JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지 Vol. 23, No. 8: 917~922, Aug. 2019

딥 러닝 기반 대학 이수학점 및 활동에 의한 교원임용 후보자 경쟁 시험 합격여부 예측

김근호[1] · 김의정[2*]

Deep learning based teacher candidate acceptance prediction using college credits and activities

Geun-Ho, Kim[1] · Eui-Jeong, Kim[2*]

[1]Graduate student, Department of Computer Education, Kongju National University, Gongju, 32588 Korea
[2*]Professor, Department of Computer Education, Kongju National University, Gongju, 32588 Korea

요 약

최근 교사 직업에 대한 선호도가 높아짐에 따라 사범대학 및 교육대학의 선호도가 높아지고 있다. 모든 학생들이 교사에 진출 할 수 있는 것이 아니라 졸업 후 교원임용 후보자 경쟁시험이라는 시험을 통과하여야한다. 그러나 인구수 감소에 따른 임용T.O는 매년 줄어들고 있고 경쟁률은 가파르게 오르고 있는 사정이다. 때문에 입학하면서 임용시험에 매진하고자 대학교가 학문을 연구하고 배우는 곳이 아닌 시험을 위한 거대한 학원이 되어가는 중이다.

이에 본 연구에서는 학점뿐이 아니라 학생들의 전반적인 학교생활 및 스터디그룹 활용 여부 등을 종합하여 딥 러닝을 이용한 교원임용 후보자 경쟁시험 합격여부의 연관성을 파악하였다. 합격에 대하여 학과활동 등이 크게 영향을 미치지는 않았고 합격률 예측에 대한 정확도는 대체적으로 70%의 정확성을 보였다.

ABSTRACT

The recent increase in preference for teacher jobs has led to a rise in preference for education colleges. Not all students can enter teachers, but they must pass the test called the competitive examination for teacher appointment candidates after graduation. However, due to the declining population, the and employment T.O.s are decreasing every year and the competition rate is rising steeply. Therefore, in order to concentrate on the recruitment exam upon entering the university, the university is becoming a huge academy for the exam, not a place to study and learn.

We found a connection between students' overall school life and their use of study groups as well as their grades and whether they passed the competition test for teachers using deep running. The academic activities did not significantly affect the acceptance process, and the accuracy of the prediction of the acceptance rate was generally 70% accurate.

키워드 : 임용시험, 딥러닝, 예측시스템, 인공지능

Keywords : Appointment test, Deep Learning, Prediction System, Artificial intelligence

Received 29 May 2019, **Revised** 3 June 2019, **Accepted** 20 June 2019

*** Corresponding Author** Eui-Jeong, Kim(E-mail:ejkim@kongju.ac.kr, Tel:+82-41-850-8823)
Professor, Department of Computer Education, Kongju National University, Gongju, 32588 Korea

Open Access http://doi.org/10.6109/jkiice.2019.23.8.917 print ISSN: 2234-4772 online ISSN: 2288-4165

ⓒThis is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(http://creativecommons.org/li-censes/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright Ⓒ The Korea Institute of Information and Communication Engineering.

Ⅰ. 서 론

최근 고용 불안과 워라벨을 중시하는 세태와 맞물려 교사 직업에 대한 선호도가 높아지고 있다. 그러나 이러한 선호도와는 다르게 실제 교원 임용 후보자 경쟁시험에 대한 염려로 인하여 원하는 사범대학에 진학하였으면서도 대학생활이나 학과 생활에 대한 고려 없이 입학과 동시에 임용 후보자 경쟁시험만을 위한 과정이 되어가고 있는 실정이다. 이에 본 연구에서는 임용시험에 합격 여부를 예측할 수 있는 모델을 설계하고자 한다. 단순히 시험공부만 열심히 한다고 합격 여부가 결정되는 것이 아님을 밝히기 위하여 학점을 전공학점 및 교육학 학점 등으로 세분화 시키고 학과 활동에 대한 등급을 부여하여 학점 및 학과 생활 등이 임용고시 합격에 미치는 영향을 파악 하고자 한다. [1]

