

SVM(Support Vector Machine) 알고리즘 기반의 EEG(Electroencephalogram) 신호 분류

이상원¹, 조한진², 채철주^{3*}

¹대구대학교 과학교육학과 박사과정, ²극동대학교 에너지IT공학과 교수, ³한국농수산대학 교양공통과 교수

EEG Signal Classification based on SVM Algorithm

Sang-Won Rhee¹, Han-Jin Cho², Cheol-Joo Chae^{3*}

¹Doctoral Course, Department of Science Education, Daegu University

²Professor, Department of Smart & PhotoVoltaic Convergence, Far East University

³Professor, Department of General Education, Korea National College of Agriculture and Fisheries

요약 본 논문에서는 사용자의 EEG(Electroencephalogram)신호를 측정하여 SVM(Support Vector Machine) 알고리즘을 이용하여 EEG 신호를 분류하고 신호의 정확도를 측정하였다. 사용자의 EEG 신호를 측정하기 위해 남-여를 구분하여 실험을 진행하였으며, EEG 신호 측정은 단채널 EEG 디바이스를 이용하였다. EEG 디바이스를 이용하여 사용자의 EEG 신호를 측정한 결과는 R을 이용하여 분석하였다. 또한 SVM의 분류 성능이 최고가 되는 특정 벡터의 조합을 적용시켜 EEG 측정 실험 데이터를 80:20(훈련 데이터: 테스트 데이터) 비율로 예측해 본 결과 인식을 93.2%의 예측 정확도를 보였다. 본 논문에서는 사용자의 EEG 신호를 약 93.2% 정도로 인식할 수 있었으며, SVM 알고리즘의 간단한 선형 분류만으로 수행이 가능하다는 점은 EEG 신호를 이용하여 생체인증에 다양하게 활용될 수 있음을 제시하였다.

주제어 : EEG, SVM, 비지도 학습, 데이터 분류, 인증

Abstract In this paper, we measured the user's EEG signal and classified the EEG signal using the Support Vector Machine algorithm and measured the accuracy of the signal. An experiment was conducted to measure the user's EEG signals by separating men and women, and a single channel EEG device was used for EEG signal measurements. The results of measuring users' EEG signals using EEG devices were analyzed using R. In addition, data in the study was predicted using a 80:20 ratio between training data and test data by applying a combination of specific vectors with the highest classifying performance of the SVM, and thus the predicted accuracy of 93.2% of the recognition rate. This paper suggested that the user's EEG signal could be recognized at about 93.2 percent, and that it can be performed only by simple linear classification of the SVM algorithm, which can be used variously for biometrics using EEG signals.

Key Words : EEG, SVM, Unsupervised Learning, Data Classification, Authentication

*This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2017R1D1A1B03032876)

*Corresponding Author : Cheol-Joo Chae(Chae.cheoljoo@gmail.com)

Received January 7, 2020

Revised February 7, 2020

Accepted February 20, 2020

Published February 28, 2020

1. 서론

뇌에서 발생하는 신호의 진동인 뇌파와 관련된 연구는 1924년 Hans Berger가 인간 뇌파를 기록한 후 다양한 분야에서 뇌파를 다루게 되었고 생체인식 및 뇌파 분석 방법의 발달로 최근에도 활발하게 연구되고 있다[1]. 초기 연구는 뇌파의 요소들을 중심으로 한 학습과 기억에 관련된 주의(attention)와 각성(awakeness)에 관한 연구에 집중되었지만 최근 10여 년간의 활발한 연구로 다양한 신경계의 활동이나 정보 처리 등 다양한 인지기능 등과 관련된 연구가 진행되고 있다[2-7]. 이와 관련해서 4차 산업혁명의 발생과 이를 지원하는 정보산업사회가 고도화 되면서 정보로써 이를 처리하고 모델링이나 공학적인 구현을 하는 IT와 BT가 융합된 연구로 진행되고 있다. 최근에는 특히 뇌 정보 처리 및 모델링 연구에 있어 게임 분야나 제어 분야 외에도, 보안이나 식별과 관련한 분야에서 많이 응용되고 있다. 이러한 연구에서는 뇌파와 관련한 정보 처리를 위해 BCI(Brain Computer Interface) 기술이 활용되고 있다[8]. EEG 신호는 복잡한 파형의 전기적인 신호를 가지고 있으며, 주파수 대역별로 분류할 수 있다. EEG 신호는 각 대역마다 특성을 가지고 있고 사용자에 따라 다른 특성을 보인다[9]. 이는 자발적 전기신호로 μV 단위로 나타나는 미세한 신호를 분석한다. 이 주파수 대역들은 무의식 상태와 깨어있는 의식 상태를 반영하는데 최근에는 네트워크 기반 서비스에서 이루어지는 사용자 인증 과정에서 제3자에게 자신을 증명하는 과정으로 일반적으로 본인만이 소유하고 있거나 알고 있는 정보를 이용하여 사용자를 인증하는 방식으로 활용되고 있다[9].

