

뇌심혈관 질환과 업무 환경의 연관성 판단을 위한 AI 모델의 개발 및 전문가 판단과의 일치도 분석

오주연 · 유기봉^{1*} · 진익훈^{2*} · 윤병윤³ · 심주호³ · 박희주 · 이종민⁴ · 이지안 · 윤진하³

연세대학교 일반대학원 보건학과, ¹연세대학교 보건행정학부, ²연세대학교 응용통계학과,
³연세대학교 의과대학 예방의학교실, ⁴연세대학교 보건대학원 산업보건전공
*유기봉과 진익훈은 공동 2저자로 이 논문에 동등하게 기여하였음.

Development of an AI Model to Determine the Relationship between Cerebrovascular Disease and the Work Environment as well as Analysis of Consistency with Expert Judgment

Juyeon Oh · Ki-bong Yoo^{1*} · Ick Hoon Jin^{2*} · Byungyoon Yun³ ·
Juho Sim³ · Heejoo Park · Jongmin Lee⁴ · Jian Lee · Jin-Ha Yoon³

Department of Public Health, Graduate School, Yonsei University, Seoul, Republic of Korea

¹Department of Health Administration, Yonsei University, Wonju, Republic of Korea

²Department of Applied Statistics, Yonsei University, Seoul, Republic of Korea

*³Department of Occupational Health, Department of Preventive Medicine,
Yonsei University College of Medicine, Seoul, Republic of Korea*

*⁴Department of Occupational Health, Graduate School of Public Health, Yonsei University,
Seoul, Republic of Korea*

**Ki-bong Yoo and Ick Hoon Jin equally contributed to this work as co-second authors.*

ABSTRACT


Introduction: Acknowledging the global issue of diseases potentially caused by overwork, this study aims to develop an AI model to help workers understand the connection between cerebrocardiovascular diseases and their work environment.


Materials and methods: The model was trained using medical and legal expertise along with data from the 2021 occupational disease adjudication certificate by the Industrial Accident Compensation Insurance and Prevention Service. The Polyglot-ko-5.8B model, which is effective for processing Korean, was utilized. Model performance was evaluated through accuracy, precision, sensitivity, and F1-score metrics.


Results: The model trained on a comprehensive dataset, including expert knowledge and actual case data, outperformed the others with respective accuracy, precision, sensitivity, and F1-scores of 0.91, 0.89, 0.84, and 0.87. However, it still had limitations in responding to certain scenarios.


*Corresponding author: Jin-Ha Yoon, Tel: +82-2-2228-5179, Fax: +82-2-392-8622, E-mail: flyinyou@gmail.com
Department of Occupational Health, Department of Preventive Medicine, Yonsei University College of Medicine 50-1
Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea


Received: April 9, 2024, Revised: June 30, 2024, Accepted: September 11, 2024


 Juyeon Oh <https://orcid.org/0000-0002-9843-4928>


 Ick Hoon Jin <https://orcid.org/0000-0003-0647-5704>


 Juho Sim <https://orcid.org/0000-0001-7652-4482>


 Jongmin Lee <https://orcid.org/0009-0004-7577-2686>

 Jin-Ha Yoon <https://orcid.org/0000-0003-4198-2955>

 Ki-bong Yoo <https://orcid.org/0000-0002-2955-6948>

 Byungyoon Yun <https://orcid.org/0000-0001-7055-6424>

 Heejoo Park <https://orcid.org/0009-0006-5365-6435>

 Jian Lee <https://orcid.org/0009-0003-9089-895X>

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Discussion: The comprehensive model proved most effective in diagnosing work-related cerebrocardiovascular diseases, highlighting the significance of integrating actual case data in AI model development. Despite its efficacy, the model showed limitations in handling diverse cases and offering health management solutions.

Conclusion: The study succeeded in creating an AI model to discern the link between work factors and cerebrocardiovascular diseases, showcasing the highest efficacy with the comprehensively trained model. Future enhancements towards a template-based approach and the development of a user-friendly chatbot webUI for workers are recommended to address the model's current limitations.

Key words: Cerebrovascular disease, occupational disease, artificial intelligence, large language model

1. 서 론

뇌심혈관 질환은 심혈관질환과 뇌혈관질환을 아울러 이르는 용어로, 혈관과 관련된 일부 병리학적 상태로 인해 발생하는 심장 및 뇌 기능의 저하로 광범위하게 정의되며(Mora & Bornstein, 1998), 고도의 스트레스, 과도한 업무부담 등이 주요 원인으로 꼽힌다. 이 질환은 전 세계적으로 주요 사망 원인 중 하나로(Deaton et al., 2011; Thrift et al., 2014; Roth et al., 2018), 세계보건기구(WHO)에 따르면 뇌심혈관 질환은 2019년 전세계 사망자의 32%를 차지하였다(WHO, 2017). 한국에서도 뇌심혈관 질환이 최근 10년 동안 주요 사망원인으로 꼽혔으며, 통계청 2021년 사망 원인통계에 따르면 심장질환으로 인한 사망이 전체 사망 원인의 2위를 차지하였다(KOSIS, 2022).

뇌심혈관 질환은 개인의 유전적 소인, 생활 습관, 기존의 건강 상태와 같은 내부 요인에 기인하게 되는 경우와 함께, 외부 환경적 요인, 특히 근무 환경이 이러한 질환의 발병에 중요한 역할을 하는 경우가 많다(Vineis et al., 2020; Kim et al., 2022; Wu et al., 2022). 이러한 연관성은 직업병으로 분류될 수 있는데, 직업병이란 특정한 업무 환경이나 작업 특성상 노출되는 물질이나 조건에 의해 발생하는 질병을 의미한다(Cherry, 1999). 근무 환경은 뇌심혈관 질환의 발병에 중요한 역할을 하며, 특히 장시간 근무와 업무부담은 질병 발병의 위험을 증가시킨다. 국내에서도 고용노동부에서 고시한 뇌심혈관 질환의 업무상 질병 인정에 필요한 사항에 따르면, 근로자가 만성적인 과중한 업무로 인해 뇌심혈관 질환이 발병된 경우, 업무상 질병으로 인정한다. 국내외로도 장시간 근무 및 업무부담가 중요인인과 뇌심혈관질환과의 연관성을 확인한 연구가 활발히 진행되었다(Park et al., 2015; Lin et al., 2018; Kim & Jung, 2021; Kamardeen, 2022). 이렇게 특정 직업

환경이 뇌심혈관 질환의 발병 및 악화에 큰 영향을 끼칠 수 있기 때문에, 직업병의 범주 내에서 근로자들의 뇌심혈관 질환의 발병 원인 및 관련 요인을 파악하고 대응하는 것이 중요하다고 판단된다.

