

생체 데이터를 이용한 프로그래머의 프로그램 난이도 예측

이설화[○], 임희석

고려대학교

whiteldark@korea.ac.kr, limhseok@korea.ac.kr

Mining Biometric Data to Predict Task Difficulty

Seolhwa Lee[○], Heuiseok Lim

Korea University

요 약

프로그래머들이 코딩을 할 때 발생하는 빈번한 실수는 많은 시간적 비용을 낭비할 수 있고 작은 실수가 전체 코드에 치명적인 에러를 유발하기도 한다. 이러한 문제점은 프로그래머들이 코드를 작성할 때 전체적인 알고리즘을 얼마나 잘 이해하는지와 이전 코드에 대한 이해력과 연관이 있다. 만약 코드에 대한 이해가 어렵다면 정교하고 간결한 코드를 작성하는데 무리가 있을 것이다. 기존 코드에 대한 난이도를 평가하는 방법은 자가평가 등을 통해 이루어져 왔다. 사람 내부 변화를 직접 측정하면 더 객관적인 평가가 가능할 것이다. 본 논문은 이런 문제들을 해결하고자 동공 추적에 가능한 아이트래커와 뇌파 측정이 가능한 EEG장비를 이용하여 습득한 생체 데이터를 통해 프로그래머들의 프로그램 난이도 예측 모델을 개발하였다.

주제어: 아이트래커, EEG, 기계학습, 프로그램 이해력, 프로그램 난이도

1. 서론

프로그래머들이 코딩을 할 때 발생하는 빈번한 실수는 많은 시간적 비용을 낭비할 수 있고 작은 실수가 전체 코드에 치명적인 에러를 유발하기도 한다. 이러한 문제점은 프로그래머들이 코드를 작성할 때 전체적인 알고리즘을 얼마나 잘 이해하는지와 이전 코드에 대한 이해력과 연관이 있다. 만약 코드에 대한 이해가 어렵다면 정교하고 간결한 코드를 작성하는데 무리가 있을 것이다.

기존의 코드에 대한 난이도를 평가하는 방법에는 코드가 얼마나 어려웠는지에 대한 자가평가를 통해 이루어져 왔다. 사람의 내부 변화를 직접 측정한다면 더 객관적인 평가가 가능할 것이다.

본 연구는 위의 문제를 해결하고자 직접적으로 프로그래머의 정신적 변화를 측정할 수 있는 생체 인식 센서를 사용하여 프로그래머의 프로그램 난이도를 측정하였다. 생체 인식 센서를 이용한 측정은 소프트웨어 프로그래머들의 프로그램 코드 이해력을 판단할 수 있는 새로운 가능성을 도출해낼 수 있다. 생체 인식 센서란 사용자 개개의 심리적 상태를 반영하여 뇌파, 눈 움직임 등의 신체적 변화 반응을 측정할 수 있는 장비를 말한다. 생체 인식 센서는 심리학 분야에서 많이 사용되어 왔는데, 사용자들의 심리상태 변화 [1,2]를 측정하는 도구로 많이 사용되어 왔다. 그러나 프로그래머들의 프로그램 코드를 이해하는 정신적 인지부하를 측정하는 분야에서는 적은 연구가 이루어져왔다.

기존 연구들 중 Parnin[3]의 연구에서는 소리없이 발성을 할 때 어려운 과제를 하고 있는지에 대한 근전도(electromyography) 검사를 수행하였다. Parnin은 소리없이 발성을 할 때 어려운 프로그램 코드를 수정한다는

연관성을 찾아냈다. 그러나 Parnin의 연구에서는 근전도 검사를 수행하여 심리적인 요인을 고려하지 않았다.

따라서, 본 연구는 뇌파를 측정할 수 있는 EEG(electroencephalographic)와 눈 움직임을 추적할 수 있는 아이트래커를 사용하여 생체 데이터를 획득하였다. 본 연구는 획득한 프로그래머들의 생체 데이터를 이용하여 기계학습 기법을 통해 프로그램 난이도를 예측하는 모델을 개발하였다.

본 논문의 기여는 다음과 같다.