본 논문에서는 EEG 센스에서 취득한 신호를 기반으로 SVM 알고리즘을 적용하여 실제 대상을 분류하고 이에 대한 정확도를 검증하기 하였다. 본 논문의 구성은 다음과 같다. 2장에서 EEG 인증 기술에 대해 기술하고 3장에서는 SVM을 이용한 EEG 신호 분류 방법에 대해 기술하였으며, 4장에서 사용자의 EEG 신호를 SVM 알고리즘을 이용한 실험 결과를 기술하였으며, 5장에서는 결론 및 향후 연구 방향에 대해 기술하였다.

2. EEG 인증 기술

EEG 신호는 뇌의 신경 사이에서 신호가 전달될 때 발생하는 미세 전기신호로써 복잡한 파형의 형태를 편의상

주파수 대역별로 Delta, Theta, Alpha, Beta, Gamma로 분류할 수 있다. 이러한 신호는 각 대역별 특성을 가지고 있으며, 사용자별로 다른 특성을 보이기 때문에 이 신호를 이용한 사용자의 인증 방법으로 사용이 가능하다[9]. 일반적인 EEG 신호 측정을 위해 두피와 귀 부분이 접지된 상태에서 Fig. 1에 표시된 20개 이상의 전극이 필요하나 최근 센서 기술의 발전으로 간편하게 측정할 수 있는 장치들이 개발되었다[10-13].

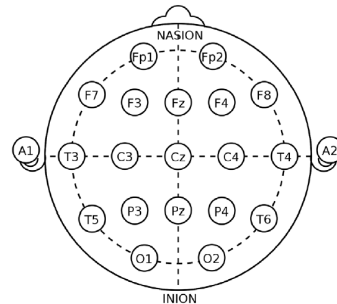


Fig. 1. Standard Electrode Map

Jasper가 제안한 전극 부착 위치를 참고하여 EEG 신호를 측정한다. Jasper가 제안한 방법은 10-20 국제 시스템(International 10-20 System)이라고 불리는 것으로, 기준 점간의 거리를 10% 또는 20% 떨어진 부위를 중심으로 측정 위치로 잡는다. Fig. 1에서 보는 바와 같이 FP는 전두엽을 나타내고 A1, A2는 좌우측 귓볼을 나타낸다. 측정되는 뇌파는 머리 전체 부위를 다채널로 측정할 수 있고, 보다 정확한 신호를 얻기 위해 위치별로 EEG 신호를 동시에 측정한다[14].

그러나 측정 장치의 발전으로 신호 수집에 필요한 전극 수는 감소하고, 측정된 신호는 블루투스 등의 무선 통신을 통한 전송 형태로 발전하였다. 대표적인 장비로는 Fig. 2에서 보는 바와 같이 EMOTIVE社의 EPOC+와 Neurosky社의 MindWave가 있는데, EEG 측정 전극의 수가 14개로 간소화 되었다[15].



Fig. 2. EEG Signal Collecting Devices

3. EEG 신호 분류 방법

3.1 SVM(Support Vector Machine)을 이용한 EEG 신호 분류

SVM 알고리즘에서는 학습 과정을 통해 사물 인식의 기준으로 사용되는 초평면(hyper-plane)을 결정하며, 초평면(hyperplane)을 기준으로 한쪽 면으로 동일한 데이터가 놓이게 평평한 경계를 만드는 것이다. Fig. 3에서 보는 것처럼 대각선 방향의 직선이 최대 여백 초평면에 해당한다. 그리고 이 직선과 가장 가까운 곳의 분류에 속한 점들을 서포트 벡터라고 한다.

