

## 딥러닝 기반의 MBTI 직업 추천 알고리즘 설계

김준겸<sup>o</sup>, 조영복<sup>\*</sup>

<sup>o</sup>대전대학교 정보보안학과,

<sup>\*</sup>대전대학교 정보보안학과

e-mail: lab8456@gmail.com<sup>o</sup>, ybcho@djju.kr<sup>\*</sup>

## Design of MBTI Job Recommendation Algorithm Based on Deep Learning

June-Gyeom Kim<sup>o</sup>, Young-Bok Cho<sup>\*</sup>

<sup>o</sup>Dept. of Information Security, Deajeon University,

<sup>\*</sup>Dept. of Information Security, Deajeon University

### ● 요약 ●

본 논문에서는 성격, 성향을 근거로 사람의 성향에 따른 직업 및 전공에 대한 만족도를 분류한 데이터셋을 구축하여 사전에 사용자의 성향을 파악하여 직업을 추천하는 알고리즘을 제안한다. 성격유형검사 뿐만이 아닌 최근 게시한 SNS 텍스트를 사전에 학습한 데이터셋에 적용해 성격유형 결과의 정확도를 상승시키고자 한다. 사전에 생성한 데이터셋 외에 대상자가 작성한 정보(직업, 전공, 직업 및 전공에 대한 만족도)로 연합 학습을 진행하여 데이터셋의 정확도를 향상시키고자 한다. 모델의 학습 및 분류의 정확도 향상을 위해 SVM, NB, KNN, SDG 알고리즘들을 비교하였고 각각 67%, 21%, 28%, 69%의 정확도를 도출하였다. 데이터 셋은 캐글에서 제공받았다.

**키워드:** 작업부하(workload), 이주(migration), 부하균등화(load balancing)

## I. Introduction

2018년에 조사한 통계청의 자료에 따르면, 첫 직장에서 이직한 청년취업자들 중 적성, 전공 등과 직무와의 불일치로 이직을 한 인원은 인구중 14%나 차지했으며, 다양한 연구들로 현재 진로미결정의 요인들이 밝혀졌다. 진로미결정의 요인들엔 성격, 환경, 인지, 적성등의 요인으로 분류 된다. 취업의 실패는 상흔효과(scarring effects)와 심리적 좌절감으로 이어지며 첫 직장에서의 지위는 이후의 지위에도 영향을 미치게 될정도로 첫 취업은 인생에서 매우 중요하기에 미리 진로를 결정하는 것이 좋다. [1]. 진로를 결정하는 것은 자신의 성향, 목표를 미리 정하여 목표를 정하는 것으로 취업자 뿐만이 아닌 중, 고등학교 학생에게도 도움이 된다. 중, 고등학생 즉 청소년기에는 진로를 결정하는 것으로도 자율적 학습동기 수준, 자기 주도학습 수준이 진로를 결정하지 못한 학생들과 차이가 났다. 현재 학교 측에서도 다양한 직업 체험 교육을 운영하는 추세로 진로를 중요시 하는 모습을 보이고 있다. [2, 3] 본 논문에서는 청년 취업자들을 대상으로한 이직 이유와 진로미결정의 요인들 중 성격, 성향을 근거로 사람의 성향에 따른 직업 및 전공에 대한 만족도를 분류한 데이터셋을 구축하여 사전에 사용자의 성향을 파악하여 직업을 추천하는 알고리즘을 제안하며, 모바일 플랫폼에서 사용가능한 어플리케이션 환경에서

작동하도록 구상한다. 데이터셋은 Kaggle에서 (MBTI) Myers-Briggs Personality Type Dataset을 제공받아 사용하고자 한다.

## II. Preliminaries

### 1. Related works

#### 1.1 MBTI 텍스트 분류

MBTI(Myers\_Briggs Type Indicator) 성격유형검사는 검사 대상자가 본인의 성격을 판단하여 검사에 적용하기에 객관성이 부족할 수 있다. 이런 문제를 해결하기 위해 개인의 작성한 텍스트를 인공지능(deep learning)에 학습시켜 대상자의 성격유형을 분류하는 연구도 진행되고 있다[4, 5, 6]. 2022년 한국예선 네이버 MBTI 심리 카페에서 데이터(게시글 텍스트, 댓글 텍스트, 작성자 ID)를 수집하여 각 유형을 이분법으로 나누어 0과 1로 나누어 진행하는 성격유형 분류 연구가 진행되었다[4]. 텍스트 전처리를 위해 KoNLpy의 Okt,

KerasTokenizer같은 형태소 분석기를 사용해 전처리를 진행하였다. 해당 논문에서는 텍스트 내의 감정 분류를 위해 LSTM방식의 Deep Neural Network 구조를 사용하였으며 입력층에 입력된 텍스트를 임베딩 하였다. 해당 논문[5]에선 기타 다른 MBTI 분류 연구의 결과들과 비교하였을때 타 논문의 정확도는 20.29%, 21%, 38%, 45%의 정확도가 나온 반면 해당 논문은 66.59%의 정확도가 도출되었으며 이는 크롭한 한국어 텍스트 데이터로도 높은 정확도의 성격 유형 예측이 가능함을 나타냈다.

