



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

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

연세대학교 대학원

보건학과

오주연

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

지도교수 윤진하

이 논문을 오주연 학위논문으로 제출함

2023년 12월

연세대학교 대학원

보건학과

오주연

## 오주연의 석사 학위논문을 인준함

심사위원 윤진하 인

심사위원 유기봉 인

심사위원 진익훈 인

연세대학교 대학원

2023 년 12 월

## <차례>

국문요약 .....	
I. 서론 .....	1
1. 연구의 배경 및 필요성 .....	1
1.1 뇌심혈관 질환과 업무상 질병과의 관계 .....	1
1.2 근로자들이 자신의 질환과 업무 관련성을 인식하기 어려운 배경 .....	3
1.3 인공지능의 활용 가능성 및 기계 학습의 필요성 .....	4
2. 연구 목적 .....	6
II. 연구 방법 .....	7
1. 사용 데이터 .....	7
2. 기반 AI 모델 및 Fine-tuning .....	13
2.1 사용한 기반 AI 모델 .....	13
2.2 모델 Fine-tuning .....	14
3. 학습 데이터에 따른 모델 결과 비교 .....	15
4. 모델의 성능 평가 및 전문가 의견과의 일치도 판정 .....	17
5. 사용 프로그램 .....	18
6. 연구 윤리 .....	19
III. 연구 결과 .....	20
1. 모델 결과 비교 .....	20
2. 직업환경의학 전문가의 의견 .....	23
IV. 고찰 .....	29
V. 결론 .....	35
참고문헌 .....	36
부록 .....	40
ABSTRACT .....	45

## 표 차례

표 1. 뇌심혈관 질환의 종류 및 의학적 정의 .....	8
표 2. 뇌심혈관 질환의 업무상 질병 인정기준 .....	9
표 3. 모델 학습에 이용한 업무상질병판정서 예시 .....	12
표 4. 각 언어모델의 특징 .....	14
표 5. 실제 업무상 질병 판정과 비교한 세 모델의 응답 결과 비교 .....	20
표 6. 세 모델의 성능 평가지표 .....	22
표 7. 복합 학습 모델에서 실제 판정과의 응답 불일치 사례에 대한 전문가 의견1.....	23
표 8. 복합 학습 모델에서 실제 판정과의 응답 불일치 사례에 대한 전문가 의견2.....	25
표 9. 복합 학습 모델에서 실제 판정과의 응답 불일치 사례에 대한 전문가 의견3.....	27
부록 표 1. 모델 테스트 시 사용한 테스트 데이터 .....	37

## 그림 차례

그림 1. gpt-3.5-turbo 모델을 사용한 학습데이터 전처리 과정.....	11
그림 2. 학습 데이터 전처리 및 각 모델의 학습 데이터 .....	16
그림 3. 학습 이전 모델의 업무상 질병 판정 불가 사례 .....	21
그림 4. 전문지식 학습모델의 업무상 질병 판정 불가 사례 .....	21
그림 5. Machine In The Loop 프로세스.....	33

## 국문요약

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

연구 배경: 뇌심혈관 질환은 전세계적으로 주요한 건강 문제로 다양한 원인에 의해 발생할 수 있다. 그 중 과로 및 과도한 업무 부담으로 인해 발병 및 악화할 수 있으며, 이는 업무상 질병으로 인정될 수 있다. 그러나 뇌심혈관 질환 및 업무상 질병 판정에 대한 지식이 부족한 일반 근로자는 본인의 질환과 업무 환경과의 관계를 인식하기 어렵다. 이에 따라 일반 근로자들도 쉽게 뇌심혈관 질환과 업무상 질병과의 관계를 이해하고 판단할 수 있게 도와주는 AI 모델을 개발하고자 한다.

연구 방법: 모델 학습에 사용한 데이터는 의학적, 법적 전문 지식과 실제 근로복지공단의 2021년도 업무상질병판정서 데이터이다. 학습에 사용한 모델은 대규모 한국어 데이터를 학습하여 한국어 성능이 좋은 Polyglot-ko-5.8B 모델을 사용하였다. 학습 이전 모델과 전문 지식만 학습한 모델, 그에 더하여 실제 업무상 질병 판정 사례까지 복합적으로 학습한 모델의 응답을 근로복지공단 실제 판정 사례와 비교한다. 학습된 모델의 성능은 정확도, 정밀도, 민감도, F1-score로 확인하였다. 마지막으로, 모델 응답을 읽고 직업환경의학 의사의 의견 및 보완점을 통해 전문가의 의견을 확보하였다.

연구 결과: 세 모델의 응답 비교 결과, 전문지식과 다양한 업무상 질병 판정 사례를 복합적으로 학습한 모델이 다른 두 모델에 비해 성능이 제일 높았다. 정확도는 학습 이전 모델, 전문지식 학습모델, 복합 학습 모델 순서대로 0.60, 0.65, 0.91이었고, 정밀도는 순서대로 0.71, 0.86, 0.84였다. 민감도는 0.54, 0.54, 0.95였고, 마지막으로 F1-score는 순서대로 0.64, 0.67, 0.93이었다. 그러나 복합 학습 모델도 몇몇 일관되지 못하거나 잘못된 응답을 출력하였다. 이에 대해 전문가는 단기 과로 및 야간근로 등 다양한 사례에 대한 학습이 추가적으로 더 필요하고, 추후에 chatbot을 위한 template 형식의 모델이 필요하다고 제안하였다.

고찰: 세 가지 모델(학습 이전 모델, 전문지식 학습모델, 복합 학습 모델)을 사용하여 뇌심혈관 질환의 업무상 질병 판정에 대한 응답을 비교하였다. 결과적으로, 복합 학습 모델이 가장 정확한 판단을 보였다. 본 연구는 근로복지공단의 업무상질병판정서라는 실제 사례를 fine-tuning 시킨, 기존에 없던 새로운 AI모델을 개발한 것에서 의의가 있다. 자신의 뇌심혈관질환과 업무적 요인 간의 연관성에 대한 배경지식이 없는 근로자와, 자신의 업무적 정보에 대해서 업무상질병 인정을 받을 수 있는지 궁금한 근로자에게 기초적인 스크리닝 도구가 될 수 있다. 그리고 한국어 친화적인 polyglot-ko 모델을 기반 모델로 사용했기 때문에 한국인 근로자에게 유용하게 사용될 수 있다. 그러나 업무적 요인 외에 다양한 요인을 학습하지 못했기 때문에 다양한 사례에 대한 판단이 어렵다. 또한 아직은 업무상질병 인정/불인정만을 판단할 수 있고, 추후에 어떤 건강관리 방안이나 어떤 절차가 필요한지 등 해결책을 제시하기엔 어려움이 있다. 추후에 업무상질병으로 인정된 사람에게 건강관리 방안이나 추후에 어떤 건강행동을 취해야 하는지 등 어떤 중재 요인을 줄 수 있는지에 대한 추가적인 연구나 개발단계가 필요하다.

결론: 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 판단이 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.

---

핵심되는 말: 뇌심혈관 질환, 업무상 질병, 인공 지능, 대규모 언어 모델

# I. 서론

## 1. 연구의 배경 및 필요성

### 1.1 뇌심혈관 질환과 업무상 질병과의 관계

뇌심혈관 질환은 심혈관질환과 뇌혈관질환을 아울러 이르는 용어로, 혈관과 관련된 일부 병리학적 상태로 인해 발생하는 심장 및 뇌 기능의 저하로 광범위하게 정의된다.<sup>1</sup> 주로 관상동맥 심장 질환, 출혈성 뇌혈관질환과 허혈성 뇌혈관질환을 포함하는 뇌졸중이 해당한다. 관상동맥 심장질환은 심장 근육에 필요한 산소를 전달하기에 국소 혈액 공급이 부족한 관상 동맥 장애로, 급성 심근경색, 협심증, 기타 허혈성 심장 질환을 포함한다.<sup>2-4</sup> 출혈성 뇌혈관질환은 뇌 조직에 직접적인 출혈이 있을 때 발생하며, 고혈압, 동맥류 파열, 혈관 기형 또는 항응고제의 합병증이 원인이 될 수 있다. 허혈성 뇌혈관질환은 혈전이 뇌로 가는 동맥을 막아 혈류가 멈출 때 발생하며, 이로 인해 뇌의 일부가 손상되는 질환을 의미한다.<sup>5</sup>

이러한 뇌심혈관 질환은 전 세계적으로 주요 사망 원인으로,<sup>6-8</sup> 세계보건기구(WHO)에 따르면 뇌심혈관 질환은 2019년 전세계 사망자의 32%를 차지하였다.<sup>9</sup> 전세계의 추세와 유사하게 한국에서도 뇌심혈관 질환이 최근 10년 동안 주요 사망원인으로 꼽혔으며, 통계청 2021년 사망 원인통계에 따르면 심장질환으로 인한 사망이 전체 사망원인의 2위를 차지하였다.<sup>10</sup> 이러한 질환의 발병과 진행에는 개인의 흡연 및 음주와 같은 생활 습관, 비만, 유전적 요인, 그리고 대기오염 및 업무적 환경과 같은 환경적 요인 등 다양한 요인들이 관련되어 있다.<sup>11-13</sup>

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. 이러한 연관성은 직업병으로 분류될 수 있는데, 직업병이란 특정한 업무 환경이나 작업 특성상 노출되는 물질이나 조건에 의해 발생하는 질

병을 의미한다.<sup>14</sup> 근로자가 특정 작업에 종사하면서 그 작업과 관련된 유해한 물질이나 조건, 또는 업무로 인한 스트레스 등에 노출되어 발병하게 된다.<sup>15</sup> 직업병은 특정 업무와 직접적인 연관성을 가지며, 업무 자체가 질병의 원인으로 작용하거나 이미 존재하는 질병을 더욱 악화시키는 경우에 해당된다.<sup>16</sup>

뇌심혈관 질환은 고도의 스트레스, 과도한 업무 부담, 장시간의 근무, 또는 특정 환경 노출로 인한 고혈압의 악화 등이 발병 또는 악화의 주요 원인이 될 수 있다. 국내에서도 고용노동부에서 고시한 뇌심혈관 질환의 업무상 질병 인정에 필요한 사항<sup>17</sup>에 따르면, 근로자가 만성적인 과중한 업무로 인해 뇌심혈관 질환이 발병된 경우, 업무상 질병으로 인정한다. 근로자의 과로 여부 판단 시 근로 시간 및 업무부담 가중요인을 고려하게 되는데, 발병 전 12주 동안 1주 평균 근로 시간이 60시간을 초과하거나 발병 전 4주 동안 1주 평균 64시간을 초과하였을 때는 업무와 질병과의 관련성이 높다고 판단하고 있다. 또한 발병 전 12주 동안 1주 평균 근로 시간이 52시간을 초과하고 교대제 업무, 휴일이 부족한 업무, 정신적 긴장이 큰 업무 등 업무부담 가중요인이 포함되어 있으면 업무와 질병과의 관련성이 높다고 판단하고 있다.

