



연합학습기반 신약개발 가속화 프로젝트

(3 세부) 연합학습 플랫폼 활용 활성화 지원

연합학습 기반 ADMET 예측을 위한 언어-그래프 앙상블 모델 개발

주관기관 : 목암생명과학연구소 (신현진)

공동기관 : 서울대학교 (장문영)

|연구 역량| 연구개발기관 및 연구책임자 소개

신약개발 전문지식을 바탕으로 다양한 인공지능 연구 수행 경험 보유

[주관기관] 목암생명과학연구소



- 목암생명과학연구소는 인공지능을 통한 신약개발을 연구하는 민간 비영리재단법인
- 국내·외 바이오 및 인공지능 분야 인력 구성 및 인공지능 신약개발 관련 내/외부 과제 다수 수행 중
- 서울대병원, KAIST 등 국내 우수 연구기관과 긴밀한 협력 추진을 통해 자체 연구 파이프라인 확대 중

연구책임자 : 신현진 소장

- 글로벌 제약사 등 우수 연구기관 인공지능 연구

참여연구자

- 자담바 에르햄바야르 책임연구원
- 이현수 연구원

연구 수행 분야 및 역량

- 현재 LLM을 적용한 신약개발 내부 연구 수행 중
 - Lipid nanoparticle(LNP) 특성 예측 및 생성을 위한 인공지능 연구
 - mRNA codon 최적화를 위한 인공지능 연구

[공동기관] 서울대학교 공과대학 인공지능 연구실



- 서울대학교 공과대학 인공지능 연구실 (Data Science and AI Lab)
- 바이오 메디컬 응용 분야 최상위 저널에 다수의 논문 발표
- NeurIPS, ICLR, ICML, CVPR 등 세계 최정상급 학술 대회에 지속적으로 우수한 성과 발표