- 코드를 작성할 때 프로그램 난이도를 자동적으로 예측함으로써 프로그래머들에게 휴식시간을 주어 불필요한 코드 에러를 방지할 수 있다.
- 프로그램 문제를 푸는 경우에는 프로그래머의 이해수준에 맞게 자동적으로 난이도 조정이 가능하다.
- 직접적인 생체 데이터를 이용하므로 기존 방법에 의존할 필요가 없다.

본 논문의 구성은 다음과 같다. 2장에서 정신적 인지부하를 측정 방법에 대한 관련 연구에 대해 다루고, 3장은 연구의 방법론, 4장은 본 연구에서 사용한 기계학습 방법, 5장은 분석결과, 6장은 결론으로 구성된다.

2. 관련 연구

눈 움직임과 뇌에 기반한 정신적 인지부하를 측정하는 센서들은 주로 심리 상태를 측정하기위하여 사용된다 [1,2].

눈의 움직임에 기반한 측정방법에는, 응시 고정, 도약 안구 운동의 횟수, 동공의 크기 등을 주로 이용한다. 특히 동공의 크기와 정신적 인지부하와의 연관성에 관한 연구가 많이 진행되어 왔다[4]. Goldberg의 연구[5]에서

는 도약 안구 운동의 횟수가 증가하면 좋지 않은 인터페이스를 가지고 있는 것으로 결론을 내렸다. 최근에는 눈 움직임에 기반한 측정방법으로 사람의 인지상태를 예측하는 연구들이 진행되고 있다[6].

뇌에 기반한 측정방법에는, 각 뇌의 위치에서 나오는 EEG(electroencephalographic) 데이터를 이용하여 다양한 정신 상태를 연관지을 수 있다[7]. 뇌파는 정신적인 활동을 할 때 마다 특정 부위의 뇌의 내부 또는 두피의 표면에서 전기적인 신호가 발생하는 것을 말한다. EEG는 각 주파수별로 크게 알파파(9-13Hz), 쉼타파(4-8Hz), 베타파(13-30Hz), 감마파(30-50Hz), 4가지로 나뉜다. Smith[8]의 연구에서는 알파파의 감소가 작업 기억과 연관이 있다고 결론내었다. 다른 관련 연구에서는 쉼타파와 델타파가 과제의 난이도를 바꿀때마다 민감하게 변화하는 것을 찾아내었다[9].

3. 방법론

3.1 실험대상

본 실험은 연구자 기관의 총 18명의 컴퓨터 과학 전공 학생 (여자 5명, 남자 13명, 평균나이 24세)이 참여하였고, 모든 참가자들은 연구기관의 기관윤리위원회의 승인을 받은 실험(KU-IRB-16-20-A-1) 절차에 따라 실험 설명을 듣고 연구 참여 동의서에 서명을 하였다.

3.2 실험절차

본 실험은 36개의 자바 코드로 구성되며 [10]의 실험을 통해 검증된 코드를 사용하였다. 사용된 코드는 팩토리얼 계산, 교체 문제, 최댓값 찾기 등의 간단한 코드로 구성된다.

실험 대상자들은 실험 전 실험에 관한 지시사항을 안내받고 36번의 코드를 반복해서 풀게 된다[11]. 문제는 직접 코딩을 하는 것이 아니고 4지선다의 출력값을 맞추는 문제이다. 과제를 완료한 직후 온라인 설문조사를 통하여 문제의 난이도를 기입한다. 문제의 난이도는 5점까지의 리커트 척도를 기입한다.