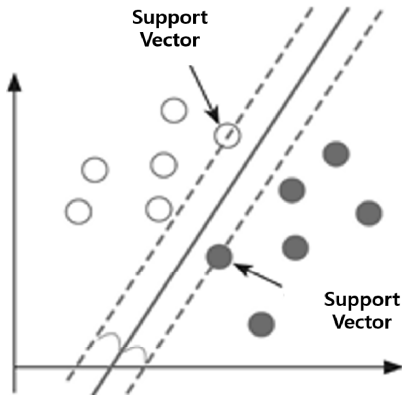


Fig. 3. SVM cognition using hyperplane

SVM 알고리즘에서 사용하는 초평면은 두 가지이다. 위의 그림처럼 사물을 선형적으로 분리할 수 있는 선형 초평면을 사용하며, 식 (1)은 선형 초평면을 이용한 선형 SVM 연산식을 나타낸다.

$$D(z) = \left(\sum_{i=0}^{m-1} a_i y_i s_i z \right) + b$$

$$= W^t \cdot z + b : W^t = \sum_{i=0}^{m-1} a_i y_i s_i \quad (1)$$

선형 SVM과 선형적으로 분리할 수 없는 경우에 사용되는 비선형 초평면이 사용되며 식 (2)는 비선형 초평면을 이용한 비선형 SVM 연산식을 나타낸다. 식 (3), (4)는 각각 비선형 SVM에서 주로 사용되는 다항식과 RBF(Radial Basis Function) 커널(kernel) 함수를 나타낸다.

$$D(z) = \left(\sum_{i=0}^{m-1} a_i y_i s_i z \right) + b$$

$$= \left(\sum_{i=0}^{m-1} a_i y_i K(s_i z) \right) + b \quad (2)$$

$$K(s_i z) = ((s_i \cdot z) + 1)^d : d > 0 \quad (3)$$

$$K(s_i z) = \exp(- \| s_i - z \|^2 / (2\sigma^2)) \quad (4)$$

본 연구에서는 최소한의 알고리즘을 활용한 편의성을 갖추어 실용성을 확보하기 위해 EEG 센서에서 측정되는 대부분의 요소를 기반으로 EEG 신호의 학습에 따른 모형을 설정하고 이를 검증한다. 선형 초평면의 경우 보통 범주 값을 유사한 데이터들로 분류하기 위해 초평면이라는 경계를 사용하는데 이는 n-1차원의 subspace를 의미하는 것이며, Fig. 4에서 보는 것과 같이 3차원의 경우 hyperplane은 2차원의 면이 되고, 2차원의 경우는 hyperplane은 1차원의 선이 된다.

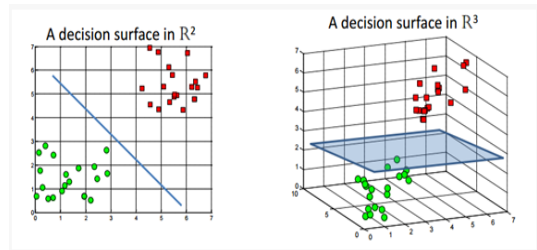


Fig. 4. Hyperplane in multi dimension

선형분리가 불가능한 데이터나 비선형 데이터의 경우에 차원을 높여서 분리하는 방법으로 Kernel Tricks 기법이 있다. 일반적으로 데이터를 분류하기 위해서는 같은 공간 내에서 패턴을 분류하기가 쉽지 않은 경우, Fig. 5, Fig. 6에서와 같이 kernel 함수를 이용하여 이들을 변형시켜 hyperplane을 이용한 분류를 쉽게 하게 된다.

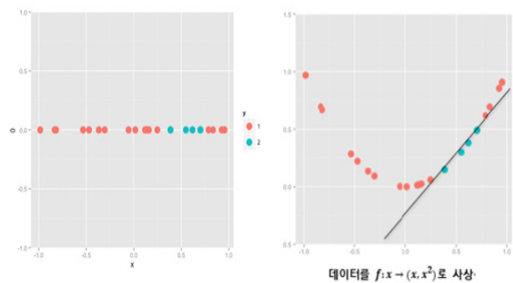


Fig. 5. Kernel Trick (1 → 2D)

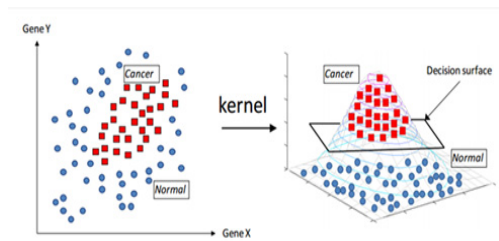


Fig. 6. Kernel Trick (2 → 3D)

3.2 EEG 데이터 수집

연구에 참석하는 대상은 동일한 연령대를 가진 성인 3 명이며, 이들의 특성으로는 EEG 신호의 대상인 3명의 특성을 살펴보면 다음과 같다. ① 50대 여. 과거 중증의 정신적 질환으로 치료를 받았고 현재에도 가벼운 증상으 로 인한 치료를 지속하고 있음. ② 50대 여. 모태 신앙으 로의 종교를 가지고 있으며 현재까지 정기적으로 독실한 신앙생활을 지속함. ③ 50대 남. 일반적인 경력을 가진. 별다른 정신적인 치료 경력은 없음.