연구책임자 : 장문영 BK부교수

- 다년간 대학 병원에서 환자 진료 및 연구

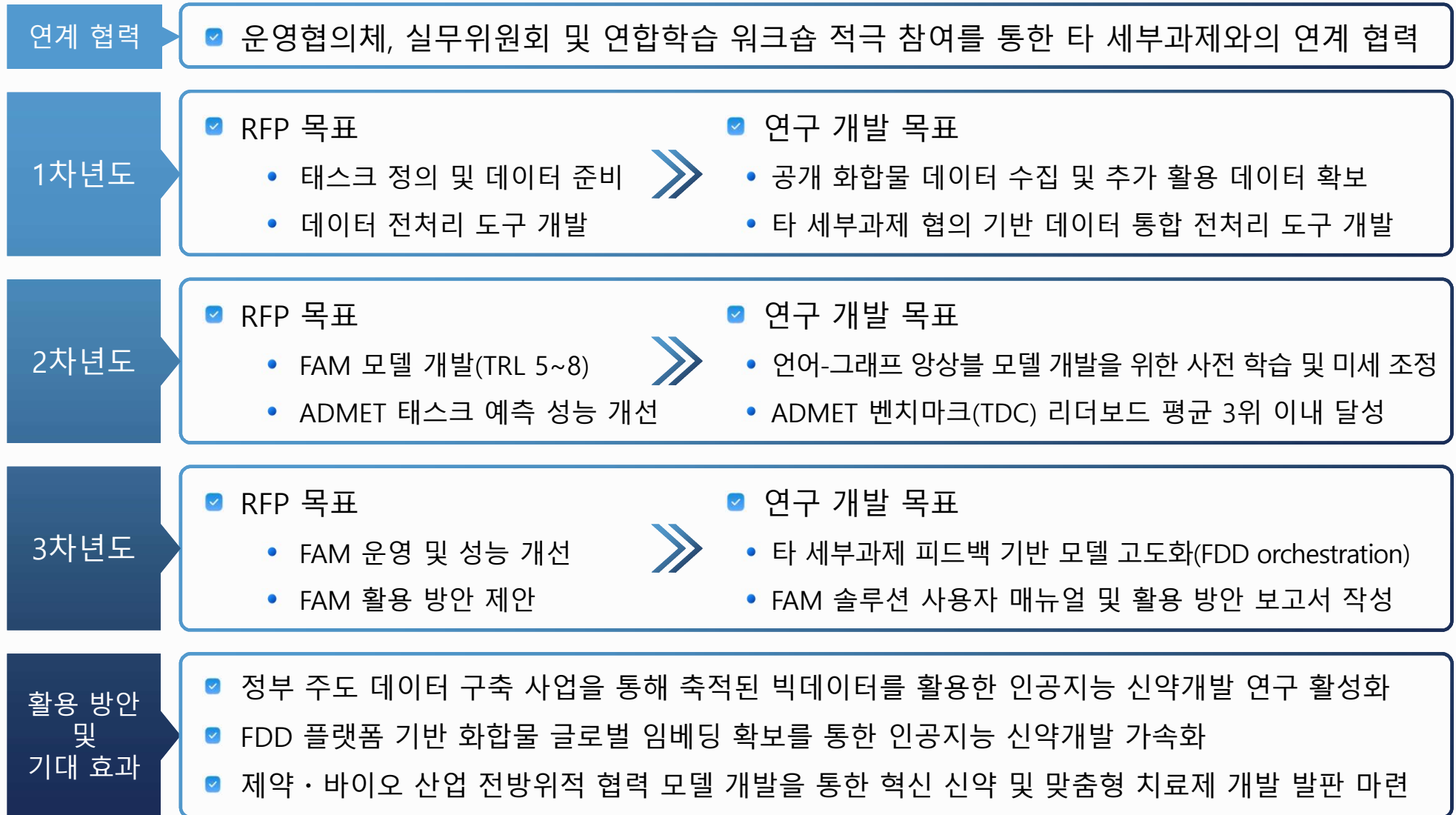
참여연구자

- 이동혁
- 박상권

연구 수행 분야 및 역량