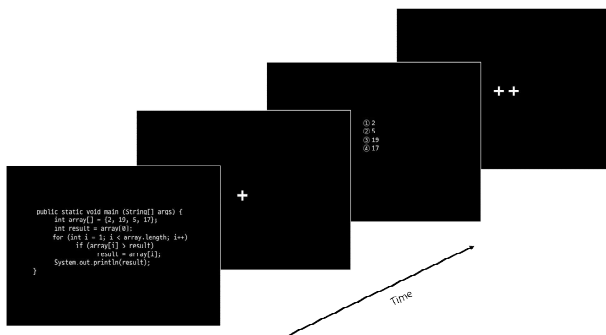


그림 1 실험 자극



그림 2 실제 실험 모습

3.3 실험 장비 및 분석 소프트웨어

EEG. 실험 장비는 Brainproduct's사의 actiCAP과 amplifier로 V-Amp 5V, 450mA를 사용하였다. 장비는 16개의 채널로 구성되며 국제 전극배치법 10-20을 따른다. 전극과 두피와의 저항값은 10KΩ 이하로 유지하였다. 실험 소프트웨어는 E-prime professional 2.0을 사용하였으며 뇌파 레코딩 소프트웨어는 Brainproduct's의 Brain Vision Recorder를 사용하였고, sampling rate는 500Hz로 기록하였다. 분석 소프트웨어는 동일회사의 Brain Vision Analyzer 2.1을 사용하여 데이터를 추출하였다.

아이트래커. 실험 장비는 SMI사의 RED-mx 아이트래커를 사용하였다. sampling rate는 60Hz로 안구 움직임을 추적하였다. 분석소프트웨어로는 동일회사의 BeGaze 3.5를 통해 데이터를 추출하였다.

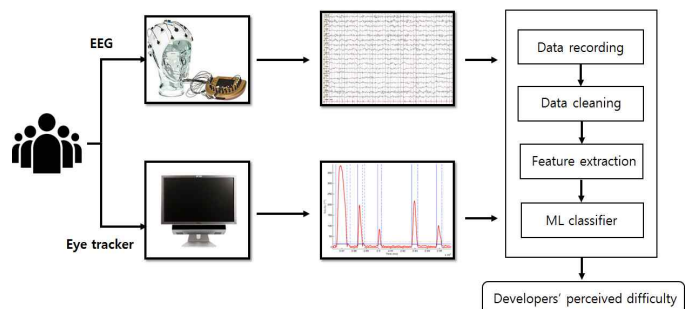


그림 3 생체 데이터를 이용한 분석과정 개요

3.4 데이터 전처리

EEG. 근육신호 등의 노이즈 데이터를 제거하기 위하여 ICA(Independence Component Analysis) 알고리즘을 사용하였고 주파수 필터링은 FFT 필터링을 사용하여 4-50Hz 구간의 채널별 알파, 쉼타, 베타, 감마 밴드의 절대파워 데이터를 얻었다. 모든 사람마다 독특한 파워스펙트럼 분포를 갖기 때문에 실험대상자 각각의 뇌파를 비교하기 위하여 각 밴드별로 상대비율을 구하여 전처리를 진행하였다.

$$\text{알파파 상대파워} = \frac{9-13\text{Hz 구간 절대파워}}{4-50\text{Hz 구간 절대파워}}$$

아이트래커. SMI사의 BeGaze소프트웨어에서 제공하는 fixation-dispersion 알고리즘을 사용하여 응시 고정, 도약 안구 운동 등을 추출하였다. 눈을 감고 뜨기 전에는 동공의 크기가 커지기 때문에 눈을 깜박인 후에 처음 있는 응시 고정 포인트는 제거하였다. 또한 동공 크기를 이용하여 집중도를 판단할 수 있는 지표로써 사용이 가능하다[12]. 응시 고정 횟수, 도약 안구 운동 횟수, 응시 시간 등의 요소를 추출하였다.

4. 기계학습

본 논문은 프로그래머의 프로그램 난이도 예측을 위하여 SVM(Support Vector Machine)을 사용하였다. SVM은 지도학습분류 모델로 최대 마진 초평면을 갖는 알고리즘으로 선형, 비선형 알고리즘보다 우수한 성능을 보인다[13]. 특히 SVM은 EEG 데이터를 분류하는데 높은 정확성을 보인다[14,15]. 또한 아이트래커를 이용한 예측 모델에서도 많이 사용되고 있다[16].

총 18명의 실험 대상자가 참여하여 36개의 문제를 수행하므로 데이터는 총 648개의 데이터로 구성된다. 또한 분류 클래스는 앞서 설문한 프로그램 난이도의 리커트척도에서 쉬움/어려움으로 분류하여 정답셋을 설정하였다.