이들의 EEG 신호를 분류하여 이들이 가진 특징을 살 펴보고자 한다. EEG 신호를 수집하기 위해 Neurosky社의 MindWave Headset을 사용하였으며, Android 기 반의 Bluetooth를 이용한 무선 데이터 전송 시스템을 이용하여 CSV 파일 형태의 데이터 작성이 가능하다.

EEG 디바이스를 활용하여 측정 가능한 데이터는 timestamp, eegRawValue, eegRawValueVolts, poorSignal, attention, meditation, blinkStrength, delta, theta, alphaLow, alphaHigh, betaLow, betaHigh, gammaLow, gammaMid, tagEvent, location 이다. 이중 일반적인 뇌파의 측정 요소인 delta, theta, alpha-, beta-, gamma- 와 timestamp, tagEvent, location 의 추가적인 정보외에 MindWave Headset에서 추가로 측정되는 주요 정보는 Attention, Meditation, Eye Blink Detection이다.

Attention 측정은 정신적인 “집중(focus)” 혹은 “주의(attention)”의 강도를 나타내며, 측정값의 범위는 0 ~ 100 이다. 이러한 측정값은 사용자가 하나의 생각이나 하거나 외부대상에 집중할 때 증가한다. Meditation 측정은 정신적인 “평온(calmness)” 혹은 “휴식(relaxation)” 을 나타내며, 불편하거나 스트레스를 받는 경우, 이를 감소하거나 완화시키는 강도를 나타내며, 측정값의 범위는 0 ~ 100 이다.

Blink Detection 측정은 사용자의 눈 깜빡임을 나타 낸다. 측정값이 클수록 더욱더 “강하거나” 혹은 “약한” 것

을 나타낸다.

4. 실험 결과 및 분석

EEG 센서에서 측정된 데이터 중, 추가된 영역에 해당 하는 내용은 Fig. 7, Fig. 8과 같다.

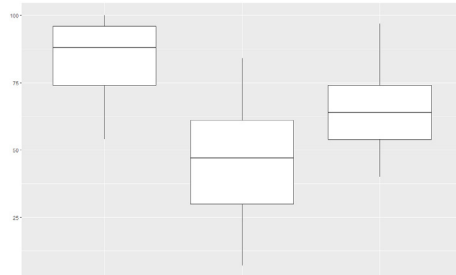


Fig. 7. Meditation by EEG Sensor

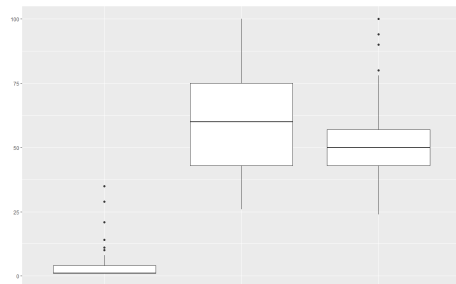


Fig. 8. Attention by EEG Sensor

Fig. 7과 Fig. 8에서 추가된 영역의 요소는 기본적인 기술 통계 값으로도 대상의 분류가 가능한 것으로 확인 된다. 이는 EEG 신호가 대상이 가진 특성을 어느 정도 반영된 것으로 확인이 가능하다. 주요적인 패턴으로 살펴 보면, 본 논문에서 EEG 센서로 확인해 본 요소중에서 Fig. 7에서 확인되는 meditation 의 경우 2번의 대상이 종교생활 등의 영향으로 제일 안정적이며 1번의 치료 경 력 등이 기본적인 생활에 대한 안정감과 스트레스에 대 한 약점으로 반영된 것으로 보여진다. Fig. 8에서 보여지 는 attention 의 경우 1번의 영향이 눈에 띄게 특이한 패턴을 보여준다.

Fig. 9는 EEG 센서에 의한 측정 데이터를 기본으로 SVM 분류와 관련한 연구에서 3개의 분류 요소에 따른 그래프를 나타낸다.