Ⅱ. 딥 러닝

2.1. 딥러닝의 개념

딥러닝은 인공지능의 접근방식의 하나로 컴퓨터에 명시적으로 프로그래밍을 하는 것이 아니라 컴퓨터가 축적된 데이터를 학습을 통해 아직 알 수 없는 특정 데이터에 대한 판단을 하는 모델을 만드는 것이다.[2]

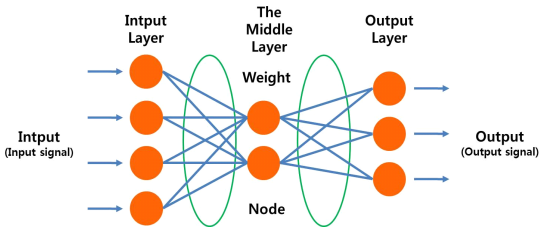


Fig. 1 Deep Learning

딥이라는 말은 신경망의 층이 깊고, 각 층마다 고려되는 변수가 많다는 의미이다 층의 깊이에 따라 구분하자면 2~3개의 층으로 되어 있는 신경망을 천층망이라고 한다. 그림 1에서 보이듯 딥러닝에서 깊이를 나타내는 층의 개수는 입력층과 출력층 사이에 있는 은닉층 개수에 하나를 더하면 된다. 이러한 이유로 딥러닝을 심층신경망이라고 한다. 딥러닝의 기본개념은 기존 신경망에 층을 증가시킨 심층신경망을 학습하여 패턴인식이나 추론에 사용하는 것을 말한다.[3]

2.2. 딥러닝 프레임워크 비교

딥러닝 프레임워크는 딥러닝 모형을 쉽고 빠르게 개발할 수 있도록 도움을 주는 소프트웨어로서 대표적인 기능으로 '자동미분'과 'GPU의활용' 등을 들 수 있다. 대학에서 개발되어 사용되어 왔던 딥러닝 프레임워크로서는 버클리 대학의 '카페(Caffe)', 몬트리올 대학의 '티아노(Theano)'이다.[4] 그리고 연전에 구글이 '텐서플로(TensorFlow)'를 오픈소스로 공개함에 따라 저변을 확대해 오고 있다. 한편, 최근에는 마이크로소프트의 딥러닝 프레임워크인 Microsoft CognitiveToolkit(CNTK) 2.0이 공개되었는데, C++와 파이썬 등의 프로그래밍 언어와 호환성을 제공하게 되면서 스크립트언어로만 사용하는 제약을 벗어나서 텐서플로나 티아노 같은 딥러닝 프레임워크와 어깨를 나란히 하게 되었다. 이를 보면 기존의 딥러닝 프레임워크는 연구용으로 주로 대학에서 개발되어 저변을 확대해 오다가, 구글과 마이크로소프트가 가세하여 딥러닝 프레임워크에 앞으로도 지속적인 투자를 할 것으로 예상이 된다.[5]

아래 표1은 현재 많이 쓰이는 딥러닝 프레임 워크를 비교한 것을 표로 정리한 것이다.

Table. 1 Deep Learning Framework comparison

Development	Deep Learning Framework	Coding ease	Execute Speed
Berkeley University	Caffe		No big difference
University of Montreal	Theano	3	
Google	TensorFlow	2	
Microsoft	CognitiveToolkit(CNTK)	1	

Ⅲ. 연구방법

3.1. 연구대상

본 연구는 공주대학교 사범대학 컴퓨터교육과 소속의 학생들의 졸업사정에 사용되었던 성적표 및 학과 활동 기록을 이용하였다. 2013년 2월 졸업생부터 2019년 2월 졸업생까지의 졸업사정 결과 및 학과활동에 참여 여부에 관련된 학과 활동 마일리지를 참고 하여 더불어

임용고사 합격 여부를 조사하여 사용하였다.[1]

아래의 표 2 에서 보이듯 7년간의 학생 136명학생의 데이터를 활용하였으며 학생 별로 졸업 평점 학점과 복수전공 여부, 스터디 유무 및 학과 활동 등으로 입력 변수를 지정하였다.