국내외로도 긴 근로 시간 및 업무부담 가중요인과 뇌심혈관질환과의 연관성을 확인하고자 하였다. Ro-Ting Lin et al<sup>18</sup>이 진행한 연구에서 근무시간이 길어질수록 뇌심혈관 질환의 상대 위험비가 증가하였고, 건설노동자를 대상으로 한 Imriyas Kamardeen et al<sup>19</sup>의 연구에서 육체적 부담이 큰 업무와 정신적 스트레스가 클수록 뇌심혈관 질환의 발생률이 높아짐을 확인하였다. 국내 연구 또한 장시간 근로 및 업무부담 가중요인이 뇌심혈관 질환에 유의한 연관성을 확인한 연구가 다수 진행되었다.<sup>20,21</sup>

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.

## 1.2 근로자들이 자신의 질환과 업무 관련성을 인식하기 어려운 배경

앞서 언급하였듯 뇌심혈관 질환은 다양한 원인에 기인하기 때문에, 뇌심혈관질환의 발병을 이해하려면 깊은 의학적 지식과 전문적인 판단이 필요하다.<sup>22</sup> 하지만 대다수의 근로자들에게는 이러한 전문적 지식이 부족하여, 자신의 건강 문제와 작업 환경 간의 연관성을 정확하게 인식하기가 쉽지 않다.

또한, 일반적으로 사람들은 일상적인 생활 습관, 유전적 요인, 환경적 요인 등 여러 가지 원인 중에서 직장 환경이 건강에 미치는 영향을 쉽게 간과하는 경향이 있다. 그뿐만 아니라, 일상적인 스트레스나 업무 중 발생하는 여러 가지 사건들을 단순히 일시적인 사건으로 간주하며 그것이 장기적으로 건강에 미치는 영향을 인식하지 못하는 경우가 많다. 이러한 관점에서 볼 때, 많은 근로자는 자신의 직무와 관련된 스트레스나 부정적인 작업 환경이 건강에 미치는 영향을 충분히 인식하지 못하고 있을 가능성이 크다.

더불어, 직업병으로 인정받기 위한 과정은 매우 복잡하다. 특히, 근로자 본인이 그 원인과 결과 사이의 명확한 연관성을 증명하는 데 필요한 다양한 의학적 검사와 증거를 제시해야 하는데,<sup>23</sup> 많은 근로자는 이러한 복잡한 절차와 그에 따른 높은 비용, 그리고 실패할 경우를 대비한 경제적 위험 때문에 직업병으로의 인정을 추구하지 않게 된다.

마지막으로, 직장 내에서 자신의 건강 문제를 직업병으로 주장하는 것이 부정적인 인식을 초래할 수 있다는 우려도 존재한다.<sup>24</sup> 특히, 일부 근로자들은 자신의 직업병 주장이 회사나 동료들로부터의 부정적인 평가나 대우를 받을 수 있다고 생각하기 때문에,<sup>25</sup> 이러한 문제를 내부적으로 해결하려는 경향이 있다.

따라서, 직업병과 관련된 지식의 부족, 복잡한 인정 절차, 그리고 직장 내에서의 부정적인 인식 등 여러 요인이 근로자들이 자신의 질병과 업무 환경 간의 연관성을 제대로 인식하고 이를 주장하는 것을 어렵게 만들고 있다.

### 1.3 인공 지능의 활용 가능성 및 기계 학습의 필요성

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 최근 발전은 이러한 문제 해결에 획기적인 방안을 제공할 수 있다는 가능성을 내비치고 있다.<sup>26</sup> 머신러닝과 딥러닝 알고리즘은 거대하고 다양한 데이터 세트에서 유의미한 패턴을 찾아내는 능력을 보유하고 있다.<sup>27</sup> 이를 활용하면, 개인의 건강 데이터와 작업 환경 데이터를 종합적으로 분석하여, 그 사이의 복잡한 연관관계를 파악하고, 근로자 개개인의 뇌심혈관 질환에 대한 업무상 질병을 판정해 주는 모델을 구축할 수 있다.

그러나, AI 기술의 빠른 발전과 뛰어난 성능에도 불구하고, 직면할 주요한 한계 중 하나는 특정 도메인 또는 특화된 분야의 지식을 내재화하지 못하는 것이다. 현대의 AI 모델, 특히 범용 AI 모델은 다양한 데이터 세트에서 훈련되기 때문에, 그들의 지식 베이스는 광범위하게 분산되어 있다. 이는 업무상 질병 판정과 같은 매우 특정한 분야에서는 그들의 성능을 제한할 수 있다. 업무상 질병의 판정은 전문성과 특정한 지식이 필요하다. 이는 단순한 의학적 진단을 넘어서, 근로자의 작업 환경, 장시간 노출된 환경의 특성, 그리고 뇌심혈관 질환의 발병에 영향을 미칠 수 있는 여러 요인을 종합적으로 고려해야 한다. 따라서, 일반적으로 훈련된 AI 모델만으로는 이러한 복잡한 판정을 제대로 수행하기 어렵다.

이 문제를 해결하기 위해서는 AI 모델을 해당 분야에 맞게 조정하는 과정이 필요하다.<sup>28</sup> 이 데이터를 활용하여 AI 모델에게 업무상 질병의 판정에 필요한 전문적인 지식과 법적 인정기준을 학습시키면 업무상 질병 판정의 정확성을 크게 향상할 것으로 예상된다.

더불어, AI의 판단 기준이 실제 의료 전문가나 직업환경의학 전문가의 판단과 얼마나 일치하는지에 대한 신뢰성 확보도 중요하다고 판단된다. 여러 AI 모델이 실제 환경에서 얼마나 효과적으로 작동하는지, 그 결과가 전문적인 의견과 얼마나 일치하는지에 대한 철저한 검증과 평가가 필요하다. 이러한 검증 과정을 통해, AI 모델이 근로자 개개인이 본인의 업무환경 및 업무 요인에 대해 대한 자가 판단 능력과 건강관리의 인식 증진에

얼마나 효과적으로 활용될 수 있는지를 탐구하고자 한다.

## 2. 연구목적

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. 연구방법

### 1. 사용 데이터

#### 1.1 학습에 사용된 데이터

AI 모델 개발을 위해 학습에 사용한 데이터는 3 가지이다. 우선 뇌심혈관 질환의 종류와 의학적 정의, 뇌심혈관 질환의 업무상 질병 인정기준을 명시한 산업재해보상보험법 및 고용노동부고시 제 2022 - 40 호와 같이 의학적 전문 지식과 법적 전문 지식을 학습시키고자 하였다. 그리고 실제 근로자들의 뇌심혈관 질환 신청 질병 내용 및 이에 대한 업무상질병판정위원회의 업무상 질병 인정 판단 내용을 확인할 수 있는 근로복지공단의 2021 년도 업무상질병판정서를 활용하여 실제 사례에 대하여 어떻게 인정 여부를 판단하는지에 대하여 학습시키고자 하였다.

‘뇌심혈관 질환’ 상병의 정의와 종류, 그리고 그에 대한 의학적 정의는 다음과 같다.<sup>9,29</sup>

**표 1 뇌심혈관 질환의 종류 및 의학적 정의**

<b>뇌심혈관 질환의 종류는?</b>	출혈성 뇌혈관질환(뇌출혈) 중 뇌실질내출혈, 뇌지주막하출혈, 허혈성 뇌혈관 질환(뇌경색), 심장질환 중 심근경색, 해리성 대동맥류가 있다.
<b>출혈성뇌혈관질환(뇌출혈)</b>	출혈성뇌혈관질환(뇌출혈)이란, 뇌혈관이 터져 출혈이 생기면서 해당 부위의 뇌 손상이 나타나는 뇌혈관 장애를 말한다.

<b>뇌실질내출혈</b>	<p>“뇌실질내출혈”은 뇌 속의 동맥이 파열되면서 출혈된 피가 뇌실질 속으로 들어가 뇌조직을 파괴하거나 출혈 부위의 정상적인 뇌조직을 압박하여 심하면 사망에 이르게 하는 질병입니다. 출혈의 기전은 고혈압에 의한 혈관 벽의 변화에 의해서 또는 고혈압에 자주 동반되는 미세경색에 의해 혈관주위의 지지조직이 약해져서 혈관 벽이 쉽게 파괴되어 출혈을 일으키거나 뇌실질 내의 작은 혈관의 분지에 있는 미세 동맥류가 파괴되어 발생하는 등 80% 이상이 고혈압으로 인한 것이며, 종양, 혈액질환, 아밀로이드혈관병변, 혈관종 등이 실질내출혈의 원인이 된다.</p>
<b>뇌지주막하출혈</b>	<p>“뇌지주막하출혈”은 지주막하강, 즉 뇌의 표면의 수막강 내에 원발성으로 혈이 일어나는 것인데, 동맥류의 파열이 가장 흔한 원인이고 그 밖에도 뇌내출혈, 동정맥기형의 파괴, 혈액질환, 혈관염, 뇌막염, 종양, 정맥성 뇌혈관질환, 외상 등에 의해 지주막하출혈이 발생한다.</p>
<b>허혈성 뇌혈관질환 (뇌경색)</b>	<p>“뇌경색”이란 본래 뇌에 혈액을 공급하는 동맥계의 일부가 동맥경화나 색전증에 의해 혈관 내강이 좁아지거나 폐색됨으로써 그 관류역의 뇌조직이 괴사에 빠지는 것을 말한다.</p>
<b>심근경색</b>	<p>“심근경색”이란 관상동맥의 폐색에 의해서 넓은 범위에 걸친 심근의 허혈성 괴사가 생긴 경우를 말한다.</p>
<b>해리성 대동맥류</b>	<p>“해리성대동맥류”란 대동맥 내막이 파열됨으로 인하여 대동맥의 진강으로부터 높은 압력의 혈액이 빠져나와서 중막의 내층과 외층을 급속히 해리시키는 질환으로 대동맥 해리 또는 대동맥 박리라고도 한다. 동맥에서 발생하는 질병 중 가장 무서운 병이라 할 수 있으며, 동맥경화가 가장 흔한 원인이고 고혈압이 중요한 선행요인(90%)으로 생각되고 있다.</p>

뇌심혈관 질환의 업무상 질병 인정기준은, 산업재해보상보험법 I(업무상 재해)의 제 34 조 제 3 항 및 제 37 조 제 5 항<sup>30,31</sup>에 따라 뇌심혈관 질환의 업무상 질병 인정 기준을 고시한 고용노동부고시 제 2022 - 40 호 『뇌혈관 질병 또는 심장 질병 및 근골격계 질병의 업무상 질병 인정 여부 결정에 필요한 사항』<sup>17</sup>을 참고하였다. 이에 따라 AI 모델에 학습시킨 내용은 다음과 같다.

