

BK 장기연수 결과보고서

연수기간	2021년 12월 30일 ~ 2022년 3월 27일
연수기관	Computational Nanoengineering Lab in The University of Texas
연구주제	머신러닝과 분자 스케일의 양자화학계산을 융합한 생체정보처리 알고리즘 설계

1. 연수활동 내용

연수목적	맞춤형 헬스케어에 위한 생체정보/머신러닝 융합 예측 알고리즘을 개발하고자 한다.
연수내용	<p>텍사스 오스틴 대학의 Computational Nanoengineering Lab와의 공동 협력 연구를 3개월간 진행하였다.</p> <p>3개월간 매주 연구 미팅에 참석하여 다른 학교의 석박사학생들의 연구 분야와 방법을 살펴볼 수 있었으며, 외부 연구원들과 교류할 수 있는 시간을 가졌다. 또한 현재 텍사스 오스틴 대학교 화학공학과에서 개설된 Molecular Modeling method and applications를 들으며 분자 모델링에 대한 이론적인 공부 또한 진행하였다. 분기별 연수내용은 다음과 같다.</p> <p>1개월차에는 LAMMPS를 활용한 분자 동역학 분석법을 익혔다. LAMMPS는 주어진 통계적 앙상블 내에서 뉴턴의 고전 운동 방정식을 이용한 수치적분을 통해 시스템에 포함된 수많은 원자들의 궤적을 생성하는 계산 프로그램이며, 고분자 기반 물질의 분자 구조와 물리적/화학적 특성을 계산하는 데 사용된다.</p> <p>2개월차에는 양자시뮬레이션과 머신러닝을 활용한 보건 의료분야 연구의 최신동향을 살펴보았다. 시뮬레이션과 머신러닝의 특성을 활용하여 알고리즘을 설계할 방안을 알아보았다.</p> <p>3개월차에는 양자시뮬레이션과 머신러닝 실습을 진행했다. 레이블된 데이터를 가지고 지도학습을 수행하는 머신러닝 코드를 작성하고 실행했다. 또한 온도에 따른 Si 산화구조를 계산하는 양자시뮬레이션을 진행했다.</p>
결과 및 시사점	<p>Molecular Modeling method에 대한 이론적인 공부와 실습을 같이 진행하며, 계산화학 분야에 대한 이해도가 깊어졌다. 또한 머신러닝 방법을 통해 데이터를 처리하는 방법을 익힐 수 있었다.</p> <p>미국은 양자 컴퓨팅 시스템 분야에 대해 부 지원 확대를 통해 양자정보과학 분야 신기술개발을 지원하고, 기술적 난제 해결 연구에 예산 우선순위를 둘 만큼, 이 분야에 대해 투자도 많이하고 이미 선구연구도 많이 진행되었다.</p> <p>이번 해외연수를 통해 이러한 해외 연구동향을 경험하고 기술교류 및 네트워킹 기회를 마련했다. 세계 수준의 경쟁력을 갖춘 우수한 글로벌인재가 되기 위한 국제적 안목을 지닌 전문가적 자질을 기를 수 있었다.</p>
향후 연구에 대한 적용방안	양자시뮬레이션과 머신러닝을 결합한 알고리즘을 통해 생체물리학 모델링에서부터 신약개발에 이르기까지 다양한 연구를 수행할 수

있다. 질병과 관련된 빅데이터를 수집하고 분석한 후, 인공지능 모델을 적용해 문제를 식별하고 해결할 수 있을 것이다. 건강관리부터 질병 치료까지 보건의료 전 영역에서 개인맞춤형 헬스케어 솔루션 개발 연구를 수행할 수 있도록 할 계획이다.

2. 일자별 활동내역(구체적으로 기술 요망)

년 월 일	연 수 내 용	비고
2021.12.30	출국	
2022.01.05	랩미팅. 텍사스 오스틴 대학의 Computational Nanoengineering Lab 과의 공동협력 연구 전, 이 연구실에서 진행해 왔고, 현재 진행중인 연구 주제에 대한 짧은 발표를 진행하였다.	
2022.01.05 -01.19	Molecular Dynamics의 계산 프로그램인 Lammmps의 기본 계산원리와 기본 계산법에 대해 배웠다. MD란 주어진 통계적 앙상블 내에서 뉴턴의 고전 운동 방정식을 이용한 수치적분을 통해 시스템에 포함된 수많은 원자들의 궤적을 생성하는 계산 기법이다.	
2022.01.12-	연구실 구경 및 출근. 코로나 상황과 기상악화 기간이 겹쳐 오랜만에 학교가 개방되었다. 텍사스대학 화학공학과 건물에 위치한 연구실로 출근하며 연구를 진행했다.	
2022.01.12.-	정기랩미팅(모두참여). 랩실 학생 모두가 참여하는 정기랩미팅은 2주에 한번씩 진행되었으며, 2주동안 연구한 내용을 요약 발표하는 시간을 가졌다. 이를 통해 Computational Nanoengineering Lab에서 진행되는 연구에 대한 설명을 들을 수 있었다.	
2022.01.19	랩미팅. 한국에서 진행하던 연구 주제에 대해 발표하였고, 이번에 진행할 새로운 프로젝트에 대한 공동연구 계획을 수립하였다. 한국에서는 VASP 프로그램을 활용한 dft 계산에 집중된 연구였다면, 이번에는 분자스케일 수준의 양자화학계산법을 익히는 방향으로 진행하기로 하였다.	
2022.01.20-	LAMMPS 기본 계산 진행. MD 코드는 고분자 기반 물질의 분자 구조와 물리적/화학적 특성을 계산하는 데 사용된다. ethanol, methanol처럼 기본 분자 시스템의 INPUT file들을 직접 코딩해보고 피드백을 받는 시간을 가졌다.	
2022.01.26.- 02.02	LAMMPS 기본 계산 진행. ethanol, methanol처럼 기본 분자 시스템의 물리적/화학적 특성을 직접 계산해보고 실험값과 비교해보았다. 또한, MD 시뮬레이션 동안 원자 운동과 상호작용을 조절하는 원자간 잠재력 (interatomic potential)의 선택에 따라 결과값의 차이도 비교해보았다.	
2022.02.02.-	Molecular Modeling method and applications 수업청강. 양자화학계산분야에서는 프로그램을 돌리는 것보다 어떤	