뇌심혈관 질환의 발병 이해는 깊은 의학적 지식과 전문적 판단을 요구하지만(Picciotto, 2010) 대부분의 근로자는 이러한 지식이 부족해 자신의 건강 문제와 작업 환경 간의 연관성을 정확히 인식하기 어렵다. 또한 종종 일상 생활 습관이나 유전적, 환경적 요인 중에서 직장 환경의 영향을 간과하고, 업무 중 발생하는 스트레스나 사건들이 건강에 장기적으로 미치는 영향을 인지하지 못한다. 마지막으로 직업병으로 인정받기 위한 과정의 복잡성, 높은 비용, 경제적 위험, 그리고 직장 내에서 부정적인 인식 등은 근로자들이 자신의 질병과 업무 환경 간의 연관성을 주장하는 것을 어렵게 만든다(Guthrie & Westaway, 2009; Kelly & Romero, 2019). 이러한 여러 요인이 근로자들이 자신의 질병과 업무 환경 간의 연관성을 제대로 인식하고 이를 주장하는 것을 어렵게 만들고 있다.

이러한 문제의식에서 출발하여, 전문지식이 없는 근로자들도 경제적 및 심리적 부담 없이 자신의 건강 문제와 업무 환경과의 관계를 쉽게 이해하고 판단할 수 있게 도와주는 도구의 필요성이 대두되었다. 인공지능(artificial intelligence, AI) 기술의 발전은 이러한 문제 해결에 획기적인 방안을 제공할 수 있다는 가능성을 내비치고 있다. 이 기술들은 거대하고 다양한 데이터 세트에서 유의미한 패턴을 찾아내는 능력을 가지고 있으며(Zhuang et al., 2017), 이를 통해 개인의 건강 데이터와 작업 환경 데이터를 종합적으로 분석하여 근로자 개개인의 뇌심혈관 질환에 대한 업무상 질병을 판정하는 모델을 구축할 수 있다.

그러나 AI 모델은 특정 도메인의 전문적 지식을 내재화하는 데 한계가 있으며, 이는 업무상 질병 판정과 같

이 전문적이고 특수한 지식을 요하는 분야에서는 AI 모델을 해당 분야에 맞게 조정하는 과정이 필요하다 (Trivedi & Mishra, 2022). 또한 이러한 AI 모델의 판단이 실제 전문가의 판단과 얼마나 일치하는지에 대한 연구도 아직 부족하여, AI 모델의 판단에 대한 신뢰성 확보가 필요하다. 따라서 AI 모델의 판단 기준이 전문가의 의견과 얼마나 일치하는지를 검증하고, 근로자가 본인의 뇌심혈관 질환과 업무관련성을 스스로 판단할 수 있는 능력과 건강관리 인식을 향상시키는 데 AI 모델이 어떻게 기여할 수 있는지 탐구해야 한다. 이에 따라 본 연구의 목적은 다음과 같다.

뇌심혈관 질환의 업무상 질병 판정 AI 모델의 개발: 근로자의 뇌심혈관 질환과 업무 환경의 연관성을 판단하기 위한 AI 모델을 개발한다. 이 모델은 근로자들의 업무 환경 데이터를 기반으로 그들의 뇌심혈관 질환과 업무 환경 간의 관계를 예측하도록 설계될 것이다.

AI 모델의 정확성 평가: 개발된 AI 모델의 예측 성능을 평가한다. 이를 위해, 실제 근로자들의 건강 데이터와 업무 환경 데이터를 모델에 적용하여 그 결과를 실제 판정 결과와 비교한다.

전문가 의견의 질적 분석: AI 모델의 판단 결과와 직업환경의학 전문가의 판단 사이의 일치도를 분석한다. 여기서는 AI 모델이 얼마나 전문가의 의견과 일치하는지, 그리고 어떤 경우에 불일치하는지를 중점적으로 조사하고자 한다.

II. 연구 방법

1. 모델 학습 데이터 및 전처리 과정

AI 모델 개발을 위해 사용된 데이터는 세 가지 주요 유형으로 구성되어 있다: 의학적 및 법적 전문 지식, 근로복지공단의 업무상질병판정서에서 얻은 실제 사례, 그리고 업무상 질병 판정 신청 사례와 판정 결과를 포함한 2021년도 근로복지공단 업무상질병판정서 홈페이지의 정보다. 뇌심혈관 질환의 업무상 질병 인정기준은, 산업재해보상보험법(업무상 재해)의 제 34조 제 3항 및 제 37조 제 5항(2023a; 2023b)에 따라 뇌심혈관 질환의 업무상 질병 인정 기준을 고시한 고용노동부고시 제 2022 - 40호 『뇌혈관 질환 또는 심장 질환 및 근골격계 질환의 업무상 질병 인정 여부 결정에 필요한 사항』을 참고하였다. 모델 학습에 사용한 뇌심혈관 질환의 업무상 질병 인정기준의 상세 내용은 다음과 같다.

- ① 발병 전 12주 동안 근로 시간이 1주 평균 60시간 초과 또는 발병 전 4주 동안 1주 평균 64시간을 초과한 경우에 업무상 질병과의 관련성이 강하다고 평가함.
- ② 발병 전 12주 동안 1부 평균 근로 시간이 52시간을 초과하는 경우, 근로 시간이 길수록 업무상 질병과의 관련성이 증가하는 것으로 평가함. 특히 다음 업무부담 가중요인 중 어느 하나에 해당하는 경우, 업무상 질병과의 관련성이 강하다고 평가함.
 - 업무가 돌발적이고 예측이 어려운 업무
 - 교대제 업무
 - 휴일이 부족한 업무
 - 유해한 작업환경
 - 육체적 강도가 높은 업무
 - 시차가 큰 출장이 잦은 업무
 - 정신적 긴장이 높은 업무
- ③ 발병 전 12주 동안 근로 시간이 1주 평균 52시간을 초과하지 않더라도 업무부담 가중요인에 복합적으로 노출되는 업무의 경우, 업무상 질병과의 관련성이 증가한다고 평가함.
- ④ 오후 10시부터 오전 6시 사이의 야간근무 경우, 주간 근무의 30%를 가산하여 근로 시간을 산출함.

업무상질병판정서는 근로자의 개인 정보, 건강검진 이력, 생활 습관, 근무 시간, 업무부담과 같은 상세 정보를 제공하여 질병 발생 배경과 업무와의 연관성을 평가하는 데 사용된다. 데이터 전처리 과정에서는 근로자의 정보와 업무상 부담을 포함한 인정사실을 gpt-3.5-turbo 모델을 이용하여 요약하고, 업무적 정보만을 바탕으로 한 판단 사례를 추출하여 위원회의 판단 및 결론을 출력 형식으로 준비한다. 이 과정을 거쳐, 구어체로 변환된 인정사실(Instruction)과 위원회의 판단 및 결론(Output)을 포함한 업무상질병판정서가 모델 학습 자료로 사용된다. 최종적으로 1,000개의 업무상질병판정서 중 800개를 학습 데이터로, 나머지 200개를 테스트 데이터로 활용하여 모델의 학습과 검증을 수행한다. 모델 학습에 사용된 업무상질병판정서의 예시는 다음과 같다.

“Instruction:

저는 발병 전 1주간의 업무 시간은 35시간이었고, 이는 발병 전 12주간의 주당 평균 업무 시간인 28시간

37분과 비교하여 30% 미증가되었습니다. 또한, 업무 환경의 변화나 특이사항은 없었습니다. 발병 전 4주 동안의 업무 시간은 1주 평균 33시간 18분이었고, 발병 전 12주 동안의 업무 시간은 1주 평균 29시간 11분이었습니다. 이 기간에도 특별한 업무 환경의 변화나 특이사항은 없었습니다. 또한, 저는 업무부담 가중요인이 없었습니다.

