCA Method for Fault Detection in Transient State Processes", Journal of Institute of Control, Robotics and Systems 22, 276-280, 2016.4.
- [4] Sung-Kwun Oh, Wook-Dong Kim, Witold Pedrycz, Su-Chong Joo, "Design of K-means clustering-based polynomial radial basis function neural networks (pRBF NNs) realized with the aid of particle swarm optimization and differential evolution", Neurocomputing, Vol. 78, No. 1, pp. 121-132, 2012.
- [5] Mingwen Zheng, Yanping Zhang, "An Algorithm to Determine RBFNN's Center Based on the Improved Density Method", Open Journal of Applied Sciences, 1-5, 2014. 04
- [6] John E. Dennis, David M. Gay, Roy E. Walch, "An Adaptive Nonlinear Least-Squares Algorithm", ACM Transactions on Mathematical Software, Vol. 7, No. 3, pp. 348-368, 1981.
- [7] Xiong Luo, Xiao hui Chang, Xiaojuan Ban "Regression and classification using extreme learning machine based on L1-norm and L2-norm", Neurocomputing, Vol 174, Part A, pp. 179-186, 2016.
- [8] Shenglei Chen, Geng Chen, Ruijun Gu, "An efficient L2-norm regularized least-squares temporal difference learning algorithm", Knowledge-Based Systems, Vol 45, pp. 94-99, 2013.

- [9] Kim, Joong-Hyo, Kwon, Sung-Dae, Hong, Jeong-Pyo, Ha, Tae-Jun, "Development of Optimum Traffic Safety Evaluation Model Using the Back-Propagation Algorithm", JOURNAL OF THE KOREAN SOCIETY OF CIVIL ENGINEERS 35(4), 2015.6, 679-690.
- [10] Jinju Hong, Sehan Kim, Jeawon Park, Jaehyun Choi, "A Malicious Comments Detection Technique on the Internet using Sentiment Analysis and SVM", Journal of the Korea Institute of Information and Communication Engineering 20, 260-267, 2016.2.
- [11] Byeong-Jin Jeong, Seung-Cheo Leel, Sung-Kwun Oh, "Design of Optimized Radial Basis Function Neural Networks Classifier Using EMC Sensor for Partial Discharge Pattern Recognition", The Korean Institute of Electrical Engineers 66(9), 1392-1401, 2017.9.

저 자 소 개



정 병 진 (Byeong-Jin Jeong)

2016년 : 수원대학교 전기공학과 졸업
2016년~현재 : 동 대학원 석사과정
관심분야 : 퍼지 이론, 컴퓨터 지능 및 지능 제어, 뉴럴 네트워크, 패턴 분류 등.
Phone: +82-31-222-6544
Fax : +82-31-220-2667
E-mail : qudwls2@suwon.ac.kr



오 성 권 (Sung-Kwun Oh)

1981년 : 연세대학교 전기공학과 공학사
1983년~1989년 : 금성산전연구소(선임연구원)
1993년 : 연세대학교 전기공학과 공학박사
1996년~1997년 : 캐나다 Manitoba 대학 전기 및 컴퓨터 공학과 Post-Doc
1993년~2004년 : 원광대학교 전기전자 및 정보공학부 교수
2005년~현재 : 수원대학교 전기전자공학부 교수
2002년~현재 : 대한전기학회 및 한국지능시스템학회 편집위원
2013년~현재 : Information Sciences 편집위원
관심분야 : 퍼지 시스템, 퍼지-뉴럴 네트워크, 자동화 시스템, 고급 Computational Intelligence, 지능제어 등.
Phone : +82-31-229-8162
E-mail : ohsk@suwon.ac.kr