	<p>파라미터를 어떤 방법으로 돌려야 실제 물리화학적 시스템을 잘 구현할 수 있는지 알아내는 것이 훨씬 중요하다. 따라서 이론적인 부분을 공부해야할 필요성이 느껴져, 텍사스 대학 화학공학과에서 진행되는 Molecular Modeling method and applications 수업을 청강 신청하고 귀국하기 전까지 청강했다.</p>	
<p>2022.02.02.- 2022.02.14</p>	<p>LAMMPS 프로그램 활용. MD는 parameter와 force field 선택에 따라 계산결과가 많이 달라진다. force field와 parameter의 물리적인 의미를 이해하고, LiTFSI을 계산한 연구 논문을 참조하여 계산해보고 비교했다.</p>	
<p>2022.02.09</p>	<p>양자 시뮬레이션을 활용한 보건의료분야의 최신 연구 동향들을 살펴보았다. monte-carlo simulation은 불확실성이 높은 모델에서 무작위 추출된 난수를 이용하여 원하는 함수의 값을 계산하기 위한 시뮬레이션 방법인데, 현재 이러한 방법은 흔히 수학이나 물리학에서 자주 사용되었다. 컴퓨터 성능이 발전하면서 이를 활용하는 분야가 많아졌고, 보건 의료분야에서는 헬스케어 서비스에 대한 평가나 예측을 하는 데 사용이 되는 것으로 알려져 있다.</p>	
<p>2022.02.14.- 02.28</p>	<p>LAMMPS 프로그램 활용. moltemplate script를 사용하여 bulk system을 만들고, npt, nvt 과정을 거쳐 물질 sampling을 했다. time step과 온도, 압력 조건을 변화시키며 계산을 진행했고, 적절한 조건을 설정하는 방법을 익혔다.</p>	
<p>2022.02.16</p>	<p>머신러닝을 활용한 보건의료분야의 최신 연구 동향들을 살펴보았다. 머신러닝 알고리즘은 대규모 데이터 세트에서 패턴과 상관관계를 찾고 분석을 토대로 최적의 의사결정과 예측을 수행하는 데이터 분석방법이다. 현재는 질병 및 치료 예측, 약물 개발 지원 등 다양한 분야에서 시험적으로 사용되고 있다.</p>	
<p>2022.02.21</p>	<p>랩미팅. 시뮬레이션 방법과 머신러닝을 결합할 방법에 대해 회의를 진행했다. 시뮬레이션과 머신러닝을 연결하는 알고리즘, 정확도를 높이는 방법에 대해 고안해보았다.</p>	
<p>2022.02.23</p>	<p>monte-carlo 개념에 대해서 배웠다. random number generator, importance sampling 등을 확률 통계방식으로 바라보며, 수식과 이를 구현하는 코딩에 대해 배웠다.</p>	
<p>2022.03.02</p>	<p>머신러닝 알고리즘과 시스템의 종류에 대해 알아보았다. 머신러닝은 크게 지도학습, 비지도학습, 강화학습으로 나뉘어 지는데, 레이블된 데이터를 가지고 지도학습을 수행하는 머신러닝 코드를 작성하고 실행해보았다.</p>	
<p>2022.03.09</p>	<p>Monte-carlo method를 구성하는 알고리즘에 대해 배웠다. selection, expansion, simulation, backpropagation 과정에 대해</p>	

	이해했다.	
2022.03.09.- 2022.03.28	monte-carlo simulation 실습. MD simulation으로 계산할 시 시간이 많이 걸리거나, random한 input으로 계산하는 게 더 유리할 경우에는 mc simulation을 쓴다. Si의 온도에 따른 산화구조를 mc method를 이용하여 계산해보았다.	
2022.03.29	귀국	

위와 같이 해외연수 결과보고서를 제출합니다.

2022년 4 월 21일

장기연수자 채소민 채소민(인)
지도교수 이상현 이상현(인)