Table. 2 Annual number of graduates

Year of Graduation	The number of graduates
2013/2	14
2013/8	2
2014/2	19
2014/8	5
2015/2	15
2015/8	3
2016/2	21
2016/8	3
2017/2	19
2017/8	1
2018/2	12
2018/8	3
2019/2	19
	136

3.2. 사전조사

아래의 표 3은 학생들에게 임용시험과 관련하여 학과활동이 임용시험을 준비하는데 방해가 된다고 생각하느냐에 대한 응답이다.

Table. 3 Pre-surface1

Very agree	agree	Neutral	disagree	Very disagree
5	31	13	11	4
7.8125%	48.4375%	20.3125%	17.1875%	6.25%

공주대학교 컴퓨터교육과 2학년~4학년 재학생 64명을 대상으로 조사하였고 75%이상의 학생이 학과 활동이 임용시험에 방해가 된다고 생각하고 있다.

Table. 4 Pre-surface2

major class	academy	online lecture	Study	Individual learning
3	3	4	5	5
15%	15%	20%	25%	25%

표 4는 학생들에게 임용시험에 본인이 가장 중요하다고 생각하는 것을 조사한 것이다. 이는 컴퓨터교육과 4학년 재학생 20명에게 설문 조사 한 것이다. 학생들의 대다수가 전공 수업보다는 스터디 및 개인학습이 중요하다고 생각하고 있다.

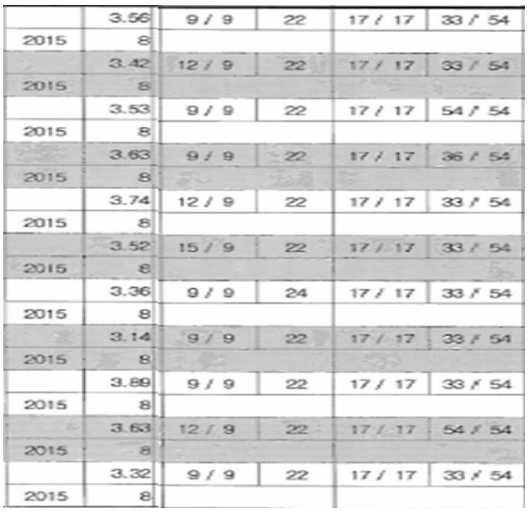
표3과 표4에서 보이듯이 학생들은 대학교 생활에서의 다른 활동이나 생활을 자신에게 도움이 된다고 느끼는 것이 아니라 단순히 임용시험에 방해가 된다고 느끼고 있으며 심지어 표 4에서 보이듯이 학과 수업보다는 학원 강의나 개별적인 스터디에서 임용고사의 도움을 많이 얻는다고 느끼고 있다.

그러나 이는 현재 재학 중인 학생들의 개인적인 생각으로서 실제 합격한 학생들의 설문조사 및 자료가 반영되지 않은 결과이다. 다음 절에서는 실제 졸업생들의 합격과 불합격 여부 및 학교생활 및 성적 등을 바탕으로 임용고사에 어떠한 사항이 도움이 되고 방해가 되는지 알아보도록 하겠다.

3.3. 전처리 과정

첫 번째 데이터를 딥러닝을 위한 데이터 전처리 과정 두 번째는 데이터를 입력받고 저장하는 데이터 처리 과정, 세 번째는 데이터를 변환하여 규칙을 찾아내고 마지막으로 합격여부를 예측 하는 과정을 거친다.[6][7]