### III. The Proposed Scheme

그림 1은 본 논문에서 제안하는 프로그램의 구성도로 프로그램의 순서는 크게 4가지로 나뉜다.

- ① API요청
- ② 정보 입력
- ③ 유형 검사
- ④ 검사 결과 및 직업 추천

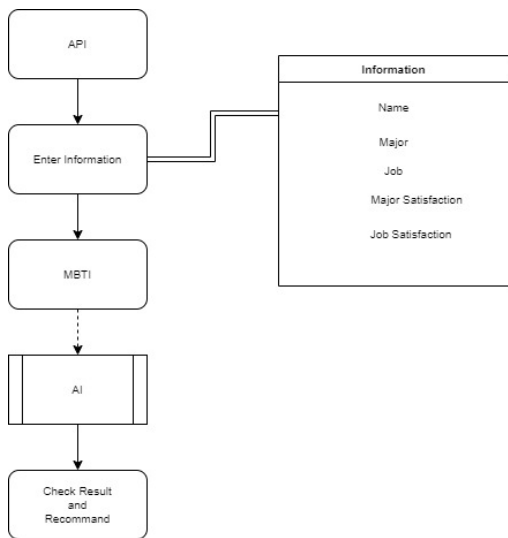


Fig. 1. System Configuration

제안하는 프로그램은 iOS, Android 두 OS에서 동작할 수 있도록 하이브리드 앱으로 개발할 수 있도록 애플리케이션 프레임워크로 React-native를 사용하고자한다. 표 3은 React-native를 구동하기 위한 두 CLI의 비교 분석 표로 각 CLI의 장, 단점을 구분해 놓았다. React-native를 구동하기 위한 모듈로써 React native-CLI와 Expo-CLI중 windows 환경에서 React native-CLI 모듈은 iOS를 구현하지 못한다는 제한사항으로 인해 Expo-CLI 모듈을 사용하여 구현하고자 한다.

Table 1. React Native CLI, Expo CLI

React Native CLI		Expo CLI	
Strengths	Weakness	Strengths	Weakness
Using a variety of libraries	Not convenient compared to Expo	Initial Configuration	Only features Provided by Expo are available
Native file control available	Initial configuration takes a long time	Automated file management	Module creation unavailable
Native file module available	Need to build a diverse development environment	Can be developed in one environment	Native files out of control
Difficult to deploy		Easy to ddeploy	

#### 1.1 알고리즘

본 논문에선 성격 유형을 분류하는 모델을 학습시키기 위해 Kaggle에서 Myers-Briggs Personality Type Dataset을 제공받아 학습하였다. 데이터셋은 MBTI 성격 유형과 텍스트로 구성됐으며, 총 데이터의 수는 8,678개로 본 논문에서는 7:3의 비율로 훈련 및 테스트 셋을 구성했다. 데이터 셋의 MBTI 성격 유형은 0~15의 숫자로 변환하며, 요청 받은 텍스트 데이터들과 단어 빈도수를 비교하여 MBTI 유형 분류의 신뢰도를 높인다. 알고리즘에 대입될 모든 데이터셋은 Count Vectorizer 모듈을 사용해 벡터화되며 불용어는 제거하는 방식으로 전처리가 진행된다. 분류 알고리즘은 SVM, SDG, NB, KNN 알고리즘을 이용하고 비교해 본 결과 SVM은 67%, SDG 69%, NB 21%, KNN 28%의 정확도가 나왔다.

	precision	recall	f1-score	support
0	0.74	0.49	0.59	41
1	0.79	0.72	0.75	67
2	0.80	0.22	0.35	18
3	1.00	0.12	0.22	8
4	0.77	0.52	0.62	33
5	0.65	0.48	0.55	54
6	0.00	0.00	0.00	10
7	1.00	0.22	0.36	9
8	0.68	0.71	0.70	294
9	0.69	0.80	0.74	366
10	0.69	0.57	0.63	135
11	0.62	0.26	0.37	38
12	0.62	0.77	0.69	218
13	0.70	0.83	0.76	261
14	0.69	0.62	0.65	137
15	0.67	0.30	0.42	46
accuracy			0.69	1735