**표 2 뇌심혈관 질환의 업무상 질병 인정기준**

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 <ul style="list-style-type: none"> <li>• 교대제 업무</li> <li>• 휴일이 부족한 업무</li> <li>• 유해한 작업환경</li> <li>• 육체적 강도가 높은 업무</li> <li>• 시차가 큰 출장이 잦은 업무</li> <li>• 정신적 긴장이 높은 업무</li> </ul>
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.

마지막으로, 실제 뇌심혈관 질환을 업무상 질병으로 인정받기 위한 판정 신청 사례와 그에 대한 업무상질병판정위원회의 판단 및 결론 자료는 근로복지공단 업무상질병판정서 홈페이지에서 확인할 수 있다. 업무상질병판정서는 발생한 뇌심혈관 질환의 업무상 질병 인정 판정을 신청한 근로자의 나이와 성별, 건강검진 이력, 흡연 및 음주와 같은 생활 습관에 대한 정보도 포함되어 있는 인정사실로, 질병의 발생 배경을 쉽게 파악할 수 있다. 또한 근무시간, 업무 부담 가중 요인, 발병 전의 특이 사항이나 돌발 상황 등의 정보가 자세히 서술되어 있어 근로자가 질병을 얻을 위험 요인에 노출되었는지를 판단하는 데 도움을 준다. 이와 같은 정보를 바탕으로 업무상질병판정위원회의 전문가들이 제시된 정보와 증거를 기반으로 업무와 질병 간의 연관성을 평가한 인정여부 판단 및 결론까지 확인할 수 있다.

## 1.2 데이터 전처리 과정

AI 모델을 학습할 때는, 앞서 언급하였던 뇌심혈관 질환의 업무상 질병 인정기준에 따라 인정 여부를 판단할 수 있게 하기 위해 전처리 과정을 거친 후 학습에 이용하였다. 근로자의 정보가 담긴 인정사실은 openAI의 gpt-3.5-turbo 모델을 이용하여 돌발상황 또는 급격한 업무환경의 변화, 발병 전 1주간의 업무시간 및 평소보다 업무시간, 업무량 30% 증가 여부 등 단기간의 업무상 부담, 발병 전 4주/12주 동안의 1주당 평균 근로 시간 등 만성적으로 과중한 업무 부담 여부, 추가적인 업무부담 가중요인만으로 정리하였다. 또한 근로자의 업무적 정보만으로 인정 여부 판단을 내린 사례로 간추려서 위원회의 판단 및 결론을 output으로 만들었다. 마지막으로, 전처리된 인정사실을 학습모델에 사람이 직접 물어보듯이 학습시키기 위해 gpt-3.5-turbo 모델을 이용하여 구어체로 문장을 변경하였다. 변경된 문장의 예시는 다음 그림과 같다.

인정사실 원본

- 근무형태 : 상용직, 고정 주간근무, 주5일  
 - 근무시간 : 08:00~17:00  
 - 업무내용 : 선박용 발판 설치 및 철거(해체) 작업

- 발병 당일인 2020. 7. 21.(화) 정상 출근하여 선박 내 지하 4층에서 평소 업무인 발판 해체 작업 수행 중 15:30경 어지럼증 최초 호소한 이후 16:46경 쓰러져 의료 기관 내원하였으며 발병 전날인 2020. 7. 20.(월)은 휴무로 근무하지 않은 것으로 확인되어 돌발 상황이나 급격한 업무환경 변화는 없었던 것으로 확인된다.

○ 만성적으로 과중한 업무(발병 전 3개월 이상)  
 - 발병 전 3개월 이상 연속적으로 과중한 부담 여부를 확인한 결과 발병 전 4주 동안 업무시간이 1주 평균 33시간 18분(휴일 12일)이고, 발병 전 12주 동안 업무 시간은 1주 평균 29시간 11분(월 평균 휴일 14일)으로 특별한 업무환경의 변화 및 특이사항은 없었던 것으로 확인된다.

○ (개인요인) 신청 상병의 발병과 관련된 개인요인으로 신청인 진술에서 다음과 같은 내용이 확인된다.  
 - 흡연력 : 약 30년, 3개비/일  
 - 음주력 : 주 1회, 소주 1병/회



업무정보만으로 간추려진 인정사실

○ 단기간 동안의 업무상 부담(발병 전 1주일 이내)  
 - 발병 전 1주간 업무시간은 35시간으로 발병 전 12주간의 주당 평균 업무시간인 28시간 37분과 비교하여 일상 업무량 및 일상 업무시간 30% 미증가 되었으며 업무환경의 변화 및 특이사항은 없었음이 확인된다.

○ 만성적으로 과중한 업무(발병 전 3개월 이상)  
 - 발병 전 4주 동안 업무시간이 1주 평균 33시간 18분(휴일 12일)이고, 발병 전 12주 동안 업무시간은 1주 평균 29시간 11분으로 특별한 업무환경의 변화 및 특이사항은 없었던 것으로 확인된다.

○ 업무부담 가중요인  
 - 업무부담 가중요인 없음

구어체 문장으로 변경된 instruction

저는 발병 전 1주간의 업무시간은 35시간이었고, 이는 발병 전 12주간의 주당 평균 업무시간인 28시간 37분과 비교하여 30% 미증가되었습니다. 또한, 업무환경의 변화나 특이사항은 없었습니다.

발병 전 4주 동안의 업무시간은 1주 평균 33시간 18분이었고, 발병 전 12주 동안의 업무시간은 1주 평균 29시간 11분이었습니다. 이 기간 동안에도 특별한 업무환경의 변화나 특이사항은 없었습니다. 또한, 저는 업무부담 가중요인이 없었습니다.

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



그림 1 gpt-3.5-turbo 모델을 사용한 학습데이터 전처리 과정

따라서, 최종적으로 모델에 학습시킬 업무상질병판정서의 구조는 구어체로 근로 시간, 업무부담 가중요인만 요약된 인정사실(Instruction)과 그에 대한 위원회의 판단 및 결론(Output)으로 이루어졌다. 모델 학습에 사용된 업무상질병판정서의 예시는 다음과 같다.

**표 3 모델 학습에 이용한 업무상질병판정서 예시**

<p><b>Instruction:</b></p> <p>저는 발병 전 1주간의 업무시간은 35시간이었고, 이는 발병 전 12주간의 주당 평균 업무시간인 28시간 37분과 비교하여 30% 미증가되었습니다. 또한, 업무환경의 변화나 특이사항은 없었습니다.</p> <p>발병 전 4주 동안의 업무시간은 1주 평균 33시간 18분이었고, 발병 전 12주 동안의 업무시간은 1주 평균 29시간 11분이었습니다. 이 기간에도 특별한 업무환경의 변화나 특이 사항은 없었습니다.</p> <p>또한, 저는 업무부담 가중요인이 없었습니다.</p> <p>이러한 정보를 바탕으로 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
<p><b>Output:</b></p> <p>발병 당일 돌발 상황이나 급격한 업무환경 변화는 없었고 발병 직전 1주간 업무시간은 35시간 21분으로 일상 업무시간보다 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.</p>

이러한 방식으로 정리된 업무상질병판정서 1,000개 중 800개를 학습 데이터, 200개를 모델 학습이 제대로 되었는지 실험해 보기 위한 테스트 데이터로 사용하였다.

## 2. 기반 AI 모델 및 Fine-tuning

### 2.1 사용한 기반 AI 모델

최근 openAI 사의 GPT, META 사의 LLaMA 등 다양한 대규모 언어 모델이 출시되고 있다. 이러한 거대 파라미터 모델에 대규모 데이터를 학습한 언어모델은 자연스러운 문장을 생성하는 등 뛰어난 성능을 보이고 있으나, 이러한 대규모의 모델을 학습시키고 연구 및 상업적으로 이용하기 위해서는 초대용량의 메모리 및 컴퓨팅 자원이 필요하고, 정확한 모델의 구조 및 모델 가중치를 공개하지 않아 더욱 연구 활용에 어려움이 있다.<sup>32</sup> 더불어 현재 대부분의 대규모 모델의 학습데이터는 영어에 치중되어 있어, 비영어권 언어에 대해서는 높은 성능을 보이지 못하고 있다. 이에 따라 영어 이외에도 비영어권 언어를 학습한 여러 다국적 언어 모델이 공개되고 있으나,<sup>33</sup> 이러한 모델의 학습데이터 역시 영어 데이터가 높은 비중을 차지하고 있어, 한국어 자료에 대해서는 여전히 성능이 만족스럽지 못하다.<sup>34</sup> 따라서 이 연구에서는 비영리 AI 연구단체인 EleutherAI 에서 만든 대규모 한국어 모델인 Polyglot-ko-5.8B 모델을 연구에 사용하였다.

Polyglot-ko 모델은 Polyglot 모델을 기반 모델로 하며, TuNiB AI 에서 수집한 1.2TB 의 한국어 데이터로 학습되었다.<sup>34</sup> 이 연구에서 Polyglot-ko 모델 중 256 개의 A100 GPU 에서 32 만 단계에 걸쳐 1,720 억 개의 토큰에 대해 GPT-NeoX 프레임워크로 학습된 Polyglot-ko-5.8B 모델을 기반 모델로 사용하였다. 5.8B 모델보다 성능이 더 뛰어난 모델인 Polyglot-ko-12.8B 모델도 있으나, 12.8B 모델은 모델 파일 크기만 26G 에 달하는 규모로 인해 상당한 용량의 GPU 메모리를 갖춘 환경을 요한다. 또한 모델의 언어 성능을 확인할 수 있는 HellaSwag 벤치마크를 활용한 10-shot 학습 조건에서 한국어에 대한 F1-score 를 비교하였을 때, 5.8B 모델은 0.598, 12.8B 는 0.610 으로,<sup>34</sup> 5.8B 로도 연구 활용에 충분하다고 판단하였다.

각 모델의 특징 및 차이점을 정리한 표는 아래와 같다.

**표 4 각 언어모델의 특징**

	GPT3.5	LLaMA2-7B	Polyglot-ko-5.8B	Polyglot-ko-12.8B
소스코드 공개 여부	비공개	공개	공개	공개
한국어 성능	낮음	낮음	높음	높음
모델크기 (파라미터 개수)	1750 억개	70 억개	58 억개	128 억개

## 2.2 모델 Fine-tuning

언어 모델의 성능과 유용성은 특정 분야나 주제에 대한 지식을 얼마나 잘 이해하고 반영하는지에 크게 좌우된다. 특히 Fine-tuning 은 기존의 대규모 언어 모델을 특정 도메인의 지식으로 더욱 특화하는 핵심 과정이다. Fine-tuning 이란, 이미 학습된 모델에 특정한 데이터를 추가로 제공하여, 모델이 그 데이터에 특화된 지식을 습득하게 하는 과정이다.<sup>28</sup> 앞서 서술하였던 전문적 지식 자료와 뇌심혈관 질환의 업무상 질병 판정 데이터를 이용하여 AI 모델을 Fine-tuning 하면, AI 는 업무상 질병의 판정에 필요한 전문적 지식과 업무상 질병 판정 패턴을 배울 수 있다. 따라서, Fine-tuned 모델은 업무상 질병 판정의 정확성을 크게 향상할 것으로 예상된다.