이러한 정보를 바탕으로 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?

Output:

발병 당일 돌발 상황이나 급격한 업무 환경 변화는 없었고 발병 직전 1주간 업무 시간은 35시간 21분으로 일상 업무 시간보다 30% 이상 증가된 사실 확인되지 않으며, 발병 전 4주 및 12주 동안 1주 평균 근무시간은 각 33시간 18분 및 29시간 11분으로 만성 과로에도 해당하지 않는다. 그 외 업무부담 가중요인은 객관적으로 확인되지 않는다. 따라서 신청인의 질병은 업무상 질병으로 불인정한다.”

본 연구는 연세의료원 연구심의 위원회(Institutional Review Board, IRB) 심의에서 IRB 면제심의 승인을 받아 진행되었다 (IRB number: 4-2023-1317).

2. 기반 모델 및 Fine-tuning

이 연구에서는 대규모 언어 모델의 한계를 극복하고, 한국어에 특화된 AI 모델 개발을 목표로 하였다. 현재 시장에는 GPT와 LLaMA와 같은 다양한 대규모 언어 모델이 존재하지만, 이들은 주로 영어 데이터에 기반하여 학습되어 비영어권 언어에 대한 성능이 상대적으로 낮고, 고용량의 메모리 및 컴퓨팅 자원을 요구한다(Lin et al., 2021; Joo et al., 2023; Ko et al., 2023). 이에 대한 대안으로, EleutherAI의 Polyglot-ko-5.8B 모델을 선택하였다. 이 모델은 1.2TB의 한국어 데이터로 학습된 것으로, 256개의 A100 GPU에서 학습되었으며, 특히 한국어 자료에 대해 높은 성능을 보인다(Ko, Yang, et al., 2023). 비록 더 높은 성능을 가진 Polyglot-ko-12.8B 모델도 존재하지만, 그 큰 용량 때문에 5.8B 모델이 연구 목적에 더 적합하다고 판단하였다.

언어 모델의 성능과 유용성은 특정 분야나 주제에 대한 지식을 얼마나 잘 이해하고 반영하는지에 크게 좌우된다 따라서 모델의 성능을 뇌심혈관 업무상질병 판정 분야에 맞추어 최적화하기 위해 fine-tuning 과정을

진행하였다. Fine-tuning이란, 이미 학습된 모델에 특정한 데이터를 추가로 제공하여, 모델이 그 데이터에 특화된 지식을 습득하게 하는 과정이다(Trivedi & Mishra, 2022). 이 과정에서는 뇌심혈관 질환에 관한 전문적 지식과 업무상 질병 판정 데이터를 추가로 학습시켜, 모델이 업무상 질병 판정에 필요한 전문 지식을 습득하도록 하였다. 이를 통해 모델은 업무상 질병 판정의 패턴을 배우고, 판정의 정확성을 크게 향상시킬 것으로 기대된다.

3. 학습 데이터에 따른 모델 결과 비교

이 연구는 언어 모델의 성능과 활용 범위를 세 가지 다른 시나리오를 통해 비교하고자 한다(Figure 1). 첫 번째 시나리오에서는 어떠한 데이터도 학습시키지 않은 초기 상태의 Polyglot-ko-5.8B 모델에게 근로 시간과 업무부담 가중 요인을 제공한 후, 뇌심혈관 질환의 업무상 질병 인정 여부에 관해 묻는다. 이 모델은 뇌심혈관 질환과 관련된 정보나 지식이 없는 상태이므로 적절하지 못한 응답을 생성할 가능성이 높다. 두 번째 시나리오에서는 뇌심혈관 질환에 대한 의학적 지식과 업무상 질병 법적 인정 기준을 학습한 모델을 사용한다. 이 모델은 뇌심혈관 질환과 업무상 질병 인정에 관한 기본 지식을 갖추고 있으며, 근로자의 근로 시간과 업무 요인을 입력했을 때 업무상 질병 판정이 가능한지 확인한다. 세 번째 시나리오에서는 뇌심혈관 질환에 대한 의학적 지식과 업무상 질병 법적 인정 기준에 더하여 실제 근로자들의 근로 시간과 업무 요인에 대한 근로복지공단의 업무상질병판정서 데이터를 학습한 모델을 사용한다. 이 모델은 전문지식에 더해 다양한 실제 인정 사례를 학습하여, 두 번째 모델보다 더 정확한 판단이 가능한지 확인한다.

4. 모델의 성능 평가 및 전문가 의견과의 일치도 판정

AI 모델을 사용하여 업무상질병의 인정 여부를 판단할 때, 그 성능 평가는 모델의 신뢰성과 적용성을 확보하기 위한 필수적인 절차로 매우 중요하다. 성능 평가를 위해 AI모델의 판단을 실제 업무상질병판정서의 결론과 비교하여 다양한 지표들을 통해 성능을 평가하고자 하였다. 모델에서 출력된 문장을 읽고 인정 또는 불인정으로 이분류하여, 모델의 정확도(accuracy), 정밀도, 민감도, F1-score를 계산하였다. 또한 실제 작업환경의학 전문가의 판단 및 이에 대한 의견을 구하여 AI

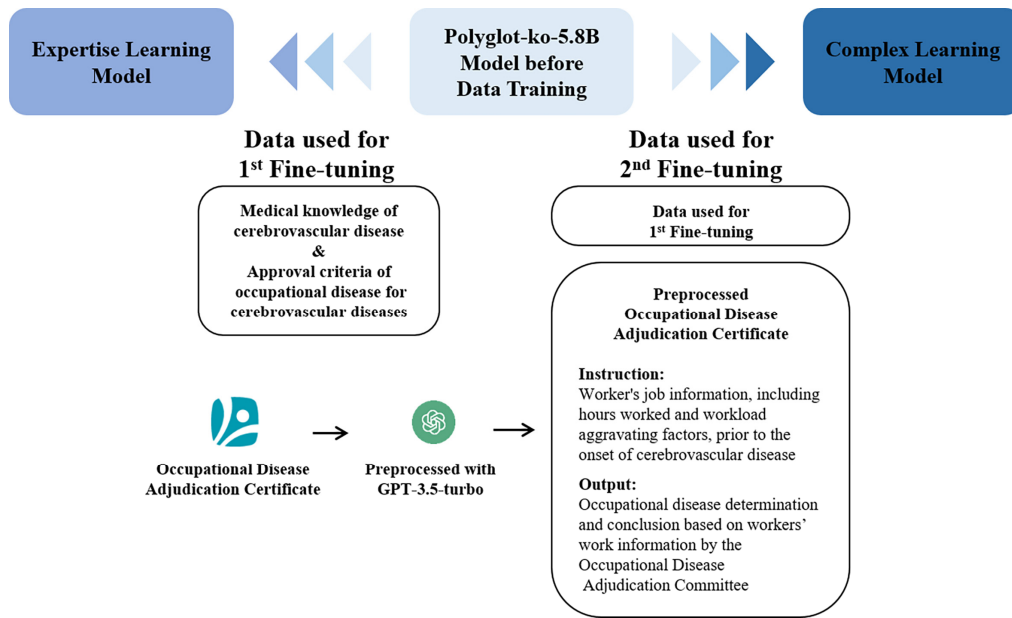


Figure 1. Preprocessing of training data using the GPT-3.5-turbo model

모델의 판단과 전문가의 판단과의 일치도를 판단해 보고자 한다. 각 사례에 대해 AI모델이 내린 판단과 전문가가 내린 판단을 비교한다. 이를 통해 어느 사례에서 일치하고 어느 사례에서 불일치하는지, 불일치한다면 어느 이유로 불일치하는지 구체적으로 파악해 보고자 한다.