적은 데이터의 양을 사용하였기 때문에 분류 모델의 성능에 대한 통계적 신뢰도를 높이기 위하여 10-겹 교차 검증(10-fold cross validation)을 수행하였다. 10-겹 교차검증은 전체 데이터를 10등분으로 쪼개어 그 중 9개(전체 데이터의 90%)는 학습데이터로 사용하고 나머지 1개(전체 데이터의 10%)의 데이터는 테스트 데이터로 순서를 바꿔가면서 검증하는 방법이다.

5. 분석 결과

본 연구는 데이터 분석 시 SVM 알고리즘 중 선형 커널을 사용하였으며 파라미터 C의 값은 1로 설정하였다. 분석은 아이트래커(EYE), EEG, 아이트래커(EYE)+EEG, 총 3가지 요소로 나누어서 분석하였다.

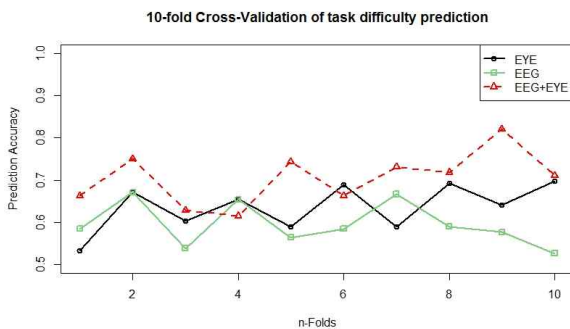


그림 4 10-겹 교차 검증을 이용한 예측 정확성 결과

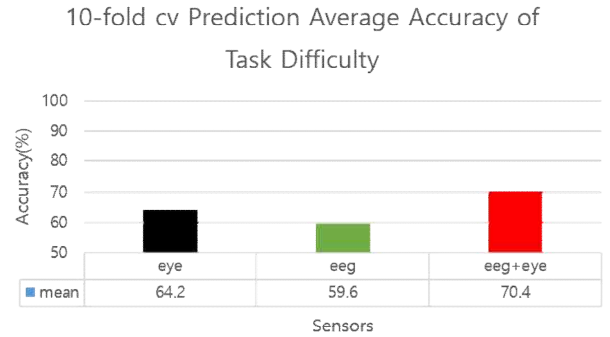


그림 5 각 센서별 평균 예측 정확성 결과

그림 4는 각 장비에서 추출한 요소를 가지고 10-겹 교차 검증을 통해 데이터를 학습시켰을 때의 모델에 대한 프로그램 난이도 예측에 대한 정확성 그래프이다. 아이트래커 데이터만 사용하여 데이터를 학습시켰을 때는 10-겹 교차 검증에서 표준편차 = 0.048 이었고, EEG 데이터만 사용하여 데이터를 학습시켰을 때는 표준편차 = 0.051, 두 장비를 모두 사용하였을 때는 표준편차 = 0.045로 결과가 나왔다.

그림 5는 각 장비별로 10-겹 교차 검증을 수행한 후의 평균 정확성을 계산한 그래프이다. 아이트래커와 EEG를 함께 사용하였을 때가 70.4%로 예측 정확성이 제일 높았으며, 아이트래커만 사용하였을 때가 64.2%, EEG만 사용하였을 때가 59.6%로 예측 정확성이 나왔다. 정확성은 두 장비를 모두 사용하였을 때가 가장 높게 나왔으며 그 다음으로 아이트래커만을 사용하였을 때가 정확성이 높게 나왔다.

비선형 SVM을 사용하여 다시 모델을 학습시켰을 경우에는 아이트래커 (정확성=73%, 표준편차=0.039), 아이트래커 + EEG (정확성=70.5%, 표준편차=0.045), EEG (정확성=65.8%, 표준편차=0.072) 으로 전체적으로 모델의 성능은 향상되었지만 선형 SVM으로 학습시켰을 때와 다르게 아이트래커만을 이용한 모델이 예측 정확도가 더 높게 나왔다.