### 3. 학습 데이터에 따른 모델 결과 비교

언어 모델의 핵심은 학습시키는 데이터의 내용이다. 언어 모델의 성능은 대부분 학습에 사용된 데이터에 근거하기 때문에, 학습 데이터와 학습 방법에 따라 그 출력 결과와 활용 범위가 크게 달라질 수 있다. 따라서 세 가지 다른 시나리오에 대해 출력값을 비교해 보고자 하였다.

우선 아무 데이터도 학습시키지 않은 AI 모델에 근로 시간 및 업무부담 가장요인을 준 후, 뇌심혈관 질환의 업무상 질병 인정이 되는지 물어보는 명령을 입력한다. 이 모델은 뇌심혈관 질환에 대한 지식, 뇌심혈관 질환의 업무상 질병 인정기준, 실제 근로자들의 뇌심혈관 업무상 질병 인정 사례 등 어떠한 관련된 내용을 학습한 적이 없는 '초기 상태'이기 때문에 뇌심혈관 질환에 대한 업무상 질병 인정에 대한 정보나 지식이 없는 상태이다. 따라서 이와 관련 질문이나 명령에 적절하지 못하거나 무의미한 응답을 생성하지 못할 가능성이 높다.

두 번째로는 뇌심혈관 질환에 대한 의학적 지식과, 뇌심혈관 질환의 업무상 질병 법적 인정기준을 학습시킨 모델이다. 뇌심혈관 질환에 대한 의학적 지식 및 뇌심혈관 질환의 업무상 질병 인정 기준에 대한 정보만으로, 근로자의 근로 시간 및 업무 요인을 입력했을 때 업무상 질병 판정이 가능할지 확인해 보고자 하였다.

세 번째로는 뇌심혈관 질환에 대한 의학적 지식과, 뇌심혈관 질환의 업무상 질병 법적 인정기준에 더하여 실제 근로자들의 근로 시간 및 업무 요인에 대하여 업무상 질병 판정 사례를 알 수 있는 업무상질병판정서 데이터까지 포함하여 학습된 모델이다. 전문지식에 더하여 다양한 실제 인정사례를 학습했을 때 전문지식만 학습한 두 번째 모델보다 좀 더 정확한 판단이 가능한지 비교해 보고자 하였다.

이와 같이 각 모델의 학습 데이터와 학습 데이터의 내용 및 전처리과정을 요약하면 다음 그림과 같다.

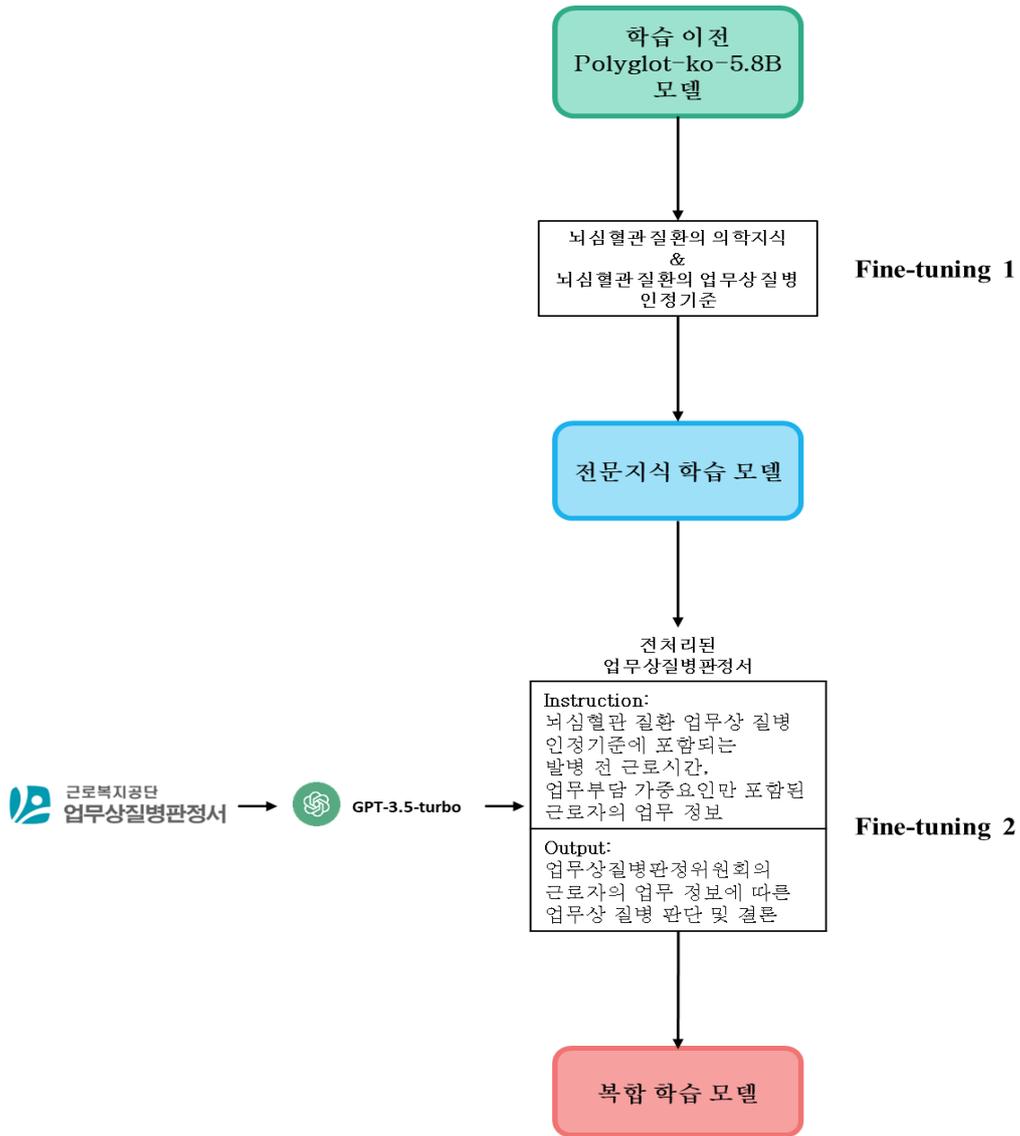


그림 2 학습 데이터 전처리 및 각 모델의 학습 데이터

#### 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 AI 모델의 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-score 를 계산하였다. 정확도, 정밀도, 민감도와 F1-score 점수를 산출하는 수식은 아래와 같다.

$$\text{정확도(Accuracy)} = \frac{TP + TN}{TP + FN + TN + FP}$$

$$\text{정밀도(Precision)} = \frac{TP}{TP + FP}$$

$$\text{민감도(Sensitivity)} = \frac{TP}{TP + FN}$$

$$\text{F1 - score} = \frac{2 * TP}{2 * TP + FN + FP}$$

여기서 TP, TN, FP, FN 은 각각 참양성(true positive), 참음성(true negative), 거짓 양성(false positive), 거짓 음성(false negative)을 의미한다. 정확도는 분류 성능 측정을 위해 사용하며, F1-score 는 모델의 정밀도와 재현율의 평균을 의미한다

또한 실제 작업환경의학 전문가의 판단 및 이에 대한 의견을 구하여 AI 모델의 판단과 전문가의 판단과의 일치도를 판단해 보고자 한다. 각 사례에 대해 AI 모델이 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 openAI 의 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 Pro 에서 A100 GPU 40GB RAM 에서 학습을 수행하였다.

## 6. 연구 윤리

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

### III. 연구 결과

#### 1. 모델 결과 비교

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

**표 5 실제 업무상 질병 판정과 비교한 세 모델의 응답 결과 비교**

실제 판정	학습 이전 모델	전문지식 학습 모델	복합 학습 모델
일치	120 건	130 건	182 건
불일치	43 건	38 건	18 건
판단불가	37 건	32 건	0 건

학습 이전 모델은 200건 중 총 80건이 실제 판정과 결과가 일치하지 않았고, 전문지식 학습 모델은 70건이 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 32건이 인정과 불인정을 판정하지 못하고 전문가의 판단이 필요하거나, 판단을 위해서는 업무와 관련한 정보가 더 필요하다는 응답을 출력하였다. 이에 관련한 질문(Instruction)에 대한 모델의 응답(Output)은 다음 그림과 같다.

Instruction

저는 돌발 상황이나 급격한 업무환경의 변화가 없었습니다. 단기간 동안의 업무상 부담으로는 업무 시간이 30% 이상 증가하지 않았습니다. 그리고 발병 전 4주간 동안의 업무시간이 1주 평균 39시간 30분이었고, 발병 전 12주 동안의 업무시간이 1주 평균 40시간 50분이었습니다. 또한, 업무부담 가중 요인은 없었습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?

Output

업무상 질병 인정 여부에 대해서는 근로자 측과 사용자 측이 서로 이견이 있을 수 있습니다. 이에 대한 판단은 근로복지공단에서 하게 될 것입니다. 근로복지공단의 판단이 있을 때까지는 귀하의 질의 내용에 대해서는 판단을 내리기가 어렵습니다.

그림 3 학습 이전 모델의 업무상 질병 판정 불가 사례

Instruction

발병 당일에는 돌발적이고 예측 곤란한 사건의 발생과 급격한 업무환경의 변화가 없었다고 확인되었습니다. 또한, 발병 전 1주일 동안 신청인은 3일 동안 업무를 수행하였으며, 총 업무시간은 44시간 45분으로 확인되었습니다. 발병 전 4주간의 주당 평균 업무시간은 42시간 20분이었고, 발병 전 12주간의 주당 평균 근무시간은 약 41시간 46분으로 확인되었습니다. 업무부담 가중요인으로는 교대제 근무가 해당됩니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?

Output

귀 공단이 관련 자료를 제출하였으므로 이를 기초로 판단이 가능하나, 관련 자료 만으로는 판정이 곤란할 것으로 판단되므로 자세한 사항은 진료시 주치의 면담을 통하여 확인하시기 바랍니다.

그림 4 전문지식 학습모델의 업무상 질병 판정 불가 사례

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 0.65으로 계산되었으며, 복합 학습 모델은 0.91로 복합 학습 모델이 가장 높았다. 정밀도는 모델 순서대로 0.71, 0.86, 0.84로 전문지식 학습 모델이 제일 높았고 그다음으로 복합 학습 모델이 가장 높았다. 민감도는 모델 순서대로 0.54, 0.54, 0.95로 복합 학습 모델이 가장 높았고, F1-score는 순서대로 0.64, 0.67, 0.93으로 복합 학습 모델이 가장 높게 계산되었다. 이에 대한 내용은 표 5에서 확인할 수 있다.