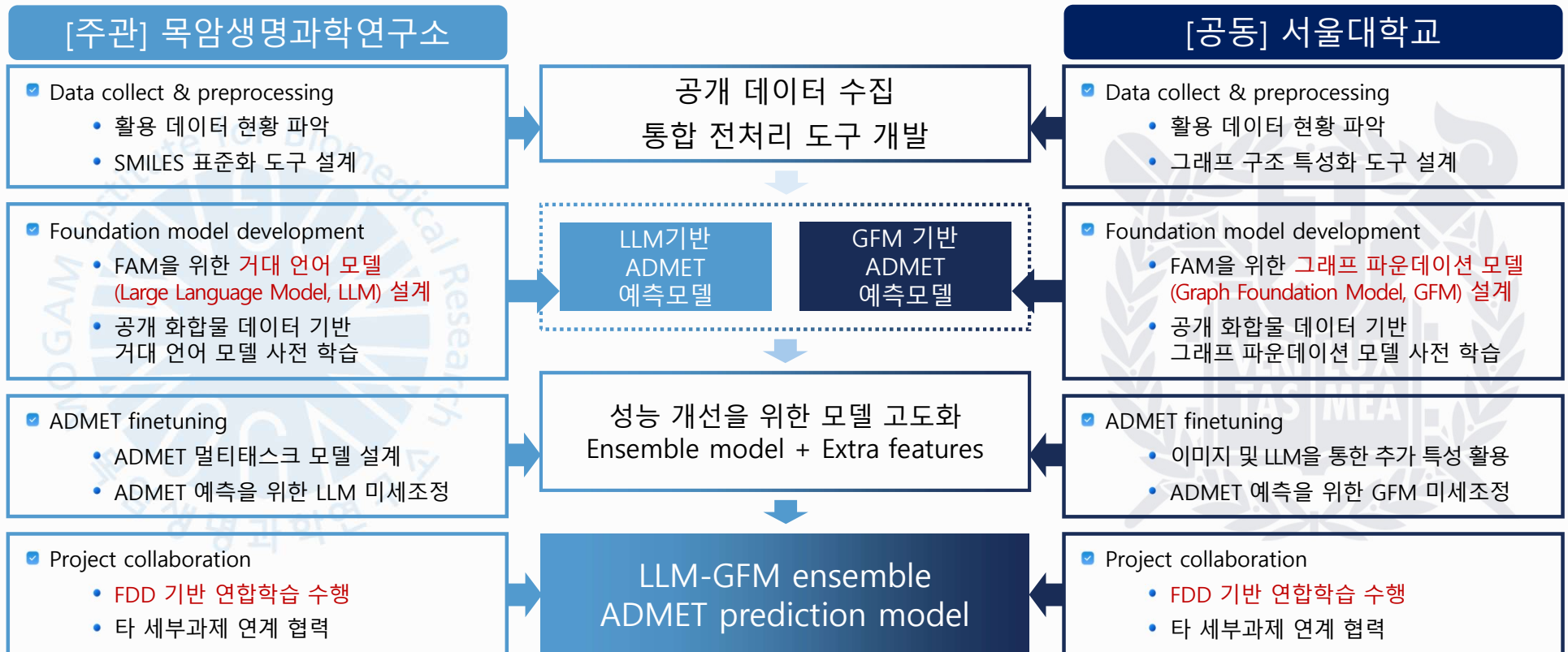
- 현재 GFM을 적용한 신약개발 내부 연구 수행 중
 - Biomedical knowledge graph를 이용한 질환 관련 유전자/단백질 예측을 위한 인공지능 연구

