

# 자연어 기반 실시간 MBTI 성격 유형 분류를 위한 머신러닝 모델 성능 비교 연구

김건민<sup>1</sup>, 김예진<sup>1</sup>, 이은성<sup>1</sup>, 장주아<sup>1</sup>, 김정백<sup>2</sup>

전남대학교 소프트웨어공학과<sup>1</sup>

전남대학교 인공지능융합학과<sup>2</sup>

204869@jnu.ac.kr, ye031010@jnu.ac.kr, 200750@jnu.ac.kr, wndk0416@jnu.ac.kr,

kyungbaekkim@jun.ac.kr

## A Comparative Study of Machine Learning Models for Real-Time MBTI Classification Using Natural Language

Geonmin Kim<sup>1</sup>, Yejin Kim<sup>1</sup>, Eunseong Lee<sup>1</sup>, Jua Jang<sup>1</sup>, Kyungbaek Kim<sup>2</sup>

Dept. of Software Engineering, Chonnam National University<sup>1</sup>

Dept. of Artificial Intelligence Convergence, Chonnam National University<sup>2</sup>

### 요약

본 연구는 자연어 기반 실시간 대화형 MBTI 검사 시스템 구현을 목표로 전통적인 머신러닝 모델 (SVM, Logistic Regression, RandomForest, XGBoost)을 대상으로 정확도 및 F1-Score와 더불어 예측 시간 및 추론 시간을 함께 측정하여 성능을 종합적으로 평가한다. 실험 결과, Linear SVC와 Logistic Regression은 전반적으로 높은 정확도와 빠른 처리 속도를 보여 실시간 대화형 시스템에 적합한 모델로 확인되었다. 이를 통하여 자연어 기반 실시간 MBTI 예측 시스템 개발을 위한 모델 선택의 실증적 근거를 제공하며, 향후 다양한 언어권 데이터셋 적용, 불균형 데이터셋을 고려한 추가 연구의 필요성을 제시한다.

## 1. 서론

성격 유형은 개인의 행동 양식, 정보 처리 방식, 감정 표현 및 의사 결정에 광범위한 영향을 미친다. 현대 사회에서는 자기 이해와 타인과의 관계 유지 등을 위해 성격 유형에 대한 정확한 인식이 더욱 중요해지고 있다. 이에 따라 많은 사람들이 자신의 성격 유형을 보다 쉽게 확인하고 이해할 수 있도록 다양한 성격 유형 검사 도구들이 개발되고 활용되어 왔다.

대표적인 성격 유형 검사 도구로 MBTI 유형 검사 (Myers-Briggs Type Indicator)[1]가 있다. MBTI는 칼 융의 심리유형론을 기반으로 개발된 심리 검사로, 개인의 성격을 외향(E)/내향(I), 직관(N)/감각(S), 사고(T)/감정(F), 판단(J)/인식(P)이라는 네 가지 이분법적 차원으로 나누는 후 이를 조합하여 총 16가지 성격 유형으로 분류한다.

기존의 MBTI 검사는 대체로 고정된 문항을 기반으로 응답을 유도하여 검사자가 질문에 대한 답변을 의도적으로 왜곡하거나, 문항 구조에 익숙해진 상태로 응답하는 등의 문제가 발생할 소지가 있다. 이러한 한계를 극복하고 사용자 개인의 성격을 보다 직관적으로 분석하기 위하여 최근에는 인공지능 및 자연어 처리(NLP) 기법을 접목하여 대화형 방식의 성격 유형 추론에 대한 연구가 활발히 진행되고 있다[2]. 대화형 검사 방식은 사전 정의된 문항이 아닌 사용자의 자연어 답변에서 텍스트의 특징을 추출하고, 이를 기반으로 성격 유형을 예측한다.

이러한 실시간 대화형 방식에서의 확장을 위해서는 반응 속도와 정확도가 모두 뒷받침되어야 한다. 이 두 가지 요소가 동시에 만족되지 않는다면, 사용자 경험 측면에서 실질적인 대화형 MBTI 검사 시스템으로의 활용이 어렵다. 따라서 본 논문에서는 실시간 대화형 MBTI 검사 시스템 구축을 위한 다양한 머신러닝 모델의 성능을 실험적으로 비교하여 정확성과 실시간 반응성을 동시에 확보할 수 있는 최적의 모델 구조를 탐색한다.

## 2. 본론

### 2.1 기존 연구의 한계

MBTI 성격 유형 분류를 위한 텍스트 기반 머신러닝 연구는 다양한 접근법으로 진행되어 왔다. 전통적인 머신러닝 기법부터 트랜스포머 기반의 딥러닝 모델에 이르기까지 다양한 연구가 진행되었으나, 대다수의 기존 연구들은 정확도와 F1-Score 등의 지표에 중점을 두어 모델을 평가하는 경향을 보여왔다[3]. 그러나, 실시간 대화형 서비스의 측면에서는 빠른 응답을 보장하기 위하여 실시간 반응성과 처리 속도에 대한 고려도 필요하다.