**표 6 세 모델의 성능 평가지표**

	학습 이전 모델	전문지식 학습 모델	복합 학습 모델
정확도	0.60	0.65	0.91
정밀도	0.71	0.86	0.84
민감도	0.54	0.54	0.95
F1-score	0.64	0.67	0.93

## 2. 직업환경의학 전문가의 의견

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

**표 7 복합 학습 모델에서 실제 판정과의 응답 불일치 사례에 대한 전문가 의견 1**

<p><b>Instruction:</b></p> <p>발병 전 4 주간 및 12 주간 주당 평균 업무시간은 54 시간이었습니다. 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?</p>
<p><b>Output:</b></p> <p>발병 전 12 주간 동안 및 업무환경의 변화 등에 따른 부담 요인으로 정신적 긴장을 겪었지만, 발병 전 1 주 동안 및 업무환경의 변화 등에 따른 부담 요인으로는 업무량 증가 및 작업환경의 변화는 없었습니다. 또한, 발병 전 1 주일 동안 및 업무환경의 변화 등에 따른 부담 요인으로는 업무시간이 30% 이상 증가하지 않았으며, 업무환경의 변화 및 업무상 부담의 증가가 없었습니다. 따라서, 신청인의 뇌심혈관 질환은 업무상 질병으로 불인정된다.</p>
<p><b>직업환경의학 의사 의견:</b></p> <p>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 덜 되었을 가능성이 있음. 상기 정보에 대해서는 추가 데이터를 수집하여 training 할 필요가 있다고 보임.</p>

실제 판정과 복합 학습 모델의 첫 번째 불일치 사례에 대하여, 모델은 환자의 정보를 바탕으로, 발병 전 12 주 동안의 부담 요인과 정신적 긴장을 인지하였지만, 발병 전 1 주 동안의 업무 부담 증가나 환경 변화 등의 정보가 제시되지 않았다고 판단하였다. 이에 따라 모델은 뇌심혈관 질환의 인정 기준 중 발병 전 12 주 동안 1 주 평균 52 시간 이상 근무하면서 업무부담 가장요인이 하나 이상이면 업무상 관련이 높다는 기준을 학습했음에도 불구하고, 신청인의 뇌심혈관 질환을 업무상 질병으로 불인정 하였다.

이에 대하여 전문가는 첫째로, 법적 필수 요소인 단기 과로 증가(30% 증가)에 대한 정보가 질문에서 명확하게 제시되지 않았다는 점을 지적하였다. 이에 따라, 모델이 이 정보를 기반으로 한 판정을 할 수 없었을 것이라는 의견이다. 둘째로, 부가 요인인 환경 변화, 스트레스, 물리적 노출 등에 대한 모델의 학습이 덜 되었을 가능성을 제기하였다. 이를 바탕으로, 전문가는 추가적인 데이터 수집 및 모델의 재학습이 필요하다고 권고하였다.

실제 판정과의 불일치 3 건에 대한 전문가의 의견 중 두 번째 건은 다음과 같다.

**표 8 복합 학습 모델에서 실제 판정과의 응답 불일치 사례에 대한 전문가 의견 2**

<p><b>Instruction:</b></p> <p>저는 최근에 발병하기 전 1 주 동안 총 업무시간이 57 시간 30 분으로, 일상 업무 시간보다 30% 이상 증가했고, 업무량도 30% 이상 증가했습니다. 또한, 발병하기 전 4 주간의 주당 평균 업무시간은 48 시간이었고, 12 주간의 1 주당 평균 업무시간은 42 시간 39 분이었습니다. 제 업무는 육체적으로 매우 힘든 업무를 수행하는 것이 가장요인입니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
<p><b>Output:</b></p> <p>신청인의 뇌심혈관 질환은 업무상 질병으로 불인정된다. 발병 전 12 주간의 평균 업무시간을 기준으로 30% 이상 증가하지 않았고, 교대제 업무를 수행하지도 않았습니니다. 또한, 발병 전 6 개월 이내에는 업무량, 환경변화, 업무상 부담 증가에 대한 업무신상요인이 없었습니다. 발병 전 1 주일 동안의 평균 업무시간은 26 시간 24 분으로, 일상 업무 시간보다 30% 이상 증가하지 않았습니니다. 또한, 발병 전 12 주간의 평균 업무시간을 기준으로 30% 이상 증가하지 않았고, 업무부담 가장요인으로는 교대제 업무 및 정신적 긴장이 높은 업무를 수행하지 않았습니니다. 그러나 주말 근무는 하지 않았습니니다. 급속도 또는 만성적인 과중한 업무를 수행한 사실은 인정되나, 신청인의 뇌심혈관 질환이 업무상 질병으로 인정되지 않는다.</p>
<p><b>직업환경의학 의사 의견:</b></p> <p>우선 48 시간의 1.3 배는 62 시간임. 따라서 현재 57 시간은 평소의 30% 증가는 아님. 다만 57 시간이라는 과로와 부가 요인인 정신적 긴장이 있었음에도, 30% 증가에 대한 복합적 판단은 어려운 것으로 보임. 질문을 줄때 법적 필수 요소인, 단기 과로 증가(30% 증가)에 대한 모델 training 이 필요할 것으로 보임. 이것이 인정되는 사례만</p>

oversampling 하여 training 이 필요할 것으로 보임. 현재를 ver1 으로 부르고 단기과로 30%를 집중적으로 학습시킨 모델이 필요하다고 판단됨.

실제 판정과 복합 학습 모델의 두 번째 불일치 사례에 대하여, 질문에서는 뇌심혈관 질환 발병 1주일 전에 일상 업무 시간보다 30% 이상 증가했고, 업무량도 30% 이상 증가했다고 서술했으나, 모델은 이를 인지하지 못하는 응답을 출력하였다. 그리고 발병 전 4주 및 12주 동안의 1주 평균 근로 시간도 과로의 기준에 해당하지 않으나, 만성적인 과중한 업무를 수행하였다고 판단하였다. 결론적으로 실제 판정은 업무상 발병 1주일 전 단기 과로로 인한 업무상 질병으로 인정된 사례였으나, 모델은 불인정으로 판단하였다.

이에 대하여 전문가는 발병 전 4주 및 12주 동안의 근로 시간뿐만 아니라 발병 1주일 전 평균 근로 시간과 같은 단기 과로의 30% 증가 사례에 대한 학습이 추가로 필요하다는 의견을 제시하였다. 이러한 특정 상황들에 대해 오버샘플링 하여 집중적으로 학습된 새로운 모델이 필요하다고 권고하였다.

실제 판정과의 불일치 3 건에 대한 전문가의 의견 중 세 번째 건은 다음과 같다.

**표 9 복합 학습 모델에서 실제 판정과의 응답 불일치 사례에 대한 전문가 의견 3**

<p><b>Instruction:</b></p> <p>저는 돌발상황이나 급격한 업무환경의 변화로 인해 뇌심혈관 질환이 발생할 수 있습니다. 제 업무량과 시간은 일상업무보다 30% 이상 증가하지 않았으며, 발병 전 4 주간의 평균 업무시간은 52 시간이었고, 발병 전 12 주간의 평균 업무시간은 56 시간 35 분이었습니다. 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?</p>
<p><b>Output:</b></p> <p>신청인의 뇌심혈관 질환은 업무상 질병으로 불인정된다. 발병 전 1 주일 동안의 평균 업무시간은 53 시간 37 분이었고, 발병 전 12 주간의 평균 업무시간은 55 시간 30 분이었습니다. 또한, 발병 전 4 주간의 평균 업무시간은 57 시간 43 분이었고, 발병 전 12 주간의 평균 업무시간은 53 시간 23 분이었습니다. 이 외에도, 업무 부담가중요인으로는 교대제 업무수행과 정신적 긴장이 높은 업무, 유해한 작업환경이 있으며, 업무상 질병으로 인한 인정기준을 충족하지 못했으므로, 신청인의 뇌심혈관 질환은 업무상 질병으로 불인정된다.</p>
<p><b>직업환경의학 의사 의견:</b></p> <p>이는 어떠한 기준이 업무 부담으로 여길지에 대한 데이터 학습이 아직 없는 모델이기 때문에 발생한 문제임. 업무 부담이 구체적으로 무엇인지 판단하기 위한 모델 학습이 필요하고, 향후 chat 방식일 경우 업무 부담에 대해서 구체적 정보를 요구하는 모델을 만들 필요가 있음. 이는 Template 형식으로 가능하다고 보임.</p>

세 번째 사례에 대하여, 모델은 신청인의 뇌심혈관 질환을 업무상 질병으로 불인정하는 판단을 내렸다. 발병 전의 업무시간과 유해한 업무 환경, 그 외의 여러 요인을 기반으로 한 판단이었다. 하지만 일부 제시된 정보와 모델의 응답에는 일관성이 없는 부분이 발견되었다.

전문가는 현재의 모델 관점에서 업무 부담에 대한 충분한 학습이 되지 않은 결과로 판단된다고 지적하였다. 업무 부담의 정의와 그에 따른 판단 기준에 대한 학습이 필요하다는 의견을 제시하였다. 특히, 향후 챗봇 방식의 모델에서는 챗봇 사용자로부터 업무 부담에 관한 구체적 정보를 받아들일 수 있는 템플릿 형식의 질문을 구성하는 것이 바람직하다고 권고하였다.

## IV. 고찰

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.

딥러닝 기반의 모델은 데이터 주도 방식으로 학습되며, 그렇기에 학습 데이터의 품질은 모델 성능에 결정적인 역할을 한다.<sup>35</sup> 특히 뇌심혈관 질환과 같은 의학적 주제는 복잡하고 다양한 변수와 패턴을 포함하고 있어, 그 특성을 정확하게 반영하는 학습 데이터가 필요하다. 모델의 '지식'은 결국 학습 데이터로부터 축적된다. 실제의 다양한 상황, 환경 변화, 환자 특성 등의 패턴과 정보는 학습 데이터를 통해 모델에게 전달된다. 그렇기 때문에, 데이터가 모델에게 어떤 지식을 제공하는지, 그리고 그 지식이 어떻게 모델의 판단 능력에 영향을 미치는지에 대한 이해는 필수적이다.<sup>36</sup> 이러한 AI 모델의 특성에 기반하여, 다양한 데이터의 학습은 AI 판단의 정확도 향상에 결정적인 역할을 한다고 판단된다. 실제 근로자들의 뇌심혈관 질환 업무상 질병 판정은 단순히 근로 시간이나 업무 환경만으로 판단하기에는 근로자의 나이, 성별, 과거 건강검진 이력, 흡연 및 음주 등 생활 습관 등 여러 가지를 종합적으로 고려해야 한다. 따라서 뇌심혈관 질환의 의학적 정의나 업무상 질병으로의 법적 인정기준만 아는 상태에서는 뇌심혈관 질환의 업무상 질병 판정은 어렵다고 판단된다.

