

## SK네트웍스 Family AI과정 3기

# 데이터 전처리 인공지능 학습 결과서

### □ 개요

- 산출물 단계 : 데이터 전처리
- 평가 산출물 : 인공지능 학습 결과서
- 제출 일자 : 2024-12-25
- 깃허브 경로 : <https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN03-FINAL-5Team>
- 작성 팀원 : 김중식

### 질문 생성 모델

#### 상황 예시 및 답변 비교

1. 이력서 키워드 추출
  - python, Docker, Azure Cognitive Services, PyTorch,
2. 질문 생성(kor)
  - Docker를 사용하여 애플리케이션을 컨테이너화할 때 고려해야 할 주요 요소는 무엇이며, 이를 통해 어떤 문제를 해결했는지 설명해 주세요.
3. 질문 생성(eng)
  - What are the key factors to consider when containerizing an application using Docker, and what problems have been solved through this process?
4. 모범답안 생성(kor)
  - Docker를 사용하여 애플리케이션을 컨테이너화할 때, 핵심 목표는 애플리케이션의 이식성과 배포 효율성을 향상시키는 것이었습니다. 하지만, 다양한 환경에서의 일관된 동작을 보장하는 것이 큰 도전과제로 작용했습니다. 이를 해결하기 위해 Dockerfile을 작성하고, CI/CD 파이프라인에 통합하여 자동화된 빌드를 구현했습니다. 그 결과, 배포 시간이 50% 단축되었고, 환경 간의 불일치 문제를 최소화하여 안정성을 크게 향상시켰습니다. 이러한 경험은 앞으로의 프로젝트에서도 컨테이너화를 통해 효율성을 극대화할 수 있는 기반이 될 것입니다.
5. 모범답안 생성(eng)
  - When containerizing an application using Docker, the core objective was to enhance the application's portability and deployment efficiency. However, ensuring consistent performance across diverse environments posed a significant challenge. To address this, we created a Dockerfile and integrated it into a CI/CD pipeline to implement automated builds. As a result, deployment times were reduced by 50%, and we minimized issues related to environmental discrepancies, significantly improving stability. This experience will serve as a foundation for maximizing efficiency through containerization in future projects.

## 결론

### 6. 평가 기준 vector db 자료(예시)

- Skill: Docker
- Question: 1. Conclusion
- Answer: DevOps technologies are growing at an exponential pace. As the systems are being more and more distributed, developers have turned towards containerization because of the need to develop software faster and maintain them better. They also aid in easier and faster continuous integration and deployment process which is why these technologies have experienced tremendous growth.

Docker is the most famous and popular tool for achieving the purpose of containerization and continuous integration/development and also for continuous deployment due to its great support for pipelines. With the growing ecosystem, docker has proven itself to be useful to operate on multiple use cases thereby making it all the more exciting to learn it! To build a good DockerFile:

<https://docs.docker.com/engine/reference/builder/>

### 7. 결론

- 시스템의 주요 기능(키워드 추출, 인터뷰 생성, Vector DB 검색, 질문 생성, 데이터 저장)이 모두 정상적으로 작동됨이 확인되었다. 특히, 키워드 추출 등을 바탕으로 사용자 정보를 활용한 맞춤형 인터뷰 설계가 가능함이 입증되었다. 다만 키워드 간의 연관성이 적고 키워드가 단어로 구성되어있는 점을 감안할 때 키워드의 검색쿼리화, 프로젝트 성과반영 등 의 기능 개선이 필요하다. 또한 향후 추가적인 데이터 확보를 통해 질문 다양성 확보가 필요하다.