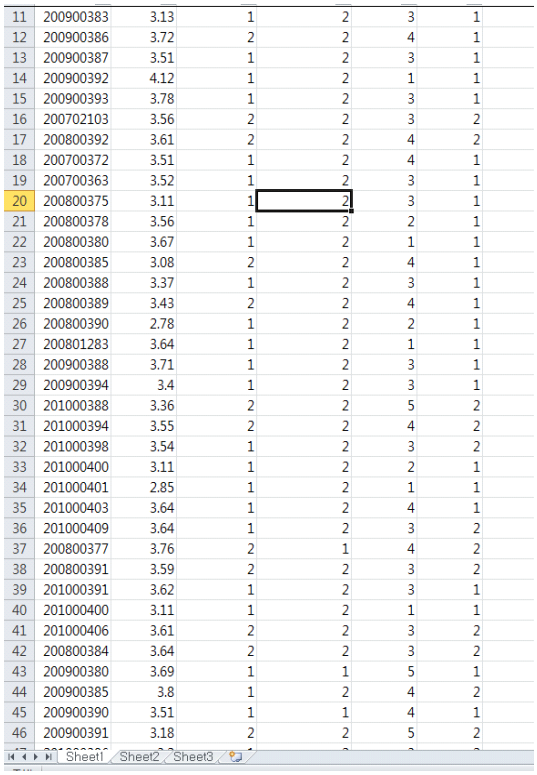
아래 그림 2는 학생들의 졸업관련 된 졸업사정 원부의 일부를 발췌한 그림이다. 이 그림을 바탕으로 데이터를 전처리를 하게 되는 것이다.



	3.56	9 / 9	22	17 / 17	33 / 54
2015	8				
	3.42	12 / 9	22	17 / 17	33 / 54
2015	8				
	3.53	9 / 9	22	17 / 17	54 / 54
2015	8				
	3.63	9 / 9	22	17 / 17	36 / 54
2015	8				
	3.74	12 / 9	22	17 / 17	33 / 54
2015	8				
	3.52	15 / 9	22	17 / 17	33 / 54
2015	8				
	3.36	9 / 9	24	17 / 17	33 / 54
2015	8				
	3.14	9 / 9	22	17 / 17	33 / 54
2015	8				
	3.89	9 / 9	22	17 / 17	33 / 54
2015	8				
	3.63	12 / 9	22	17 / 17	54 / 54
2015	8				
	3.32	9 / 9	22	17 / 17	33 / 54
2015	8				

Fig. 2 Pre-treatment data

아래 그림 3은 그림 2를 바탕으로 전처리한 학생의 데이타 일부이다. 학생들의 졸업사정시 제출하였던 이수영역별 성적표 및 학생생활 기록 등을 참고 하여 텐서플로우를 활용하기 위한 전처리 과정으로 엑셀에 데이터를 정리한 것이다.



11	200900383	3.13	1	2	3	1
12	200900386	3.72	2	2	4	1
13	200900387	3.51	1	2	3	1
14	200900392	4.12	1	2	1	1
15	200900393	3.78	1	2	3	1
16	200702103	3.56	2	2	3	2
17	200800392	3.61	2	2	4	2
18	200700372	3.51	1	2	4	1
19	200700363	3.52	1	2	3	1
20	200800375	3.11	1	2	3	1
21	200800378	3.56	1	2	2	1
22	200800380	3.67	1	2	1	1
23	200800385	3.08	2	2	4	1
24	200800388	3.37	1	2	3	1
25	200800389	3.43	2	2	4	1
26	200800390	2.78	1	2	2	1
27	200801283	3.64	1	2	1	1
28	200900388	3.71	1	2	3	1
29	200900394	3.4	1	2	3	1
30	201000388	3.36	2	2	5	2
31	201000394	3.55	2	2	4	2
32	201000398	3.54	1	2	3	2
33	201000400	3.11	1	2	2	1
34	201000401	2.85	1	2	1	1
35	201000403	3.64	1	2	4	1
36	201000409	3.64	1	2	3	2
37	200800377	3.76	2	1	4	2
38	200800391	3.59	2	2	3	2
39	201000391	3.62	1	2	3	1
40	201000400	3.11	1	2	1	1
41	201000406	3.61	2	2	3	2
42	200800384	3.64	2	2	3	2
43	200900380	3.69	1	1	5	1
44	200900385	3.8	1	2	4	2
45	200900390	3.51	1	1	4	1
46	200900391	3.18	2	2	5	2

Sheet1 Sheet2 Sheet3

Fig. 3 Pre-treatment data

그림 3을 바탕으로 다음 표 5와 같이 몇 가지 항목을 정리하여 데이터를 변환하고 전처리 하였다.