그러나 복합 학습 모델도 한계점은 있었다. 발병 전 4 주 및 12 주 동안의 근로 시간 및 업무부담 가중요인으로 판단할 수 있는 만성 과로뿐만 아니라, 발병 전 1 주일 동안의 근로 시간 및 업무 요인으로 확인할 수 있는 단기 과로나 야간 근무 등 다른 업무상

질병 인정 요소들을 종합적으로 판단하지 못하는 모습을 보였다. 이에 대하여 전문가는 절대적인 학습 데이터의 수가 부족하고, 또한 현재 학습 데이터에 만성 과로 요인에 대한 내용으로 치중되어 있어 과적합 현상이 발생했다고 판단하였다. 과적합(Overfitting) 현상이란, 머신 러닝 및 딥러닝의 주요 문제점 중 하나로, 학습 모델이 학습 데이터에 과도하게 최적화되어 실제 환경에서나 테스트 데이터에 대해 일반화된 성능을 보이지 못하는 현상을 의미한다. 이 문제는 학습 과정에서 모델이 데이터의 기본적인 패턴과 노이즈까지 구분하지 못하고 모두 학습하게 되어 발생한다. 이 경우, 학습 모델이 학습 데이터에 있는 오류 및 편향까지 학습해 버려서 응답의 품질이 낮아질 수 있다.<sup>37</sup> 모델이 과적합되면, 학습 데이터에 대해서는 정확도가 매우 높게 나오지만, 새로운 질문, 즉 테스트 데이터에 대해서는 정확도가 낮고 일관성 없는 응답을 도출할 수 있다.

따라서 이러한 결과는 모델이 단기 과로와 같은 복잡하고 다양한 패턴을 갖는 사례에 대해 충분히 학습하지 못했음을 시사한다. 이는 학습 데이터의 다양성 부족이나, 특정 유형의 사례에 대한 데이터가 불충분했을 가능성을 나타낸다. 실제 업무상질병판정위원회에서 업무상 질병 판정 시 단기 과로 및 발병 전 돌발상황, 야간 근무 여부가 업무상 질병 인정 기준에 부합하는 사례가 적기 때문에, 학습 데이터에서도 만성 과로를 기준으로 인정 및 불인정 판정 사례가 다수였던 것으로 판단된다. 또한, 단기 과로나 야간 근무와 같은 사례는 만성 과로와 달리 더욱 복잡한 인과 관계와 상황적 요소들을 포함하고 있을 수 있으며, 이는 AI 모델이 그러한 요소들을 충분히 이해하고 반영하는 데 한계가 있음을 보여준다.

이 연구에서는 모델의 업무상질병 판정 여부에 대하여 인정과 불인정으로 결과를 이분류하고, 분류모델 성능 평가지표를 활용하여 모델의 성능을 평가하였다. 그 결과, 정확도는 복합 학습 모델에서 가장 높았으며, 이는 복합 학습 모델이 전문지식과 실제 사례 데이터를 모두 통합하여 학습함으로써, 더 정확한 판단을 내릴 수 있음을 시사한다. 그러나 복합 학습 모델의 정밀도는 0.84 로, 전문지식 학습 모델의 0.86 에 비해 낮았다. 즉 복합 학습 모델이 불인정 사례를 예측하는 데 있어 좀 더 포괄적이어서, 더 많은 사례를 불인정으로 분류했지만 그 중 일부는 실제로는 '인정' 사례였을 수 있다. 이는 실제 사례에 대한 업무상질병판정위원회 판단의 다양성을 반영하면서 발생하는 특정 유형의 오류를

허용하는 경향이 있음을 시사한다. 마지막으로 복합 학습 모델의 민감도와 F1-score 는 다른 모델에 비해 압도적으로 높았는데, 이는 복합 학습 모델이 정밀도와 민감도 사이의 균형을 잘 맞추며 전반적으로 높은 성능이 보임을 나타낸다.

이에 따라 연구의 향후 방향은 이러한 유형의 사례에 대한 데이터를 추가하고, 모델이 다양한 업무상 질병 판정의 복잡성을 더욱 잘 이해하고 처리할 수 있도록 하는 것에 초점을 맞추어야 한다고 판단된다. 이는 모델의 일반화 능력을 강화하고, 실제 업무 환경에서 발생할 수 있는 다양한 상황에 더욱 효과적으로 대응할 수 있게 만들어 줄 것이다.

또한 전문가는 업무 부담에 대한 구체적 정보를 수집하고 분석하는 능력을 갖춘 모델의 개발이 필요함을 강조했다. 특히, 챗봇 형태의 모델이 사용자로부터 업무 부담과 관련된 상세 정보를 효과적으로 요구하고 처리할 수 있어야 한다고 제언했다. 이는 Template-based 방식을 통해 구현될 수 있으며, 사용자가 자신의 업무 환경과 건강 상태에 대해 보다 명확하고 상세하게 기술할 수 있도록 돕는 역할을 할 것으로 판단된다. 이러한 접근은 사용자의 개별적인 상황을 더 잘 이해하고, 그에 따른 적절한 판단을 제공하는 데 중요할 것이다. 더불어 챗봇 개발을 위해서는 몇 가지 중요한 요소들이 고려되어야 한다. 우선, 챗봇이 사용자로부터 업무 부담에 관한 구체적이고 다양한 정보를 수집하고, 이를 효과적으로 분석할 수 있는 기능이 필요하다. 이는 사용자의 업무 환경, 작업 조건, 생활 습관, 건강 상태 등을 포괄하는 데이터를 기반으로 할 것이다. 또한, 챗봇은 사용자에게 친숙하고 직관적인 인터페이스를 제공하여, 사용자가 쉽게 정보를 제공하고 상호작용할 수 있도록 해야 한다. 이러한 접근은 챗봇이 업무상 질병 판정 및 예방, 관리에 있어 중요한 도구로서의 역할을 수행하는 데 도움이 될 것이다.

본 연구의 몇 가지 강점 중 하나는, 근로복지공단의 업무상질병판정서를 기반으로 한 새로운 AI 모델을 개발함으로써 뇌심혈관 질환과 업무 환경 간의 연관성을 파악하는 데 중요한 진전을 이루었다. 특히, 만성 과로와 같은 업무상 질병 인정 여부를 정확하게 판단하는 데 이 모델이 상당한 성능을 보인 것은 주목할 만한 성과이다. 이는 업무 환경과 건강 상태 간의 복잡한 상관관계를 이해하는 데 중요한 역할을 할 수 있으며, 특히 배경지식이 부족한 사람들에게 자신의 건강 상태와 업무 환경 간의 관계에 대한 기초적인

스크리닝을 제공하는 유용한 도구로 활용될 수 있다. 두 번째로, 한국어에 특화된 Polyglot-ko 모델을 사용함으로써 한국어 데이터에 대한 모델의 성능을 극대화했다. 마지막 연구의 강점은, 업무상질병판정서 전처리 및 데이터 문장 변환, 변환된 데이터 검증 등의 프로세스를 ‘Machine In The Loop’로 진행하였다는 점이다. 근로복지공단의 업무상질병판정서에서 모델 학습에 필요한 업무환경 관련 요인만 추출, 그 후 모델 학습에 용이하게 구어체 문장으로 변환하는 과정을 gpt-3.5-turbo 모델을 활용하였다. ‘Human In The Loop’ 프로세스로 진행하면 연구자가 직접 전처리 및 변환 과정에 가담하기 때문에 소량의 데이터 전처리 과정에서는 좀 더 연구자의 목적에 부합하게 데이터를 만들 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. ‘Machine In The Loop’를 설명한 그림은 아래와 같다.

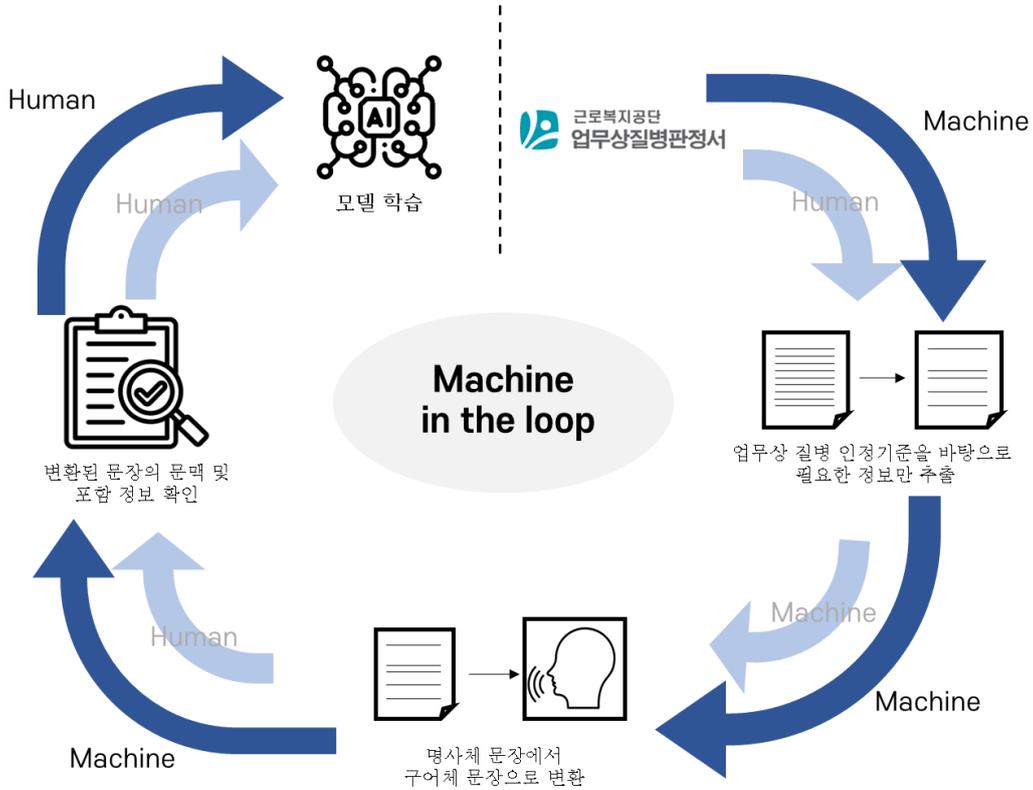


그림 5 Machine In The Loop 프로세스

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 않음으로써, 모델의 판정 범위와 정확성에 제한을 두었다. 둘째, 모델은 단기 코로나 야간 근로와 같은 특정 사례에 대해 상대적으로 낮은 성능을 보였다. 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. 마지막으로, 현재 모델은 업무상 질병의 인정 여부만을 판단할 수 있으며, 질병 관리 방안이나 후속 조치에 대한 조언을 제공하는 데에 어려움이 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. 결론

본 연구에서는 뇌심혈관 질환에 대한 전문 지식과 업무상 질병 인정에 대한 정보가 없는 근로자들도 본인의 뇌심혈관 질환과 근로 시간 및 업무 요인의 연관성 및 업무상 질병 인정 여부를 알 수 있는 AI 모델을 개발하고자 했다. 어느 정도의 정보를 학습하여야 업무상 질병 판정이 가능한지 알아보기 위해 학습 이전 모델, 전문지식 학습모델, 복합 학습 모델의 응답 결과를 비교하였고, 뇌심혈관 질환의 업무상 질병과 관련한 전문 지식과 함께 실제 업무상질병판정 사례를 학습한 복합 학습 모델의 성능이 다른 모델보다 더 높은 성능을 보여주었다.