5. 사용 프로그램

업무상질병판정서의 전처리를 위해서 openAI의 GPT-3.5-turbo 모델을 이용하였고, 이는 R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria)에서 API 활용 방식으로 모델을 사용하였다. 또한 Polyglot-ko-5.8B 모델을 학습시키기 위해 Colab Pro에서 A100 GPU 40GB RAM에서 학습을 수행하였다.

III. 결 과

전문지식 학습모델은 예폭 수 10으로 학습하였으며, 학습 결과 training loss 0.0763으로 학습 완료되었다. 복합 학습 모델은 전문지식 학습모델의 데이터에 더하여 300건의 업무상 질병 인정사례와 500건의 업무상 질병 불인정 사례로 이루어진 업무상질병판정 사례를 예폭 수 10으로 학습하였으며, 학습 결과 최종 training loss 0.225로 학습 완료되었다. 각 모델에 대한 테스트는 70건의 업무상 질병 인정 사례와 130건의 업무상 질병 불인정 사례로 이루어진 테스트 데이터로 수행하였다. 각각의 모델은 동일한 테스트 데이터에 대해 판정을 진행하였다.

Table 1에서는 실제 업무상 질병 판정과 비교한 세

Table 1. Comparison of model responses with actual occupational disease adjudication outcomes

Consistent with actual adjudication	Actual adjudication of occupational disease adjudication certificate	Model before training	Expertise learning model	Complex learning model
Consistent	Approval	46	56	59
	Disapproval	74	78	123
Inconsistent	Approval	11	6	11
	Disapproval	32	34	7
Unable to judge	Approval	13	8	0
	Disapproval	24	18	0

Table 2. Result of evaluation metrics for models

	Model before training	Expertise learning model	Complex learning model
Accuracy	0.6	0.67	0.91
Precision	0.66	0.8	0.89
Sensitivity	0.45	0.52	0.84
F1-score	0.53	0.63	0.87

모델의 응답을 비교한 결과를 확인할 수 있다. 학습 이전 모델은 200건 중 총 80건이 실제 판정과 결과가 일치하지 않았고, 전문지식 학습 모델은 66건이 실제 판정과 일치하지 않았다. 복합 학습 모델은 총 18건이 실제 판정과 결과가 일치하지 않았다. 학습 이전 모델은 불일치 80건 중 37건, 전문지식 학습 모델은 70건 중 26건이 인정과 불인정을 판정하지 못하고 전문가의 판단이 필요하거나, 판단을 위해서는 업무와 관련한 정보가 더 필요하다는 응답을 출력하였다.

모델의 성능을 측정하기 위해 실제 판정과 각 모델의 응답 값을 비교하여 정확도, 정밀도, 민감도, 그리고 F1-score를 계산하였다. 모델의 응답을 읽고 업무상 질병으로 인정하였는지 불인정하였는지 구분하여 인정과 불인정 두 그룹으로 분류하여 계산하였다. 학습 이전 모델과 전문지식 학습 모델에서 업무상 질병 판정을 하지 못한 사례에 대해서는 실제 판정과 반대로 판정하였다고 설정한 후 두 지표를 계산하였다.

계산 결과, 학습 이전 모델은 정확도 0.60, 전문지식 학습 모델은 정확도 0.65로 계산되었으며, 복합 학습 모델은 0.91로 복합 학습 모델이 가장 높았다. 정밀도는 모델 순서대로 0.66, 0.80, 0.89로 복합 학습 모델이 제일 높았고 그다음으로 전문지식 학습 모델이 가장 높았다. 민감도는 모델 순서대로 0.45, 0.52, 0.84로 복합 학습 모델이 가장 높았고, F1-score는 순서대로 0.53, 0.63, 0.87으로 복합 학습 모델이 가장 높게 계산되었다. 이에 대한 내용은 Table 2에서 확인할 수 있다.

복합 학습 모델에서, 실제 판정과 비교하여 응답이 달랐던 사례에 대하여 왜 모델의 응답이 실제 판정과 다르게 나왔을지에 대한 작업환경의학 전문가의 의견을 구하였다. 모델이 실제로 응답으로 도출했던 문장을 읽고, 전문가와 함께 각 사례의 특성과 모델의 판정 과정, 그리고 실제 판정과 차이점을 검토하였다. 실제 인정 판정과의 불일치 건 중 3건에 대한 전문가의 의견은 다음과 같다.

Instruction:
 발병 전 4주간 및 12주간 주당 평균 업무 시간은 54시간이었습니다. 업무의 양, 강도, 책임 및 업무 환경의 변화 등에 따른 부담 요인으로 정신적 긴장을 겪었습니다. 이러한 정보를 바탕으로 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?

Output:
 발병 전 12주간 동안 및 업무 환경의 변화 등에 따른 부담 요인으로 정신적 긴장을 겪었지만, 발병 전 1주 동안 및 업무 환경의 변화 등에 따른 부담 요인으로는 업무량 증가 및 작업 환경의 변화는 없었습니다. 또한, 발병 전 1주일 동안 및 업무 환경의 변화 등에 따른 부담 요인으로는 업무 시간이 30% 이상 증가하지 않았으며, 업무 환경의 변화 및 업무상 부담의 증가 없었습니다. 따라서, 신청인의 뇌심혈관 질환은 업무상 질병으로 불인정 된다.

직업환경의학 의사 의견:
 질문을 줄때 법적 필수 요소인, 단기 과로 증가 (30% 증가) 없는 경우 없다고, 또는 모른다고 질문에 명시하는 방안, 부가 요인 (환경변화, 스트레스, 물리적 노출)에 대해서는 모델 학습이 될 되었을 가능성이 있음. 상기 정보에 대해서는 추가 데이터를 수집하여 training 할 필요가 있다고 보임.

모델과 실제 판정의 첫 번째 불일치 사례에서, 모델은 발병 전 12주 동안의 부담 요인을 파악했지만, 발병 전 1주 동안의 업무부담 정보가 부족했다고 판단했다. 전문가는 이 정보의 부재로 모델이 판정을 내릴 수 없었을 것이라고 지적했으며, 추가적인 데이터 수집과 모델 재학습을 권고했다.

Instruction:
 저는 최근에 발병하기 전 1주 동안 총 업무 시간이 57시간 30분으로, 일상 업무 시간보다 30% 이상 증가했고, 업무량도 30% 이상 증가했습니다. 또한, 발병하기 전 4주간의 주당 평균 업무 시간은 48시간이었고, 12주간의 1주당 평균 업무 시간은 42시간 39분이었습니다. 제 업무는 육체적으로 매우 힘든 업무를 수행하는 것이 가중요인입니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?