연합학습 기반 ADMET 예측을 위한 언어-그래프 앙상블 모델 개발



연합학습 협의체 기반 병렬 연구 수행

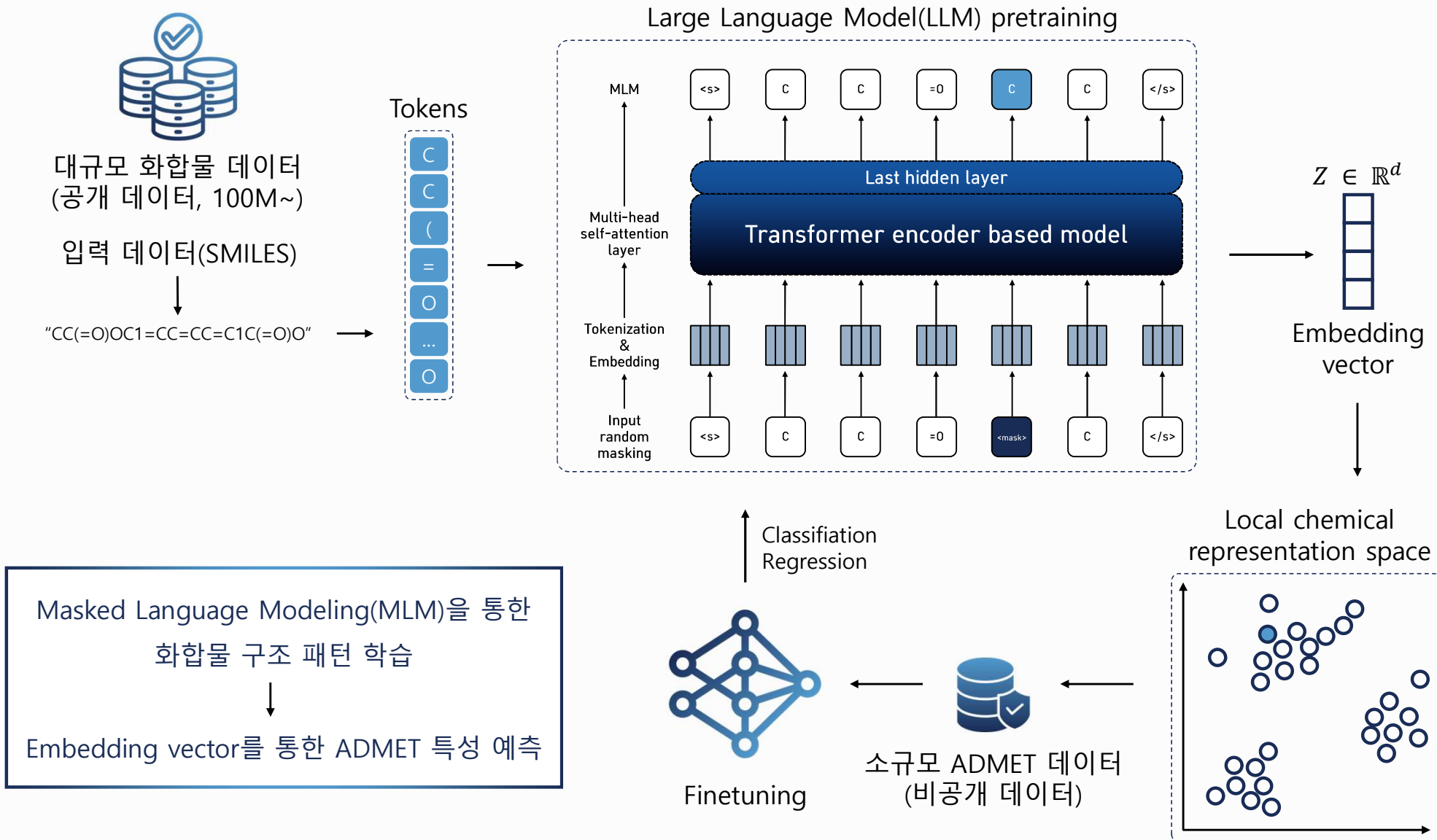
운영협의체 - 실무위원회 - 연합학습 워크숍



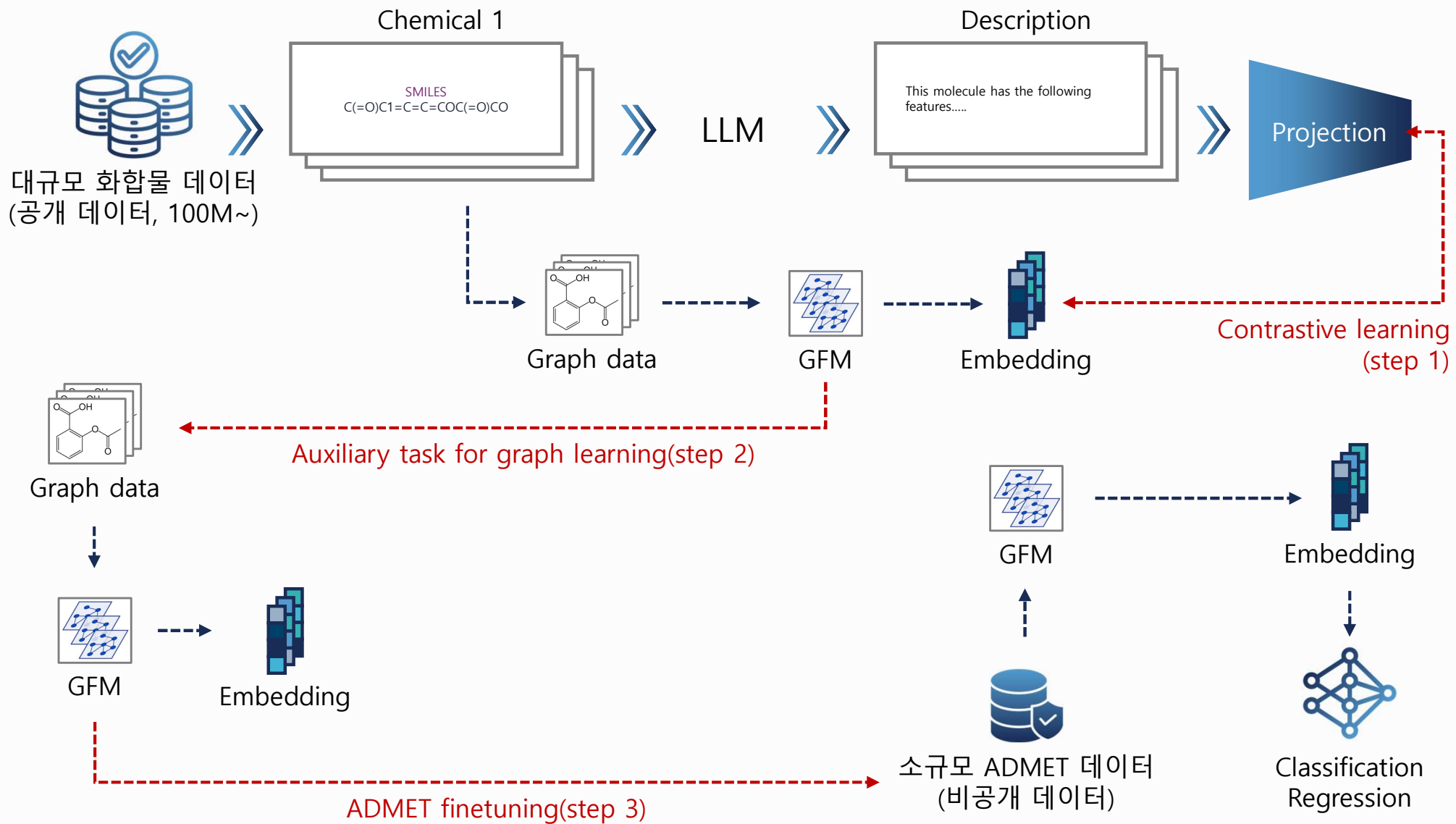
주요 마일스톤

- ✓ (1차년도) LLM/GFM 학습을 위한 공개 화합물 데이터 수집 및 전처리
- ✓ (2차년도) LLM/GFM 사전 학습 및 미세 조정을 통한 FAM 확보
- ✓ (3차년도) FDD 플랫폼 내에서 ADMET 태스크별 연합학습 수행 및 모델 고도화

거대 언어 모델(Large Language Model, LLM) 학습 전략

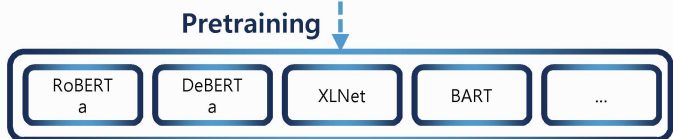