이번 연구를 통해 실제 뇌심혈관 질환의 업무상질병 판정에 대하여 높은 성능을 보이는 한국어 친화적인 AI 모델을 개발하였다는 점, 실제 한국인 근로자에게 건강과 업무환경의 연관성에 대한 기초적인 스크리닝 도구가 될 수 있다는 점에서 의의가 있다. 그러나 아직 다양한 사례에 대해 판단이 미흡하다는 점, 추후 건강관리 방안을 제공하지 못하는 점에서 한계점을 가진다.

따라서 이러한 한계점을 보완하여, 실제 근로자들이 자신의 업무 요인과 뇌심혈관 질환의 연관성 및 업무상 질병 인정 여부를 판단해보고 건강관리 방안을 제공해주는 챗봇 형식의 WebUI 개발이 필요하다고 판단된다. 본 연구의 모델 응답 성능을 높인 후, Template-based 방식의 모델이 필요하며, 이를 통해 실제 근로자들이 본인의 건강과 업무 환경간의 연관성을 인지하고 스스로 건강관리를 하는 데에 도움을 줄 수 있는 도구가 될 수 있을 것이다.

## 참고문헌

1. Mora CD, Bornstein RA. Evaluation of Cerebrovascular disease. In: *Neuropsychology*. Springer; 1998:171-186.
2. Genders TS, Hunink MM. Epidemiology of coronary artery disease. *Clinical applications of cardiac CT*. 2012:3-6.
3. Bishop T. Coronary heart disease. *Primary Health Care*. 2016;26:14-14.
4. Sandoval Y, Thygesen K, Jaffe AS. The universal definition of myocardial infarction: present and future. *Circulation*. 2020;141(18):1434-1436.
5. HATHAWAY LR. Stroke (brain attack). *Nursing2023*. 2005;35(11):49.
6. 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. *Lancet (London, England)*. 2018;392(10159):1736-1788.
7. Deaton C, Froelicher ES, Wu LH, Ho C, Shishani K, Jaarsma T. The global burden of cardiovascular disease. *European journal of cardiovascular nursing*. 2011;10 Suppl 2:S5-13.
8. Thrift AG, Cadilhac DA, Thayabaranathan T, et al. Global stroke statistics. *International journal of stroke : official journal of the International Stroke Society*. 2014;9(1):6-18.
9. Organization WH. Cardiovascular diseases (CVDs). 2017; [http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
10. 사망원인별 사망률 추이. 2022; [https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx\\_cd=1012](https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=1012).
11. Wu Y, Xiong Y, Wang P, et al. Risk factors of cardiovascular and cerebrovascular diseases in young and middle-aged adults: A meta-analysis. *Medicine*. 2022;101(48):e32082.
12. Vineis P, Demetriou CA, Probst-Hensch N. Long-term effects of air pollution: an exposome meet-in-the-middle approach. *International journal of public health*. 2020;65(2):125-127.

13. Kim UJ, Choi WJ, Kang SK, et al. Standards for recognition and approval rate of occupational cerebro-cardiovascular diseases in Korea. *Annals of occupational and environmental medicine*. 2022;34:e30.
14. Cherry N. Recent advances: occupational disease. *BMJ (Clinical research ed)*. 1999;318(7195):1397-1399.
15. Vieriu D. General Considerations Regarding the Classification, Record and Prevention of Occupational Diseases. 2015.
16. van der Molen HF, Frings-Dresen MH. Occupational diseases: from cure to prevention. In: Vol 8: MDPI; 2019:1681.
17. 뇌혈관 질병 또는 심장 질병 및 근골격계 질병의 업무상 질병 인정 여부 결정에 필요한 사항. In: 고용노동부 고시 제2022-40호, 2022.4. 28. 발령, 2022. 7. 1. 시행) 제1호가목.
18. Lin R-T, Chien L-C, Kawachi I. Nonlinear associations between working hours and overwork-related cerebrovascular and cardiovascular diseases (CCVD). *Scientific Reports*. 2018;8(1):9694.
19. Kamardeen I. Work stress related cardiovascular diseases among construction professionals. *Built Environment Project Asset Management*. 2022;12(2):223-242.
20. Kim S, Jung Y. Effect of Long Working Hours on Cardiovascular Disease in South Korean Workers: A Longitudinal Study. *Asia-Pacific journal of public health*. 2021;33(2-3):213-219.
21. Park S, Nam J, Lee JK, Oh SS, Kang HT, Koh SB. Association between night work and cardiovascular diseases: analysis of the 3rd Korean working conditions survey. *Annals of occupational and environmental medicine*. 2015;27:15.
22. Picciotto D. [Occupational cardiovascular diseases and phlebopathies]. *Giornale italiano di medicina del lavoro ed ergonomia*. 2010;32 4 Suppl:160-161.
23. Dovgusha LV, Petruhin NN. On the issue of criteria for recognition of occupational disease as an insured event in health care workers. *Russian Journal of Occupational Health and Industrial Ecology*. 2020.
24. Guthrie R, Westaway JA. Emerging legal concerns with chronic diseases in the Australian workplace: pre-employment medicals, functional capacity evaluations,

- workers' compensation and disability discrimination. *Journal of law medicine*. 2009;16 5:803-821.
25. 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.
  26. Mirza B, Syed TQ, Khan B, Malik Y. Potential Deep Learning Solutions to Persistent and Emerging Big Data Challenges—A Practitioners' Cookbook. *ACM Computing Surveys*. 2021;54:1 - 39.
  27. Zhuang Y-tW, Fei Chen, Chun Pan, Yun-he. Challenges and opportunities: from big data to knowledge in AI 2.0. *Frontiers of Information Technology Electronic Engineering*. 2017;18:3-14.
  28. Trivedi UB, Mishra P. Fine Tuning (Diagnosis) of Machine Learning Algorithm (Model) for optimization. Paper presented at: 2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)2022.
  29. 노동특수이론 및 업무상재해관련소송. 사법연수원.
  30. 산업재해보상보험법 시행령 제37조제5항. In:[시행 2023. 7. 1.] [법률 제18928호, 2022. 6. 10., 일부개정].
  31. 산업재해보상보험법 시행령 제34조제3항 및 별표 제 1호가목 단서. In:[시행 2023. 7. 1.] [법률 제18928호, 2022. 6. 10., 일부개정].
  32. 주하영, 오현택, 양진홍. 오픈 소스 기반의 거대 언어 모델 연구 동향: 서베이. *한국정보전자통신기술학회 논문지*. 2023;16(4):193-202.
  33. Lin XV, Mihaylov T, Artetxe M, et al. Few-shot learning with multilingual language models. *arXiv preprint arXiv:10668*. 2021.
  34. Ko H, Yang K, Ryu M, Choi T, Yang S, Park S. A Technical Report for Polyglot-Ko: Open-Source Large-Scale Korean Language Models. *arXiv preprint arXiv:02254*. 2023.
  35. Jang SS, Hwang KH, Ha YG. High Quality Training Set Collection using Generative Adversarial Network. Paper presented at: 2020 IEEE International Conference on Big Data and Smart Computing (BigComp)2020.

36. 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.
37. Montesinos López OA, Montesinos López A, Crossa J. Overfitting, model tuning, and evaluation of prediction performance. In: *Multivariate statistical machine learning methods for genomic prediction*. Springer; 2022:109-139.

## 부록

부록 표 1 모델 테스트 시 사용한 테스트 데이터

No.	Instruction
1	<p>증상이 발생하기 전 24시간 이내에 돌발 상황이나 급격한 업무환경의 변화는 없었습니다. 증상이 발생하기 전날에는 평소와 같이 업무를 수행했습니다. 발병 전 1주간의 업무시간은 총 57시간 00분이었고, 발병 전 1주 이내에는 업무환경의 변화, 강도, 책임, 업무량의 변화는 없었으며 평소와 같이 업무를 수행했습니다. 또한, 발병 전 4주 동안의 주당 평균 업무시간은 55시간 16분이었고, 발병 전 12주 동안의 주당 평균 업무시간은 44시간 36분이었습니다. 이러한 정보를 바탕으로 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
2	<p>발병 전 4주간 및 12주간 주당 평균 업무시간은 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?</p>
3	<p>발병 전일에는 돌발 상황이나 급격한 업무환경의 변화는 없었습니다. 발병 전 1주일 동안의 업무량, 강도, 책임, 환경의 변화를 확인한 결과, 일상 업무시간은 30% 증가하지 않았으며, 그 외에도 업무환경의 변화나 특이사항은 없었습니다. 발병 전 4주 동안의 업무시간은 1주당 평균 21시간 52분이었고, 12주 동안의 업무시간은 1주당 평균 18시간 7분이었습니다. 또한, 업무환경의 변화나 특이사항은 없었습니다.</p> <p>업무부담 가중요인과 관련하여 특이사항은 없었습니다. 이러한 정보를 바탕으로 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
4	<p>저는 증상이 발생하기 전 24시간 이내에 돌발 상황이 발생하거나 급격한 작업환경의 변화가 없었습니다. 발병 전 4주 동안의 주당 평균 업무시간은 31시간 48분이었고, 발병 전 12주 동안의 주당 평균 업무시간은 36시간 53분이었습니다.</p>

	다. 이러한 정보를 바탕으로 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?
5	저는 발병 당일에는 돌발적이고 예측 곤란한 사건의 발생과 급격한 업무 환경의 변화가 없었으며, 발병 전 1주 이내에는 일상 업무시간 및 업무량이 30% 이상 증가되지 않았습니다. 또한, 발병 전 4주간의 주당 평균 업무시간은 약 26시간 49분이었고, 발병 전 12주간의 주당 평균 근무시간은 약 26시간 44분이었습니다. 따라서, 업무부담 가중요인은 해당되지 않습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?
6	저는 증상이 발생하기 전 24시간 이내에 돌발적인 상황이나 급격한 업무환경 변화가 없었습니다. 발병하기 전 1주일 동안에는 업무량, 강도, 책임, 환경변화에 대한 특이사항은 없었고, 일상업무보다 30% 이상 증가하지 않았습니다. 발병하기 전 4주 동안에는 주당 평균 30시간 2분의 업무시간이 있었고, 발병하기 전 12주 동안에는 주당 평균 28시간 21분의 업무시간이 있었습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?
7	저는 돌발적 사건이나 급격한 업무환경 변화로 인해 발생한 질환이 없습니다. 또한, 단기간 동안의 업무상 부담도 없었습니다. 제가 겪은 업무는 일상 업무시간과 비교해도 업무량이나 시간이 30% 이상 증가하지 않았습니다. 발병 전 4주 동안 평균적으로 1주당 49시간 42분을 일하고, 발병 전 12주 동안 평균적으로 1주당 45시간 45분을 일했습니다. 그러나, 업무부담 가중요인은 해당되지 않습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?
8	신청인은 발병 당일에는 돌발적이고 예측 곤란한 사건의 발생과 급격한 업무 환경의 변화가 없었다고 확인되었습니다. 또한, 발병 전 1주일 동안 신청인은 3일 동안 업무를 수행하였으며, 총 업무시간은 44시간 45분으로 확인되었습니다. 발병 전 4주간의 주당 평균 업무시간은 42시간 20분이었고, 발병 전 12주