Output:
 신청인의 뇌심혈관 질환은 업무상 질병으로 불인정 된다. 발병 전 12주간의 평균 업무 시간을 기준으로 30% 이상 증가하지 않았고, 교대제 업무를 수행하지도 않았습니다. 또한, 발병 전 6개월 이내에는 업무량, 환경변화, 업무상 부담 증가에 대한 업무상 요인이 없었습니다. 발병 전 1주일 동안의 평균 업무 시간은 26시간 24분으로, 일상 업무 시간보다 30% 이상 증가하지 않았습니다. 또한, 발병 전 12주간의 평균 업무 시간을 기준으로 30% 이상 증가하지 않았고, 업무부담 가중요인으로 는 교대제 업무 및 정신적 긴장이 높은 업무를 수행하지 않았습니다. 그러나 주말 근무는 하지 않았습니다. 급속도 또는 만성적인 과중한 업무를 수행한 사실은 인정되나, 신청인의 뇌심혈관 질환이 업무상 질병으로 인정되지 않는다.

직업환경의학 의사 의견:
 우선 48시간의 1.3배는 62시간임. 따라서 현재 57시간은 평소의 30% 증가는 아님. 다만 57시간이라는 과로와 부가 요인인 정신적 긴장이 있었음에도, 30% 증가에 대한 복합적 판단은 어려운 것으로 보임. 질문을 줄 때 법적 필수 요소인, 단기 과로 증가 (30% 증가)에 대한 모델 training이 필요할 것으로 보임. 이것이 인정되는 사례만 oversampling 하여 training이 필요할 것으로 보임. 현재를 ver1으로 부르고 단기과로 30%를 집중적으로 학습시킨 모델이 필요하다고 판단됨.

두번째 건은 업무량과 근로 시간이 30% 이상 증가한 사례였으나, 모델은 이를 감지하지 못하고 만성적인 과중한 업무를 인지했다. 이에 대하여 전문가는 단기 과로 상황에 대한 추가 학습이 필요하며, 이를 위해 해당 상황을 중점적으로 학습하는 새로운 모델이 필요하다고 제안했다.

Instruction:
 저는 돌발상황이나 급격한 업무 환경의 변화로 인해 뇌심혈관 질환이 발생할 수 있습니다. 제 업무량과 시간은 일상업무보다 30% 이상 증가하지 않았으며, 발병 전 4주간의 평균 업무 시간은 52시간이었고, 발병 전 12주간의 평균 업무 시간은 56시간 35분이었습니다. 또한, 제 업무는 유해한 작업환경에 노출되고 육체적으로도 강도가 높은 업무입니다. 이러한 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?

Output:
 신청인의 뇌심혈관 질환은 업무상 질병으로 불인정 된다. 발병 전 1주일 동안의 평균 업무 시간은 53시간 37분이었고, 발병 전 12주간의 평균 업무 시간은 55시간 30분이었습니다. 또한, 발병 전 4주간의 평균 업무 시간은 57시간 43분이었고, 발병 전 12주간의 평균 업무 시간은 53시간 23분이었습니다. 이외에도, 업무부담 가중요인으로는 교대제 업무수행과 정신적

긴장이 높은 업무, 유해한 작업환경이 있으며, 업무상 질병으로 인한 인정기준을 충족하지 못했으므로, 신청인의 뇌심혈관 질환은 업무상 질병으로 불인정 된다.

직업환경의학 의사 의견:
 이는 어떠한 기준이 업무부담으로 여길 지에 대한 데이터 학습이 아직 없는 모델이기 때문에 발생한 문제임. 업무부담이 구체적으로 무엇인지 판단하기 위한 모델 학습이 필요하고, 향후 chat 방식일 경우 업무부담에 대해서 구체적 정보를 요구하는 모델을 만들 필요가 있음. 이는 Template형식으로 가능하다고 보임.

세번째 건에서, 모델은 신청인의 뇌심혈관 질환을 업무상 질병으로 불인정하였으나, 일부 정보와 모델의 응답에 일관성이 없는 부분이 있었다. 전문가는 이러한 결과는 업무부담에 대한 충분한 학습이 이루어지지 않았기 때문이라고 지적하며, 업무부담에 대한 정의와 판단 기준을 학습해야 한다고 제안했다. 또한, 향후 챗봇 방식의 모델에서는 사용자로부터 구체적인 업무부담 정보를 받아들일 수 있는 템플릿 형식의 질문을 구성하는 것이 중요하다고 권고하였다.

실제 불인정 판정과의 불일치 건 중 2건에 대한 전문가의 의견은 다음과 같다.

Instruction:
 저의 일상 업무량이 30% 이상 증가하였고, 일상 업무 시간은 30% 이상 미증가하였습니다. 또한, 발병 전 1주간의 주당 평균 업무 시간은 54시간, 발병 전 12주간의 주당 평균 업무 시간은 51시간, 발병 전 4주간의 주당 평균 업무 시간은 54시간이었습니다. 그러나 업무부담 가중요인 갯수는 0개입니다. 이러한 정보를 바탕으로, 저의 뇌심혈관질환이 업무상 질병으로 판정받을 수 있을까요?

Output:
 신청인의 뇌심혈관 질환 상병은 업무상질병으로 인정된다. 업무상질병판정위원회는 신청인의 뇌심혈관 질환 상병이 업무상 질병으로 인정되기 위해서는 발병 전 24시간 이내에 업무와 관련된 뇌심혈관 질환의 증상이 발생하였어야 하고, 업무 시간이 12주간은 주당 평균 51시간, 12주간은 주당 평균 50시간, 4주간은 주당 평균 53시간이었다. 또한 업무부담 가중요인 갯수는 0개였다.

직업환경의학 의사 의견:
 상병 발생 1주전의 업무 시간은 52시간을 조금 넘긴 54시간이었지만, 4주간의 주당 평균 업무는 54시간 등으로 매우 높은 과로 수준은 아니었음. 여기에 가중 요인을 얼마나 높게 여겨냐에 따라 업무상질병 여부가 판단될 수 있다고 보임. 그러나, 가중요인의 심각성에 대한 정보를 학습하지 않은 모델로, 그 역할에 제한점이 있을 것으로 보임. 따라서 향후 가중 요인의 심각도를 학습하도록 발전시킬 필요가 있음.

<p>Instruction: 제 일상 업무량과 업무 시간이 각각 30% 이상 증가하지 않았으며, 발병 전 1주간의 주당 평균 업무 시간이 51시간, 발병 전 12주간의 주당 평균 업무 시간이 44시간 2분, 발병 전 4주간의 주당 평균 업무 시간이 46시간 45분이었습니다. 또한 업무부담 가중요인 갯수는 2개였습니다. 이러한 정보를 바탕으로 제 뇌심혈관질환이 업무상질병으로 판정받을 수 있는지 알려주실 수 있을까요?</p>
<p>Output: 신청인의 뇌심혈관 질환 상병은 업무상질병으로 인정된다.</p>
<p>직업환경의학 의사 의견: 1주간 평균 51시간이며, 12주 평균 44시간으로 만성 과로가 있다고 판단하기 어려움. 그러나 가중 요인이 3개 있는 경우로, 아직 모델이 가중 요인의 심각도를 학습하지 못하여 발생하는 부분으로 보임. 추후 가중 요인의 심각도를 학습시킬 필요가 있음.</p>

모델은 일부 불인정 사례에서, 업무부담 가중요인을 잘 인식하지 못하였다. 이에 대해 전문가는 만성과로가 없는 상태에서 가중요인이 여러 개인 경우 모델은 가중요인에 대한 심각성을 얼마나 고려할지 모르고 있는 상태라고 판단하였다. 가중요인의 영향력을 아직 학습하지 못한 상태로, 이를 보완하기 위해서는 추후 여러 상황에서의 업무부담 가중요인에 대한 개별적 학습이 필요하다고 판단하였다.