6. 결론

본 논문은 생체 데이터를 이용하여 프로그래머의 프로그램 난이도를 예측하는 모델을 개발하였다. 모델 평가 결과는 전반적으로 아이트래커만을 사용하거나 아이트래커와 EEG장비를 모두 사용하는 경우가 프로그램의 난이도를 예측하는데 더 정확하였다. 본 연구는 사람으로 대상으로 실험하였기 때문에 데이터를 수집하고 적은 데이터로 모델을 개발하여 추후 특징 선택 (feature selection)을 사용하여 이런 한계점을 해결할 필요성이 있다.

Acknowledgement

"이 논문은 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. R1610941)."

"본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음. [2016(B0101-16-0340), 개인과 집단지성의 디지털 콘텐츠화를 통한 유통 및 확산 서비스 기술 개발]"

참고문헌

- [1] J. A. Veltman and A. W. K. Gaillard, "Physiological workload reactions to increasing levels of task difficulty," *Ergonomics*, vol. 41, no. 5, pp. 656-669, 1998.
- [2] Wierwille, W. W., & Eggemeier, F. T, "Recommendations for mental workload measurement in a test and evaluation environment," *Human Factors*, vol. 35, 263-281, 1993.
- [3] C. Parnin, "Subvocalization-toward hearing the inner thoughts of developers," In *Proc. of the 19th Int'l Conf. on Program Comprehension (ICPC)*, pp 197-200, 2011.
- [4] Pedrotti, M., Lei, S., Dzaack, J., & Rötting, M, "A data-driven algorithm for offline pupil signal preprocessing and eyeblink detection in low-speed eye-tracking protocols," *Behavior Research Methods*, vol. 43, pp 372-383, 2011.
- [5] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: methods and constructs," *Int'l. Journal of Industrial Ergonomics*, vol. 24(6), pp 631- 645, 1999.
- [6] Jaana Simola, Jarkko Salojärvi, and Ilpo Kojo, "Using Hidden Markov to uncover Processing States from Eye Movements in Information Search Tasks," *Cognitive Systems Research*, vol.9, pp 237-251, 2008.
- [7] Klimesch, W, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Res Rev*, vol. 29, pp169-195, 1999.
- [8] M. E. Smith and A. Gevins," Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator," In *Defense and Security*, pp 116-126. *Int'l. Society for Optics and Photonics*, 2005.
- [9] Hankins, T. C., & Wilson, G. F, "A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight," *Aviation, Space, and Environmental Medicine*, vol. 69, pp 360-367, 1998.
- [10] Siegmund, J., Kästner, C., Apel, S., Parnin, C., Bethmann, A., Leich, T., ... & Brechmann, A. "Understanding understanding source code with functional magnetic resonance imaging." ,In *Proceedings of the 36th International Conference on Software Engineering* (pp. 378-389). ACM. 2014.
- [11] Seolhwa, Lee, Kichun, Nam, Jae-bum, Jung, Song-hyun Kim, Heuseko, Lim, "An EEG Experimental Design for Understanding Differences in Understanding Programming Language between Novice and Expert Programmers" ,*The 3rd International Conference for Small & Medium Business*, 2016.
- [12] Lee, E.C., Park, K.R., Whang, M. and Min, K. Measuring the degree of eyestrain caused by watching LCD and PDP devices. *International Journal of Industrial Ergonomics*, 39(5), pp.798-806, 2009.
- [13] Noor Aishah Atiqah Zulkifli, Sawal Hamid Md Ali, Siti Anom Ahmad, Md Shabiul Islam," Review on Support Vector Machine (SVM) Classifier for Human Emotion Pattern Recognition from EEG Signals," *Asian Journal of Information Technology*, Vol. 14, pp 135-146. 2015.
- [14] Rani, P., Liu, C., Sarkar, N., Vanman, E. "An empirical study of machine learning techniques for affect recognition in human robot interaction." *Pattern Analysis and Applications*, Vol. 9, pp 58-69, 2006.
- [15] Chen, G., Hou, R. "A New Machine Double-Layer Learning Method and Its Application in non-Linear Time Series Forecasting." In: *Proceedings of the 2010 International Conference on Mechatronics and Automation (ICMA)*, 2007.
- [16] Rello, Luz, and Miguel Ballesteros. "Detecting readers with dyslexia using machine learning with eye tracking measures." In *Proceedings of the 12th Web for All Conference*, p. 16. ACM, 2015.