	<p>간의 주당 평균 근무시간은 약 41시간 46분으로 확인되었습니다. 업무부담가 중요인으로는 교대제 근무가 해당됩니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
9	<p>저는 돌발 상황이나 급격한 업무환경의 변화에 처해있었습니다. 또한, 일상 업무 외에도 다른 업무를 수행하였습니다. 제 뇌심혈관 질환이 발병하기 전 1주일 동안은 일상 업무 시간이나 업무량이 30% 이상 증가하지 않았습니다. 발병하기 전 4주 동안 제 업무시간은 1주당 평균 65시간이었습니다. 발병하기 전 12주 동안 제 업무시간은 1주당 평균 61시간 50분이었습니다. 제 업무는 교대제로 격일근무를 수행하는 것이었습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
10	<p>저는 돌발 상황이나 급격한 업무환경의 변화가 없었습니다. 단기간 동안의 업무상 부담으로는 업무 시간이 30% 이상 증가하지 않았습니다. 그리고 발병 전 4주간 동안의 업무시간이 1주 평균 39시간 30분이었고, 발병 전 12주 동안의 업무시간이 1주 평균 40시간 50분이었습니다. 또한, 업무부담가중 요인은 없었습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
11	<p>저는 최근에 발병하기 전 1주 동안 총 업무시간이 57시간 30분으로, 일상 업무 시간보다 30% 이상 증가했고, 업무량도 30% 이상 증가했습니다. 또한, 발병하기 전 4주간의 주당 평균 업무시간은 48시간이었고, 12주간의 1주당 평균 업무시간은 42시간 39분이었습니다. 제 업무는 육체적으로 매우 힘든 업무를 수행하는 것이 가장요인입니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
12	<p>저는 돌발 상황이나 급격한 업무환경의 변화가 없었습니다. 발병 전 1주일 이내에는 업무량, 강도, 책임, 환경의 변화가 확인되지 않았습니다. 또한, 발병 전 4주간 동안의 업무시간은 1주 평균 32시간 12분이었고, 발병 전 12주 동안의 업무시간은 1주 평균 35시간 38분이었습니다. 따라서, 저의 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>

13	<p>저는 돌발 상황이나 급격한 업무환경의 변화는 없었습니다. 발병 전 1주일 동안의 업무량, 강도, 책임, 환경의 변화도 없었습니다. 또한, 업무 시간이 30% 이상 증가하지 않았습니다. 그리고 발병 전 4주간 동안의 업무시간은 1주 평균 18시간 30분이었고, 발병 전 12주 동안의 업무시간은 1주 평균 19시간 10분이었습니다. 따라서, 저의 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
14	<p>저는 돌발상황이나 급격한 업무환경의 변화로 인해 뇌심혈관 질환이 발생할 수 있습니다. 제 업무량과 시간은 일상업무보다 30% 이상 증가하지 않았으며, 발병 전 4주간의 평균 업무시간은 52시간이었고, 발병 전 12주간의 평균 업무시간은 56시간 35분이었습니다. 또한, 제 업무는 유해한 작업환경에 노출되고 육체적으로도 강도가 높은 업무입니다. 이러한 뇌심혈관 질환 업무상 질병 인정 기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
15	<p>저는 돌발적으로 발병한 뇌심혈관 질환이 있습니다. 발병 당일에는 평소와 같이 업무를 수행하였으며, 돌발적인 상황이나 급격한 업무환경 변화는 없었습니다. 또한, 발병 전 1주 동안에는 업무의 양이 30% 이상 증가하지 않았고, 발병 전 1주간의 업무시간도 30% 이상 증가하지 않았습니다. 또한, 발병 전 4주 동안의 업무시간은 1주 평균 60시간 03분이었고, 발병 전 12주 동안의 업무시간은 1주당 평균 62시간 25분이었습니다. 또한, 업무부담 가중요인에 해당하는 사항은 없습니다. 이러한 뇌심혈관 질환이 뇌심혈관 질환 업무상 질병 인정 기준을 바탕으로 업무상 질병으로 인정될 수 있을까요?</p>
16	<p>제 뇌심혈관 질환이 업무와 관련한 돌발상황 또는 급격한 업무환경의 변화로 인해 발생하지 않았습니다. 제 뇌심혈관 질환이 업무상 부담으로 인해 발생하지 않았습니다. 제 업무시간은 30% 이상 증가하지 않았습니다. 발병 전 4주 동안의 1주당 평균 업무시간은 45시간 2분이었고, 발병 전 12주 동안의 1주당 평균 업무시간은 42시간 50분이었습니다. 또한, 제 업무는 유해한 작업환경에 노출되었습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>

17	<p>저는 돌발상황이나 급격한 업무환경의 변화와 관련된 업무를 수행하지 않았습니다. 또한, 발병 전 1주간의 업무 시간은 일상업무 시간보다 30% 이상 증가하지 않았으며, 업무량도 30% 이상 증가하지 않았습니다. 또한, 발병 전 4주간의 주당 평균 업무시간은 65시간 07분이었고, 발병 전 12주간의 1주당 평균 업무시간은 60시간 05분이었습니다. 따라서, 업무부담 가중요인은 해당되지 않습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
18	<p>저는 돌발상황이나 급격한 업무환경의 변화와는 관련이 없습니다. 발병 전 1주간의 업무 시간은 일상업무 시간과 비교하여 30% 이상 증가하지 않았으며, 업무량도 30% 이상 증가하지 않았습니다. 또한, 발병 전 4주간의 주당 평균 업무시간은 31시간 45분이었고, 발병 전 12주간의 1주당 평균 업무시간은 35시간 35분이었습니다. 업무부담 가중요인에 해당하는 사항은 없습니다. 따라서, 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
19	<p>발병 전 1주간의 업무시간은 일상업무 시간보다 30% 이상 증가하지 않았고, 업무량도 30% 이상 증가하지 않았습니다. 발병 전 4주 동안 1주 평균 업무시간은 39시간 30분이었고, 발병 전 12주간의 1주당 평균 업무시간은 37시간 35분이었습니다. 제 업무는 유해한 작업환경에서 이루어지고 있습니다. 뇌심혈관 질환 업무상 질병 인정기준을 바탕으로, 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>
20	<p>발병 전 12주간 주간 평균 72시간 42분 일하였고, 4주간 주간 평균 67시간 50분 일하였습니다. 정신적으로 스트레스가 큰 업무를 수행했고, 소음이 큰 유해한 환경에 노출되었습니다. 이러한 정보를 바탕으로 제 뇌심혈관 질환이 업무상 질병으로 인정될 수 있을까요?</p>

## ABSTRACT

### **Development of an AI model to assess the association between cerebrovascular disease and the work environment and analysis of consistency with expert judgment**

Juyeon Oh

*Department of Public Health  
The Graduate School, Yonsei University*

(Directed by Professor Jin-Ha Yoon)

**Background:** Cerebrovascular disease(CVD) is a major health problem worldwide and can be caused by a variety of factors. Among them, overwork and excessive workload can cause and exacerbate the condition, which can be recognized as an occupational disease. However, due to the lack of knowledge on CVD and occupational disease determination, it is difficult for the average worker to recognize the relationship between their condition and their work environment. Therefore, we aimed to develop an AI model that can help general workers easily understand and judge the relationship between CVD and occupational diseases.

**Methods:** The data used to train the model includes medical and legal expertise and actual cases of occupational diseases. The model used for training is the Polyglot-en-5.8B model, which has good performance in Korean by training on large-scale Korean data. The responses of the model before training, the model that learned only specialized knowledge, and the model that learned a combination of specialized knowledge and actual occupational disease cases are compared with the actual cases. The performance of the trained model was checked by accuracy, precision, sensitivity, and F1-score. Finally, expert opinions were obtained by reading the model responses and reviewing the opinions and points of improvement from occupational and environmental medicine doctor.

**Results:** As a result of comparing the responses of the three models, the model that learned a combination of expert knowledge and various occupational disease determination cases had the highest performance compared to the other two models. The accuracy was 0.60, 0.65, and 0.91 for the pre-learning model, the expert knowledge learning model, and the complex learning model, in that order, and the precision was 0.71, 0.86, and 0.84 in that order. Sensitivity was 0.54, 0.54, and 0.95, and finally, F1-score was 0.64, 0.67, and 0.93 in that order. However, the complex learning model also output some inconsistent or

incorrect responses. In response to this, the expert suggested that additional learning is needed on various cases such as short-term overwork and night work, and that a template-type model for chatbot is needed in the future.

**Discussion:** Responses to the occupational disease determination of CVD were compared using three models (transfer learning model, expert knowledge learning model, and complex learning model). As a result, the complex learning model showed the most accurate judgment. This study is significant in that it developed a new, unprecedented AI model that fine-tuned an actual case study called the Korea Workers' Compensation and Welfare Service's Occupational Disease Declaration. It can serve as a basic screening tool for workers who do not have background knowledge about the relationship between their cerebrovascular disease and work-related factors, and for workers who are curious about whether their work-related information can be recognized as an occupational disease. And because the Korean-friendly polyglot-ko model was used as the base model, it can be useful for Korean workers. However, it is difficult to make judgments about various cases because various factors other than business factors have not been learned. In addition, it is currently only possible to determine recognition/disapproval of occupational diseases, and it is difficult to suggest solutions such as what health management measures or procedures are needed in the future. Additional research or development steps are needed to determine what intervention factors, such as health management plans or what health behaviors to take in the future, can be given to people who are later recognized as having an occupational disease.

**Conclusion:** This study developed an AI model that can determine the relationship between cerebrovascular diseases, working hours, work factors, and whether occupational diseases are recognized. The complex learning model showed the highest performance by learning expert knowledge and actual occupational disease diagnosis cases, but it had limitations such as lack of judgment for various cases and difficulty in providing future health management plans. Therefore, after improving the performance of the model, it is judged that it is necessary to develop a webUI in the form of a chatbot so that actual workers can use it. To this end, it is judged that it is necessary to supplement with a template-based model that will reflect the detailed information of workers.

---

Key words: Cerebrovascular disease, Occupational Disease, Artificial Intelligence, Large Language Model