IV. 고 찰

본 연구에서는 학습 데이터에 따라 학습된 AI 모델의 응답 결과를 비교하고 성능을 평가하였다. 학습 이전 모델은 아무 데이터를 학습하지 않은 모델로 실험하였고, 전문지식 학습모델은 의료 및 법적 분야의 전문 지식을 학습하였다. 마지막으로 복합 학습 모델은 전문지식에 더하여 실제 뇌심혈관 질환의 업무상 판정 사례까지 학습하였다. 이 모델은 뇌심혈관 질환의 업무상 질병 판정에 대한 다양한 상황과 시각을 학습하였으므로, 학습 이전 모델 및 전문지식만 학습한 모델보다 훨씬 더 유연한 판단 능력을 기대하였다. 세 모델 모두 동일한 테스트 데이터로 응답을 출력하였을 때, 복합 학습 모델이 전반적으로 더 정확한 판단을 하였다.

이러한 결과는 데이터 주도 방식으로 학습되는 딥러닝 기반 모델에서 학습 데이터의 품질이 모델 성능에 매우 중요함을 강조한다(Jang et al., 2020). 특히 뇌심

혈관 질환과 같은 의학적 주제를 다룰 때는 다양한 변수와 패턴을 포함하는 정확한 학습 데이터가 필수적이다(Ding et al., 2019). 실제 뇌심혈관 질환 업무상 질병 판정에는 단순히 근로 시간, 업무 환경뿐만 아니라 근로자의 나이, 성별, 건강 이력, 생활 습관 등 여러 요소를 종합적으로 고려해야 하므로, 뇌심혈관 질환의 의학적 정의나 업무상 질병으로의 법적 인정기준만 아는 상태에서는 뇌심혈관 질환의 업무상 질병 판정은 어렵다고 판단된다. 실제로 세 모델 중 복합 학습 모델의 모든 성능 지표가 다른 모델보다 높았는데, 이는 복합 학습 모델이 불인정 사례를 실제 판정과 일치하게 판단한 사례 수가 다른 모델에 비해 압도적으로 많았기 때문이다. 따라서 복합 학습 모델이 다양한 실제 불인정 판정 사례를 학습하며 학습 이전 모델 및 전문지식 학습모델보다 불인정 사례를 비교적 더 정확하게 판단하였을 것으로 판단된다(Montesinos López et al., 2022).

그러나 복합 학습 모델도 발병 전의 근로 시간과 업무부담 등을 고려하여 업무상 질병을 판정하는 데 있어서 한계를 보였다. 특히, 단기 과로나 야간 근무와 같은 다양한 업무상 질병 인정 요소들을 충분히 판단하지 못했으며, 다양한 조건에서의 업무부담 가중 요인을 제대로 판단하지 못하였다. 이는 주로 학습 데이터의 수가 부족하고 만성 과로 요인에 치중된 데이터로 인한 과적합 현상 때문으로 판단된다. 과적합은 학습 모델이 실제 환경이나 테스트 데이터에 대해 일반화된 성능을 보이지 못하게 만드는 주요 문제이다. 이로 인해 모델은 단기 과로와 같이 복잡한 패턴을 갖는 사례에 대해 충분히 학습하지 못했으며, 이는 학습 데이터의 다양성 부족 및 특정 유형의 데이터가 불충분했을 가능성을 나타낸다. 실제로 업무상 질병을 판정할 때 단기 과로와 같은 요소가 중요한 기준임에도 불구하고, 학습 데이터에는 이러한 사례가 충분히 반영되지 않았던 것으로 보인다. 연구 결과는 모델이 다양한 업무상 질병 판정의 복잡성을 더 잘 이해하고 처리할 수 있도록 학습 데이터의 다양성을 증가시키고, 특히 단기 과로나 야간 근무와 같은 사례에 대한 데이터를 추가하는 것에 초점을 맞추어야 한다는 방향을 제시한다. 이러한 접근은 모델의 일반화 능력을 강화하고 실제 업무 환경에서 발생할 수 있는 다양한 상황에 보다 효과적으로 대응할 수 있게 할 것이다.

추가적인 모델의 발전 방향에 대하여 전문가는 업무부담에 대한 구체적 정보를 효과적으로 수집하고 분석

하는 능력을 갖춘 모델의 개발이 필요하며, 이를 위해 챗봇 형태의 모델이 사용자로부터 업무부담과 관련된 상세 정보를 요구하고 처리할 수 있어야 한다고 강조하였다. 이러한 접근은 사용자의 개별적인 상황을 더 잘 이해하고, 그에 따른 적절한 판단을 제공하는 데 중요할 것으로 판단되며, 챗봇의 개발을 위해서는 사용자가 쉽게 정보를 제공하고 상호작용할 수 있도록 친숙하고 직관적인 인터페이스를 제공하는 것이 필요하다고 지적하였다.

한편 모델을 위해 10-fold cross-validation을 적용하여 모델 성능을 보다 신뢰성 있게 평가하였다. 10-fold cross-validation을 통해 각 모델의 평균 accuracy를 산출한 결과, 첫 번째 모델은 61%, 두 번째 모델은 66.7%, 세 번째 모델은 87.4%의 정확도를 보였다. 초기 모델과 전문지식 학습 모델의 경우 교차검증 전후의 accuracy가 소폭 증가하거나 거의 차이가 없었으나, 복합 학습 모델의 경우 교차검증 전의 91% accuracy에서 교차검증 후 87.4%로 감소한다. 이는 교차검증 전의 복합 학습 모델이 overfitting 또는 underfitting의 영향을 받았음을 시사한다.

복합 학습 모델의 성능을 더욱 면밀히 분석하기 위해 예폭 수를 변경하여 추가적으로 fine-tuning을 수행하였다. 본 논문에서는 예폭 수 10의 결과를 기재하였으나, 추가적으로 복합 학습 모델에 대하여 예폭 수 5와 15로 fine-tuning을 진행하였다. 그 결과, 예폭 수 5와 15에 대한 각각의 training loss는 0.412, 0.218로 산출되었다. 또한 각 예폭 수에서의 모델 test 수행 후 정

확도 계산 결과, 85%, 89.5%로 계산되었다. 이는 예폭 수와 모델의 accuracy가 반드시 정비례하지 않음을 시사한다. 즉, 더 많은 예폭 수를 사용하여 모델을 학습시키는 것이 항상 더 나은 성능을 보장하지 않는다는 점을 확인할 수 있었으며, 적절한 예폭 수를 선택하는 것이 중요함을 강조한다.