Fig. 2. SDG result

	precision	recall	f1-score	support
0	0.63	0.66	0.64	41
1	0.68	0.73	0.71	67
2	0.67	0.22	0.33	18
3	0.00	0.00	0.00	8
4	0.84	0.48	0.62	33
5	0.54	0.48	0.51	54
6	0.00	0.00	0.00	10
7	0.67	0.22	0.33	9
8	0.67	0.71	0.69	294
9	0.60	0.96	0.71	366
10	0.72	0.50	0.59	135
11	0.73	0.21	0.33	38
12	0.76	0.63	0.69	218
13	0.76	0.77	0.77	261
14	0.66	0.71	0.69	137
15	0.79	0.24	0.37	46
accuracy			0.67	1735

Fig. 3. SVM result

	precision	recall	f1-score	support
0	0.00	0.00	0.00	41
1	0.00	0.00	0.00	67
2	0.00	0.00	0.00	18
3	0.00	0.00	0.00	8
4	0.00	0.00	0.00	33
5	0.00	0.00	0.00	54
6	0.00	0.00	0.00	10
7	0.00	0.00	0.00	9
8	0.00	0.00	0.00	294
9	0.21	1.00	0.35	366
10	0.00	0.00	0.00	135
11	0.00	0.00	0.00	38
12	0.00	0.00	0.00	218
13	0.00	0.00	0.00	261
14	0.00	0.00	0.00	137
15	0.00	0.00	0.00	46
accuracy			0.21	1735

Fig. 4. NB result

	precision	recall	f1-score	support
0	0.00	0.00	0.00	41
1	0.12	0.06	0.08	67
2	0.00	0.00	0.00	18
3	0.00	0.00	0.00	8
4	0.19	0.27	0.22	33
5	0.39	0.17	0.23	54
6	0.00	0.00	0.00	10
7	1.00	0.11	0.20	9
8	0.27	0.62	0.38	294
9	0.29	0.68	0.41	366
10	0.34	0.08	0.13	135
11	0.50	0.08	0.14	38
12	0.43	0.06	0.10	218
13	0.37	0.03	0.05	261
14	0.83	0.07	0.13	137
15	0.50	0.02	0.04	46
accuracy			0.29	1735

Fig. 5. KNN result

#### IV. Conclusions

본 논문에서는 취업자들을 대상으로 한 이직 이유와 진로미결정의 요인들 중 성격, 성향을 근거로 사람의 성향에 따른 직업 및 전공에 대한 만족도를 분류한 데이터셋을 구축하여 사전에 사용자의 성향을 파악하여 직업을 추천하는 알고리즘을 제안하며, 모바일 플랫폼에서 사용 가능한 어플리케이션 환경에서 작동하도록 구상한다. 성격유형 검사 뿐만이 아닌 최근 게시한 SNS 텍스트를 사전에 학습한 데이터셋에 적용해 성격유형 결과의 정확도를 상승시키고자 한다. 모델에 사용될 알고리즘으로 SDG, SVM, NB, KNN 알고리즘을 사용하였고 정확도를 비교한결과 SDG 알고리즘이 69%로 가장 높은 결과가 나왔다. 향후 MBTI와 홀랜드 직업 적성 검사의 연관성을 찾아 추가적

으로 직업 추천을 할 수 있도록 해야한다.

#### REFERENCES

- [1] J. M. Lee, H. J. Kil, "A Study on the effect of college graduate's education and job preparation experiences on the major-job match," *Korean Journal of Educational Adiministration*, vol. 37, no. 5, pp. 127-148, Dec. 2019.
- [2] C. J. Moon, M. W. Lee, "A Study on the factors affecting the results and the performance of middle and high school career experience," *The Journal of Career Education Research*, vol. 35, no. 2, pp. 1-22, Jun. 2022.
- [3] G. H. Kim, C. I. Jeong, "A Study on the Path Recommendation System Using MBTI and Performance," *Korea Institute of Information and Communication Engineering*, vol. 24, no. 1, pp.49-52, Jul. 2020.
- [4] J. M. Park, J. M. Kim, "Deep Learning Based MBTI Personality Type Classification Study," *Korean Journal of Educational Adiministration*, vol. 37, no. 5, pp. 127-148, Dec. 2019.
- [5] Choong EJ, Varathan KD. 2021. Predicting judging-perceiving of Myers-Briggs Type Indicator(MBTI) in online social forum. *PeerJ* 9:e11382 <https://doi.org/10.7717/peerj.11382>
- [6] Y. H. Zhang, B. H. Lee, "Job Recommender Service based on Deep Neural Network Reflecting MBTI Personality Type" *The Korean Academic Society of Business Administration*, Vol. 21, no. 4, pp. 99-113, Aug. 2021.