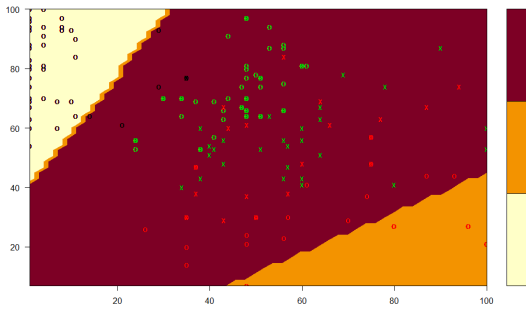


Fig. 9. SVM classification plot

Fig. 9에서 보여지는 패턴은 3개의 대상이 분류되는 신호 분류를 위해 Fig. 7의 최대 마진을 갖는 최적 분류 초평면을 찾아 최적의 대역적인 식별이 가능하며, SVM 분류 알고리즘에 의한 EEG 신호 분류가 진행되었다. 전체 데이터를 훈련셋 과 모델 검증셋을 80 : 20 의 비율로 무작위로 나누고, 이를 기준으로 훈련시킨 모델을 기반으로 실제 데이터와의 예측을 수행해 본 결과는 Table 1과 같다.

이와 관련해서 전체적인 모형의 정확도는 93.2 %로 파악되며, 훈련셋을 이용한 모델의 정확도가 상당한 수준으로 검증된다. 이는 SVM의 가장 간단한 Linear 모형의 커널을 이용한 알고리즘으로 EEG 신호를 활용하여 대상에 대한 분류가 상당한 정도의 활용가능성을 보여준다. 실제로 Fig. 10에서 보여지는 EEG 신호의 3가지 요소 (Attention, Meditation, BlinkStrength)만으로 확인해 본 3명의 대상에 대한 데이터를 나타낸다. 그림에서 파악되는 실제 분류가 시각적인 분류로도 대상을 구분가능한 형태로 파악된다.

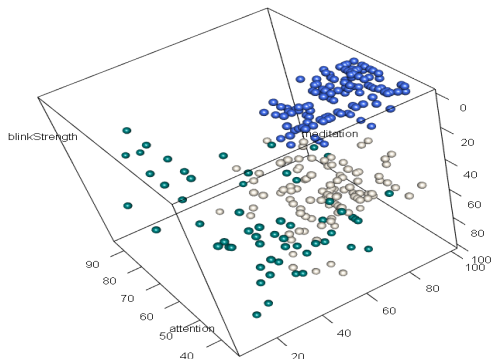


Fig. 10. Plot 3 Dimension as 3 factor (Attention, Meditation, BlinkStrength)

Table 1. Predict between real & model (93.2%)

	User 1	User 2	User 3
User 1	6367	0	0
User 2	0	5098	417
User 3	0	853	6062

5. 결론 및 향후 연구

본 논문에서는 EEG 신호에서 파악되는 대상별 분류는 약 93.2% 정도의 정확성으로 예측가능하다는 점을 보여주고 있다. 특히 SVM 알고리즘의 가장 간단한 선형 (Linear) 분류만으로 수행이 가능하다는 측면은 응용적인 관점에서 개발자의 용이함과 사용자의 편리함에 대한 정보기술적인 특이나 공학적인 측면의 어플리케이션의 활용을 제시해 주고 있다. 논문에서의 실험 결과 정신적인 활동이나 사회생활의 경력이 상당히 반영되는 요소로 파악되며, 이는 추가적인 연구의 대상으로 연령대별, 직업의 형태나 경력별, 혹은 정신적인 치료나 활동 영역별 등의 다양한 요소에 따른 향후 연구가 필요하다고 생각된다.

REFERENCES

- [1] H. Jeon & S. H. Lee. (2016). Electroencephalography of learning and memory, *Korean J Biol Psychiatry*, 23(3), 102.
- [2] O. Jensen, J. Kaiser & J. P. Lachaux. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in neurosciences*, 30(7), 317-324.
DOI : 10.1016/j.tins.2007.05.001
- [3] R. Llinás, F. J. Urbano, E. Leznik, R. R. Ramírez & H. J. van Marle. (2005). Rhythmic and dysrhythmic thalamocortical dynamics: GABA systems and the edge effect. *Trends in neurosciences*, 28(6), 325-333.
DOI : 10.1016/j.tins.2005.04.006
- [4] U. Ribary. (2005). Dynamics of thalamo-cortical network oscillations and human perception. *Progress in brain research*, 150, 127-142.
DOI : 10.1016/S0079-6123(05)50010-4
- [5] A. Schnitzler & J. Gross. (2005). Normal and pathological oscillatory communication in the brain. *Nature reviews neuroscience*, 6(4), 285-296.
DOI : 10.1038/nrn1650