이 연구의 주요 강점은 근로복지공단의 업무상질병판정서 데이터를 활용하여 개발된 새로운 AI 모델을 통해 뇌심혈관 질환과 업무 환경 간의 연관성을 깊이 있게 탐색하는 데 중요한 발전을 이루었다는 점이다. 특히, 이 모델은 업무상 질병, 예를 들어 만성 과로의 인정 여부를 정확하게 판단할 수 있는 뛰어난 성능을 보여, 업무 환경과 건강 상태 간의 복잡한 상호작용을 이해하는데 크게 기여할 수 있다. 이는 업무 환경이 개인의 건강에 미치는 영향을 파악하는 데 있어 기초적인 스크리닝 도구로 사용될 수 있으며, 특히 배경 지식이 부족한 사람들에게 유용하게 활용될 수 있다. 또한, 이 연구는 한국어에 특화된 Polyglot-ko 모델을 사용하여 한국어 데이터의 처리 성능을 향상시켰다. 이는 한국어로 된 업무상질병판정서 데이터의 특성을 더 잘 반영하고 이해할 수 있게 해, 모델의 정확도와 효율성을 극대화하는데 기여했다. 연구의 또 다른 주요 강점은 ‘Machine In The Loop’ 프로세스를 통해 데이터 전처리 및 변환 과정을 자동화한 점이다(Figure 2). ‘Human In The Loop’ 프로세스로 진행하면 연구자가 직접 전처리 및 변환 과정에 가담하기 때문에 소량의 데이터 전처리 과정에서 좀 더 연구자의 목적에 부합하게 데이터를 만

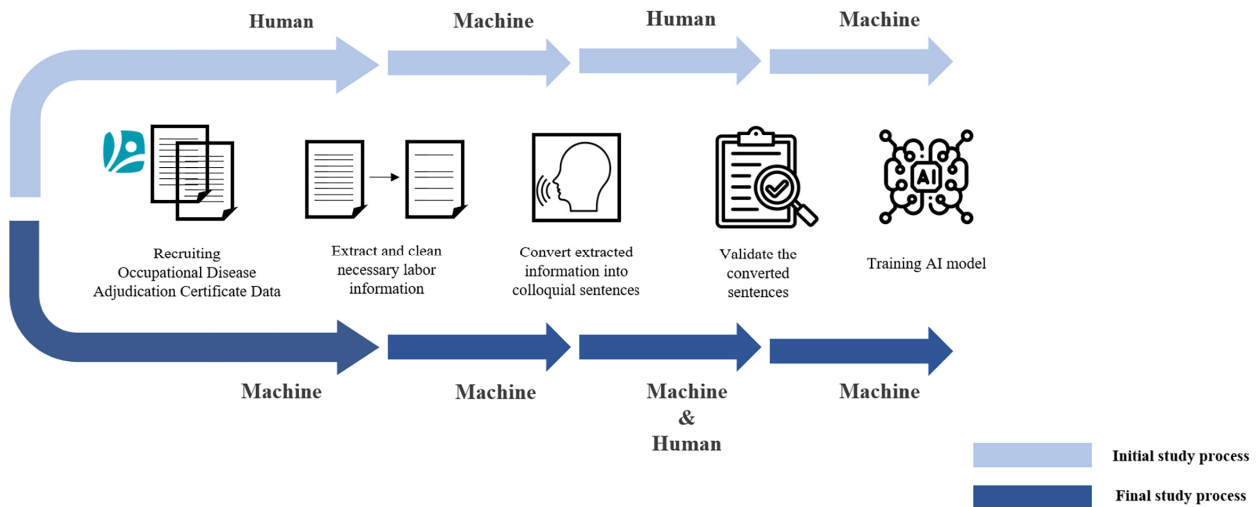


Figure 2. The process of machine-in-the-loop

들 수 있으나, 이는 추후에 대량의 데이터를 처리하기엔 처리 시간 및 효율성 측면에서 다소 무리가 있다. 따라서 데이터 전처리 과정에 AI 모델을 도입하여 데이터 처리 작업의 상당 부분을 자동화하여, 작업에 드는 시간을 줄이고 효율성을 극대화하였다. 또한 이러한 ‘Machine In The Loop’ 프로세스는 추후에 학습데이터 증량 및 학습모델을 챗봇에 응용하는 등 연구 규모 확장에 유연하게 대응할 수 있을 것으로 판단된다.

그러나 이 연구는 몇 가지 중요한 단점과 한계를 가지고 있다. 첫째, 모델은 주로 업무적 요인에 초점을 맞추었으나, 뇌심혈관 질환의 발병과 악화에는 근로자의 생활 습관, 건강 이력 등 다양한 요인이 영향을 미친다. 이러한 다양한 요인들이 모델에 포함되지 않아, 모델의 판정 범위와 정확성에 제한이 있었다. 둘째, 모델은 일부 특정 사례에 대한 성능이 낮았고, 질병 관리에 대한 조언을 제공하는 데에 한계가 있었다. 이는 모델이 이러한 복잡하고 다양한 패턴을 갖는 사례를 충분히 학습하지 못했기 때문일 수 있으며, 특히 이러한 유형의 사례는 더욱 복잡한 인과 관계와 상황적 요소들을 포함하고 있을 가능성이 있다. 셋째, 업무상질병판정서의 전처리를 위해 gpt-3.5-turbo 모델을 사용하였는데, GPT 모델의 특성상 동일한 입력에 대해서도 결과가 매번 동일하지 않을 수 있다. 이는 직업적 특성이 제대로 추출되지 않아서 불일치한 결과가 생성될 가능성을 내포한다. 이를 최소화하기 위해 업무상질병판정서 결론의 문체 및 단어 등을 텍스트를 동일한 포맷으로 맞추는 후 GPT 모델을 활용하여 전처리를 진행하였다. 그러나 이러한 노력에도 불구하고, 모델의 랜덤한 출력으로 인해 성능에 영향을 줄 수 있는 가능성을 완벽히 배제하지는 못하였고, 연구자가 일부 개입하여 검증을 수행하였다. 추후에 더 많은 데이터를 빠르게 전처리하기 위한 ‘Machine in the Loop’ 프로세스를 위해서는 추가적인 방법론적 개선이 필요하다. 마지막으로, 10-fold cross-validation을 통해 모델의 성능을 평가한 결과, 복합 학습 모델의 accuracy가 교차검증 전후로 감소하였다. 이는 교차검증 전 복합 학습 모델이 overfitting 되었음을 시사한다. 그러나 교차검증 후의 87.4% accuracy는 여전히 높은 수치로, 모델 개발 초기 단계로서는 유망한 결과로 판단된다. 그럼에도 불구하고 향후 연구에서는 더 다양한 데이터와 고도화된 모델링 기법을 적용하여 이러한 overfitting 문제를 해결할 필요성이 있다.

이러한 장단점을 고려할 때, 연구의 향후 방향은 다양한 유형의 업무상 질병 사례에 대한 데이터를 모델에 통합하고, 모델이 다양한 상황에 더욱 효과적으로 대응할 수 있도록 개선하는 데 초점을 맞추어야 할 것이다. 이를 통해 모델의 일반화 능력을 강화하고, 실제 업무 환경에서 발생할 수 있는 다양한 상황에 더욱 정밀하고 신뢰할 수 있는 업무상 질병 판정 도구로서의 역할을 수행할 수 있을 것이다.