Table. 5 Preprocessed data item

Category	Variable
Student ID	Student ID
Grade	Graduation Grade
Multiple major status	1= Multiple major status Uncomplete 2= Multiple major status Complete
Study	1= Study X 2= Study ○
Major activity	Scale 1~5(best~worst)
Passing the appointment examination	1= Fail 2= Pass

표 5의 데이터를 좀 더 세분화 하여 응시 지역 및 임용고시 응시 횟수 등을 분류하고 싶었으나 해당 분류는 좀 더 나중의 연구를 위하여 남겨두기로 하고 본 연구에서는 임용고사 합격의 학과활동의 영향성을 판단하기로 해서 학과활동의 척도에 중점을 두었다.

Ⅳ. 텐서플로우를 이용한 진로예측 모델 설계

4.1. 순서도

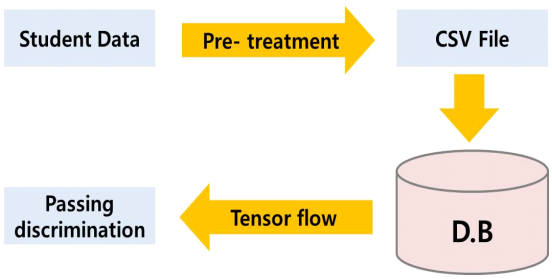


Fig. 4 Flow chart

그림 4는 플로우 차트를 나타내고 있다. 학생들의 데이터를 기반으로 전처리를 하여 딥러닝을 할 수 있는 CSV파일로 만들어 DB를 생성하여 텐서플로우를 이용하여 규칙을 생성하고 결론을 도출하는 것이다. 18년도 8월 데이터까지 테스트 데이터를 이용하고 19년도 2월 데이터를 바탕으로 정확도를 측정하였다. 아래 표 6에서 보이듯 레이어를 늘려서 정확도를 높이고자 했으나 일정수의 레이어 층이 늘어나면 오히려 정확도가 떨어지는 경향을 보인다. 이는 딥 러닝을 무작정 많이 훈련을 한다고 해서 성능이 좋아지지 않는다는 것을 시사한다. 반복적이고 다양한 시도를 해서 정확도를 높일 필요가 있을 것이다.

Table. 6 Accuracy

number of layer	Rmse	Accuracy
1	1.07	65.84%
3	0.96	71.67%
5	0.85	70.83%
7	0.94	62.50%

본 연구의 결과는 학점이 높고 공부를 열심히 한다고 해서 임용고사에 합격한다는 것도 아니고 학과생활을 열심히 하고 많이 한다고 해서 임용고사에 떨어지는 것은 아니라는 것이다.

학점이 높지만 임용고사에 떨어진 학생들도 있는 반면에 학점이 낮더라도 합격한 학생들도 있고 학과생활을 전혀 하지 않고 임용시험에만 매진한 학생들은 떨어졌지만 학과 활동을 열심히 한 학생들은 떨어지는 경우도 있었다.

학생들의 학점 및 복수전공 이수 여부 스터디 등의 다양한 수업관찰데이터를 이용하여 학생들의 임용시험 합격여부를 예측하고자 모델을 설계하였다. 텐서플로를 이용한 딥러닝을 활용하여 임용고사 합격여부를 예측하였지만 데이터의 빈약성과 컴퓨터교육과에 한정된 데이터라는 사안으로 모든 사범대학에 적용시킬 만큼 신뢰도를 얻기는 힘들 것이다. 그러나 위에서 기술하였듯이 단순히 공부만으로 임용시험에 합격하는 것이 아니고 학과활동으로 인하여 임용시험에 불이익이 있다는 것도 아니라는 점은 명확하게 알 수가 있다.