[6] P. J. Uhlhaas & W. Singer. (2006). Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron*, 52(1), 155-168.
DOI : 10.1016/j.neuron.2006.09.020

[7] X. J. Wang. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, 90(3), 1195-1268.
DOI : 10.1152/physrev.00035.2008

[8] K. Lee, C. H. Lee, J. Bae & J. Lee. (2015). EEG Signal Classification Algorithm based on DWT and SVM for Driving Robot Control. *Journal of the Institute of Electronics and Information Engineers*, 52(8), 117-125.
DOI : 10.5573/IEIE.2015.52.8.117

[9] Y. K. Kim, C. J. Chae & H. J. Cho. (2018). User Authentication Method using EEG Signal in FIDO System. *Journal of the Korea Convergence Society*, 9(1), 465-471.
DOI : 10.15207/JKCS.2018.9.1.465

[10] H. G. Ko, J. M. Cho & D. Choi. (2017). An Incremental Elimination Method of EEG Samples Collected by Single-Channel EEG Measurement Device for Practical Brainwave-Based User Authentication. *Journal of the Korea Institute of Information Security and Cryptology*, 27(2), 383-395.
DOI : 10.13089/JKIISC.2017.27.2.383

[11] W. Khalifa, A. Salem, M. Roushdy & K. Revett. (2012, May). A survey of EEG based user authentication on Informatics and Systems (INFOS) (pp. BIO-55). IEEE.

[12] D. Starling. (2003). *Temporal Analysis of EEG patterns in a biofeedback based Brain Computer Interface* (pp. 1-36). Tech Report No. CYB/2003/UG/DJS/1.

[13] G. Costantini et al. (2009, July). SVM classification of eeg signals for brain computer interface. In *Proceedings of the 2009 conference on Neural Nets WIRN09: Proceedings of the 19th Italian Workshop on Neural Nets, Vietri sul Mare, Salerno, Italy, May 28-30 2009* (pp. 229-233). IOS Press.

[14] C. H. Lee, J. W. Kwon, G. D. Kim, J. E. Hong, D. S. Shin & D. H. Lee. (2009). A study on EEG based concentration transmission and brain computer interface application. *Journal of the Institute of Electronics Engineers of Korea SC*, 46(2), 41-46.

[15] J. Chuang, H. Nguyen, C. Wang & B. Johnson. (2013, April). I think, therefore i am: Usability and security of authentication using brainwaves. In *International Conference on Financial Cryptography and Data Security* (pp. 1-16). Springer, Berlin, Heidelberg.

이 상 원(Sang-Won Rhee)

[정회원]



- 1996년 6월 ~ 2005년 2월 : 삼성SDS
- 2013년 3월 ~ 2016년 2월 : 대구대 과학교육과(교육학 석사)
- 2016년 3월 ~ 2018년 2월 : 대구대 과학교육과(박사수료)
- 2017년 6월 ~ 현재 : 소프트개발사 대표

- 관심분야 : 정보보호, 인공지능, 딥러닝, 빅데이터
- E-Mail : cupsarang@hanmail.net

조 한 진(Han-Jin Cho)

[정회원]



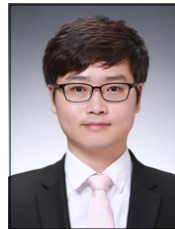
- 1999년 2월 : 한남대학교 컴퓨터공학과(공학석사)
- 2002년 8월 : 한남대학교 컴퓨터공학과(공학박사)
- 2002년 8월 ~ 현재 : 극동대학교 에너지IT공학과 교수
- 관심분야 : 정보보호, 스마트폰 보안,

모바일 콘텐츠

- E-Mail : hanjincho@hotmail.com

채 철 주(Cheol-Joo Chae)

[정회원]



- 2009년 8월 : 한남대학교 컴퓨터공학과(공학박사)
- 2009년 9월 ~ 2013년 4월 : 한국전 자동차연구원 선임연구원
- 2013년 4월 ~ 2016년 8월 : 한국과학기술정보연구원 선임연구원
- 2016년 9월 ~ 현재 : 한국농수산대학

교양공통과 교수

- 관심분야 : 정보보호, 바이오 보안, 네트워크 보안
- E-Mail : chae.cheeljoo@gmail.com