V. 결 론

본 연구는 뇌심혈관 질환과 근로 시간, 업무 요인 간의 관계 및 업무상 질병 인정 여부를 파악할 수 있는 AI 모델을 개발하였다. 복합 학습 모델은 전문 지식과 실제 업무상 질병 판정 사례를 학습하여 가장 높은 성능을 보였지만, 다양한 사례에 대한 판정이 부족하다는 점과 추후 건강관리 방안 제공이 어렵다는 점 등의 한계점이 있었다. 따라서 모델의 성능 향상 후, 실제 근로자들이 사용할 수 있게 챗봇 형식의 webUI 개발이 필요하다고 판단되며, 이를 위해 근로자의 상세한 정보를 반영할 Template-based 방식의 모델로의 보완이 필요하다고 판단된다.

References

- Enforcement decree of the industrial accident compensation insurance act (2023. 7. 1.) article 37 paragraph 5 act no. 18928. 2023a
- Enforcement decree of the industrial accident compensation insurance act article 34 paragraph 3 and subparagraph 1 (a) of the appendix act no. 18928. 2023b
- Cherry N. Recent advances: Occupational disease. *Bmj* 1999;318(7195):1397-1399 (<https://doi.org/10.1136/bmj.318.7195.1397>)
- Deaton C, Froelicher ES, Wu LH, Ho C, Shishani K et al. The global burden of cardiovascular disease. *Eur J Cardiovasc Nurs* 2011;10 Suppl 2:S5-13 ([https://doi.org/10.1016/s1474-5151\(11\)00111-3](https://doi.org/10.1016/s1474-5151(11)00111-3))
- Ding J, Li X, Kang X, Gudivada VN. A case study of the augmentation and evaluation of training data for deep learning. *Journal of Data and Information Quality (JDIQ)* 2019;11(4):1-22 (<https://doi.org/10.1145/3317573>)
- Guthrie R & Westaway J. Emerging legal concerns with

- chronic diseases in the australian workplace: Pre-employment medicals, functional capacity evaluations, workers' compensation and disability discrimination. *J Law Med* 2009;16(5):803-821
- Jang SS, Hwang KH, Ha YG, editors. High quality training set collection using generative adversarial network. 2020 IEEE International Conference on Big Data and Smart Computing (BigComp); 2020: IEEE; (<https://doi.org/10.1109/BigComp48618.2020.00-27>)
- Joo HY, Oh H, Yang J. A survey on open source based large language models. *The Journal of Korea Institute of Information, Electronics, Communication Technology* 2023;16(4):193-202 (<https://doi.org/10.17661/jkiiect.2023.16.4.193>)
- Kamardeen I. Work stress related cardiovascular diseases among construction professionals. *Built Environment Project Asset Management* 2022; 12(2):223-242 (<https://doi.org/10.1108/BEPAM-06-2021-0081>)
- Kelly S & Romero A. Consequences of sharing invisible chronic health conditions at work: Implications for business communication education and training. *Business Communication Research Practice* 2019; 2(1):13-21 (<https://doi.org/10.22682/bcrp.2019.2.1.13>)
- Kim S & Jung Y. Effect of long working hours on cardiovascular disease in south korean workers: A longitudinal study. *Asia Pac J Public Health* 2021; 33(2-3):213-219 (<https://doi.org/10.1177/1010539520979927>)
- Kim UJ, Choi WJ, Kang SK, Lee W, Ham S et al. Standards for recognition and approval rate of occupational cerebro-cardiovascular diseases in korea. *Ann Occup Environ Med* 2022;34:e30 (<https://doi.org/10.35371/aoem.2022.34.e30>)
- Ko H, Yang K, Ryu M, Choi T, Yang S et al. A technical report for polyglot-ko: Open-source large-scale korean language models. arXiv preprint arXiv:02254 2023 (<https://doi.org/10.48550/arXiv.2306.02254>)
- Kosis. Mortality trend by cause of death 2022: Korean Statistical Information Service; 2022 [Available from: https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=1012]
- Lin RT, Chien LC, Kawachi I. Nonlinear associations between working hours and overwork-related cerebrovascular and cardiovascular diseases (ccvd). *Scientific Reports* 2018;8(1):9694 (<https://doi.org/10.1038/s41598-018-28141-2>)
- Lin XV, Mihaylov T, Artetxe M, Wang T, Chen S et al. Few-shot learning with multilingual language models. arXiv preprint arXiv:10668 2021 (<https://doi.org/10.48550/arXiv.2112.10668>)
- Montesinos López OA, Montesinos López A, Crossa J. Overfitting, model tuning, and evaluation of prediction performance. *Multivariate statistical machine learning methods for genomic prediction*: Springer; 2022. p. 109-139 (https://doi.org/10.1007/978-3-030-89010-0_4)
- Mora CD & Bornstein RA. Evaluation of cerebrovascular disease. *Neuropsychology*: Springer; 1998. p. 171-186 (https://doi.org/10.1007/978-1-4899-1950-2_8)
- Park S, Nam J, Lee JK, Oh SS, Kang HT et al. Association between night work and cardiovascular diseases: Analysis of the 3rd korean working conditions survey. *Ann Occup Environ Med* 2015;27:15 (<https://doi.org/10.1186/s40557-015-0064-1>)
- Picciotto D. [occupational cardiovascular diseases and phlebopathies]. *G Ital Med Lav Ergon* 2010;32(4 Suppl):160-161
- Roth GA, Abate D, Abate KH, Abay SM, Abbafati C et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: A systematic analysis for the global burden of disease study 2017. *The lancet* 2018;392(10159):1736-1788 (https://doi.org/10.1007/978-1-4899-1950-2_8)
- Thrift AG, Cadilhac DA, Thayabaranathan T, Howard G, Howard VJ et al. Global stroke statistics. *Int J Stroke* 2014;9(1):6-18 (<https://doi.org/10.1111/ijss.12245>)
- Trivedi UB & Mishra P, editors. Fine tuning (diagnosis) of machine learning algorithm (model) for optimization. 2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT); 2022: IEEE; (<https://doi.org/10.1109/ICAECT54875.2022.9807982>)
- Vineis P, Demetriou CA, Probst-Hensch N. Long-term effects of air pollution: An exposome meet-in-the-middle approach. *Int J Public Health* 2020; 65(2): 125-127 (<https://doi.org/10.1007/s00038-019-01329-7>)
- Who. Cardiovascular diseases (cvds): World Health Organization; 2017 [Available from: [http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))]
- Wu Y, Xiong Y, Wang P, Liu R, Jia X et al. Risk factors of cardiovascular and cerebrovascular diseases in

young and middle-aged adults: A meta-analysis. *Medicine (Baltimore)* 2022;101(48):e32082 (<https://doi.org/10.1097/md.00000000000032082>)

Zhuang YT, Wu F, Chen C, Pan YH. Challenges and opportunities: From big data to knowledge in ai 2.0. *Frontiers of Information Technology Electronic Engineering* 2017;18:3-14 (<https://doi.org/10.1631/>

FITEE.1601883)

<저자정보>

오주연(연구원), 유기봉(교수), 진익훈(교수), 윤병윤(중개 연구조교수), 심주호(연구조교수), 박희주(연구원), 이종민(연구원), 이지안(연구원), 윤진하(교수)