Ⅴ. 결 론

학생들의 임용시험 합격여부를 예측하는 것은 사범대학 학생들의 미래 및 대학생활을 하기 위하여 아주 중요한 일이다. 기존에는 선배들의 조언 및 합격 후기 등에 의존하다 보니 여러 가지 학과생활이 임용합격에 방해가 일 뿐이다 혹은 학과 수업보다는 인터넷 강의 및 강사들에 더 의존하면 된다는 등의 부정확한 정보에 학생들이 휩쓸려서 학과 생활 및 대학생활을 포기하고 학교를 하나의 학원 혹은 단순히 자격증을 받기 위한 수단으로만 여겨졌다.

본 연구에서는 딥러닝을 하기 위한 표본으로는 데이터의 숫자도 적고 변수의 숫자도 적다고 할 수 있다. 그러나 서문에서 밝혔듯이 일단 학생들에게 학과생활이 임용고사에 크게 방해가 되지 않는다는 것을 밝히기 위한 연구였기에 이 정도에서 만족을 하고 앞으로 학생들에게 임용고사의 가이드라인이나 합격을 위한 정확한 예측을 하기 위해서는 다양한 변수 및 많은 데이터가 필요 할 것이다. 연구를 좀 더 확장하여 학생들의 전공 세부학점 필수과목의 학점 및 교양 학점 등으로 세분화 시키고 학과 활동 및 복수전공 이수 여부 인터넷 강의 및 직강 수강 여부 등의 입력 항목을 좀 더 세분화 시킨다면 사범대학 학생들의 들의 임용고사 공부를 위한 지표 및 학교생활을 위한 이정표가 되어 줄 수 있을 것이라 사료된다. 더불어 학교에 남아있는 자료에 졸업생들의 직접적인 설문조사를 통하여 공부 시간 등에 대한 자료까지 추가된다면 좋을 것이다.

REFERENCES

[1] G. H. Kim, C. I. Jeong, C. S. Kim, E. J. Kim, and C. S. Kang, “A Study on the Competitive Admission of Candidates for Teacher's Employment by Using Deep Learning,” *Conference on korea information and communication engineering*, vol. 23, no. 1, pp.182-185, 2019.

[2] G. H. Kim, and E. J. Kim, “Design of a Hopeful Career Forecasting Program for the Career Education,” *International Journal of Information and Communication Engineering Journal of the Korea Institute of Information and Communication Engineering*, vol. 22, no. 8, pp.1055- 1060, 2018.

[3] C. S. Kim, *Big Data and Deep Learning practice practice*, 1th ed., Seoul, : JoEunGlTer, 2019.

[4] Y. B. Yang, “A Study on Prediction of Parent School Satisfaction Using Educational Data Mining,” Master. dissertation, Korea University, Seoul: DE, 2018.

[5] Y. J. Jung, S. M. Ahn, J. H. Yang, and J. J. Lee, “Comparison of Deep Learning Frameworks: About Theano, Tensorflow, and Cognitive Toolkit,” *Journal of Intelligence and Information Systems*, vol. 23, no. 2, pp. 1-17, 2017.

[6] Y. H. Lee, and D. H. Koo, “A Study on Development Deep Learning Based Learning System for Enhancing the Data Analytical Thinking,” *Journal of The Korean Association of Information Education*, vol. 21, no. 4, pp. 393-401, 2017.

[7] J. H. Seo, “A Comparative Study on the Classification of the Imbalanced Intrusion Detection Dataset Based on Deep Learning,” *Journal of Korean Institute of Intelligent Systems*, vol. 28, no. 2, pp. 152-159, Apr. 2018.

김근호(Geun-Ho Kim)

2009년 공주대학교 컴퓨터과학과 석사
2016년 공주대학교 컴퓨터교육학과 박사 수료
2011년~현재 공주대학교 컴퓨터교육과 조교
※관심분야 : 컴퓨터교육, 인공지능, 가상현실, 패턴인식

김의정(Eui-jeong Kim)

1997년 충남대학교컴퓨터공학 박사자통신공학과 공학박사
1997년~1998년 ETRI 연구원
1998년~현재 공주대학교 컴퓨터교육과 교수
※관심분야 : 컴퓨터비전, 패턴인식, 가상현실, 컴퓨터교육