### 2.2 실험 계획

이러한 문제 의식을 바탕으로, 본 연구에서는 다양한 분류 모델의 실시간 서비스 도입 적절성을 평가하기 위하여 네 가지 모델(Linear SVC, LogisticRegression,

RandomForest, XGBoost)를 대상으로 정확도와 F1-Score 그리고 예측 시간과 추론 시간을 추가로 측정하는 실험을 진행한다. 추론 시간은 각 모델이 예측 시간을 바탕으로 1000개의 입력을 처리할 수 있는 시간이다. 학습 데이터로 Myers-Briggs Personality Type Dataset(MBTI\_1) MBTI Personality Types 500 Dataset(MBTI\_500)를 바탕으로 둘을 결합한 Combined Dataset을 추가로 구성한다. 각 데이터셋에 담긴 텍스트에 소문자 변환, 특수문자 제거, 불용어 제거, 표제어 추출 등의 전처리 과정을 적용하여 데이터 형식의 일관성을 확보한다.

(표 1) 데이터셋 별 샘플 수

Dataset	Count
Myers-Briggs Personality Type Dataset	8,675
MBTI Personality Types 500 Dataset	106,067
Combined Dataset	122,553

### 2.3 실험 결과

(표 2) 모델 별 Accuracy 및 F1-Score

	Dataset	Accuracy	F1-Score
Linear SVC	MBTI_1	0.6367	0.6285
	MBTI_500	0.8152	0.8139
	Combined	0.8099	0.8086
Logistic Regression	MBTI_1	0.6505	0.6283
	MBTI_500	<b>0.8204</b>	<b>0.8175</b>
	Combined	<b>0.8138</b>	<b>0.8107</b>
RandomForest	MBTI_1	0.6367	0.6311
	MBTI_500	0.7245	0.7224
	Combined	0.7223	0.7213
XGBoost	MBTI_1	<b>0.6679</b>	<b>0.6543</b>
	MBTI_500	0.7807	0.7777
	Combined	0.7773	0.7746

(표 3) 모델 별 예측 시간 및 추론 시간

	Dataset	Prediction Time(ms)	Inference Time(ms)
Linear SVC	MBTI_1	3.0	1.71
	MBTI_500	<b>25.70</b>	<b>1.21</b>
	Combined	28.50	<b>1.24</b>
Logistic Regression	MBTI_1	<b>2.60</b>	<b>1.5</b>
	MBTI_500	26.10	1.23
	Combined	<b>28.40</b>	<b>1.24</b>
RandomForest	MBTI_1	59.50	34.31
	MBTI_500	806.60	38.03
	Combined	771.50	33.62
XGBoost	MBTI_1	14.80	8.51
	MBTI_500	112.20	5.29
	Combined	116.10	5.06

네 가지 모델은 앞서 언급한 세 가지 데이터셋에 대해 각각 학습과 검증을 거쳤다. 실험 결과, 대용량 데이터를 처리하는 상황에서 빠른 추론 속도를 보이며 정확도와 F1-Score가 고르게 높은 것은 Linear SVC와 Logistic Regression이었다. 이 둘은 빠른 처리 속도를 가지면서도

정확도와 F1-Score에서 우수한 성능을 보였다. 반면, Random Forest와 XGBoost 모델은 정확도와 F1-Score 측면에서 경쟁력 있는 수치를 달성하였으나, 예측 시간 및 추론 시간이 비교적 길었다. 이는 데이터셋 규모가 클수록 앙상블 모델의 트리 구조가 복잡해지는 특성에서 기인한 것으로 추정된다. 실험 결과, 정확도 유지 및 실시간 처리에 LinearSVC, Logistic Regression가 적절함을 확인하였다.

### 3. 결론 및 향후 연구

본 연구는 대규모 텍스트 데이터를 짧은 시간 안에 처리해야 하는 대화형 MBTI 분류 시스템 개발에 있어 정량적 성능 지표와 실시간 응답성을 동시에 고려해야 함을 논하고, 이를 위하여 사용자의 자연어 텍스트를 활용해 MBTI 성격 유형을 분류하는 머신러닝 모델의 정확도와 실시간 응답성을 측정하는 실험을 진행하였다. 실험 결과, Linear SVC와 Logistic Regression이 평균적으로 가장 높은 정확도와 F1-Score, 그리고 빠른 추론 속도를 모두 충족함을 확인하였다.

향후 연구에서는 다양한 언어권의 텍스트 데이터셋을 활용한 모델 검증이 필요하다. 또한, 16가지 유형의 데이터가 균등하게 분포하지 않는 점을 고려하여 오버샘플링이나 데이터 증강 기법 등을 통해 데이터셋을 보완하는 연구가 요구된다. 또한, 보다 정교한 성격 유형 분류를 위하여 BERT, RoBERTa 등 사전학습 기반 딥러닝 모델을 적용하거나, 사용자 대화 데이터를 기반으로 한 LLM(Language Model) 파인튜닝을 통해 모델의 문맥 이해도와 분류 성능을 향상시키는 연구도 진행할 계획이다.

### ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 소프트웨어중심대학사업의 연구결과로 수행되었습니다.(2021-0-01409)

### 참고 문헌

[1] Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator (3rd ed.). Consulting Psychologists Press.

[2] A. S, A. M. Joy and J. N. D'Souza, "Prediction of MBTI with textual data using different pre-trained transformer models," 2023 Fourth International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), Bengaluru, India, 2023, pp. 1-7

[3] M. N. Sahono et al., "Extrovert and Introvert Classification based on Myers-Briggs Type Indicator(MBTI) using Support Vector Machine (SVM)," 2020 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 2020, pp. 572-577