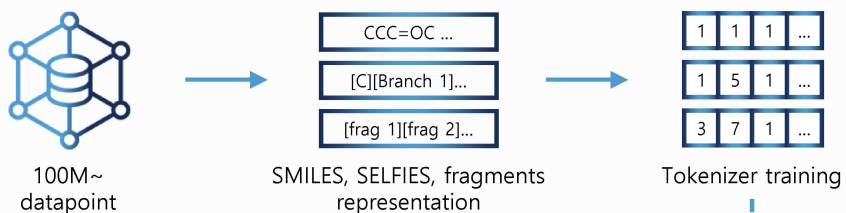


그래프 파운데이션 모델(Graph Foundation Model, GFM) 학습 전략



LLM-GFM 앙상블 전략 수립

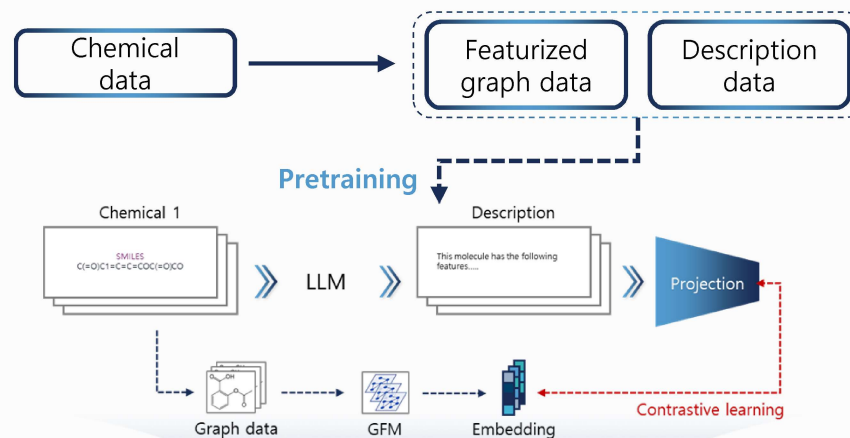
Large Language Model(LLM)



Benchmark task finetuning

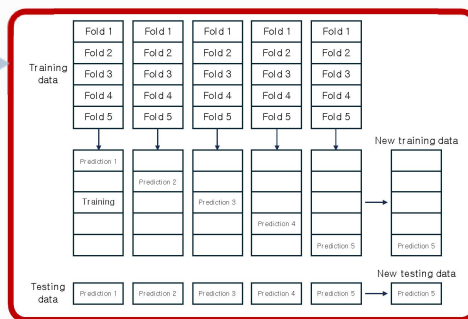
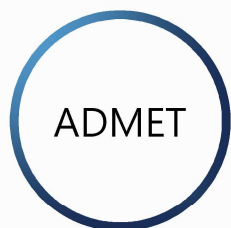
Finetuned LLMs

Graph Foundation Model(GFM)



Auxiliary task finetuning

Finetuned GFMs



Stacking ensemble

Final predictions for ADMET

