

ScAlence

1. 프로젝트를 소개해주세요!

우리 scAlence 팀은 Graph Attention Network를 기반으로 화합물의 쓴맛 여부를 예측하는 모델을 만드는 프로젝트를 진행했습니다. 화합물의 고유한 특성을 분석하여 화합물의 어떤 부분이 쓴맛을 내도록 하는지 heatmap으로 나타내었습니다. GUI를 활용해서 AI나 화학을 잘 모르는 분들도 쉽게 사용할 수 있도록 했습니다.

2. 프로젝트 하면서 어떤 문제를 겪었나요?

예상했던 성능보다 성능이 낮게 나오는 문제가 있었는데, Hyperparameter tuning을 통해서 성능을 기존의 모델보다 향상시키는 것을 성공했습니다.

Github 사용이 익숙하지 않아서 코드 공유에 어려움이 있었습니다. Github flow 전략을 사용했는데 그 흐름이 잘 와닿지가 않아서 커밋도 많이 꼬이고 사용에 어려움이 많았던 것 같습니다. 그래서 git 공부를 더 해야겠다는 생각을 했습니다...

3. 프로젝트 하기 전후 달라진 점이 있다면?

Attention 기법에 대한 이해도가 더 높아졌으며, pytorch 기반의 모델 아키텍처에 익숙해질 수 있었습니다. Hyperparameter 최적화를 직접 해본 것은 처음인데, 모델의 성능에 큰 영향을 미친다는 것을 알게 되었습니다.

4. 프로젝트를 시작하는 팀에게 전해줄 꿀팁을 말해주세요!

본인이 흥미있는 주제를 잡는 것이 중요한 것 같습니다. 그래야 평소에도 해당 주제에 대해 끊임없이 생각하고 해결책을 생각해낼 수 있습니다.

또, 본인이 알고 있다고 생각한 것을 실제로 구현해보는 경험을 해보는 것이 중요한 것 같습니다. 아이디어에서 그치지 않고 구현하여 작동하기까지 막히는 부분이 많기 때문에 이러한 부분에 신경쓰며 진행하면 더 퀄리티 높은 프로젝트가 될 수 있을 것 같습니다.

5. 개발자로서의 역량 향상을 위해 어떤 노력을 기울였으며, 이 프로젝트를 통해 어떤 기술적 성장을 이루었나요?

화합물의 특징 데이터를 학습 데이터로 사용해본 적이 없었기 때문에 배경 지식부터 학습해야했습니다. 이 과정에서 학습 데이터에 대한 배경지식이 많이 늘었고, Preprocessing 기술도 성장한 것 같습니다. 또한 Attention 기법을 수식을 통해서만 알고 있었지만 실제로 구현해본적은 없었기 때문에 이번 기회를 통해서 기존에 있던 AttentiveFP 모델을 분석해가면서 해당 모델의 구조를 더 상세하게 알아볼 수 있었고, 수정하면서 어떻게 작동하는지 더 자세하게 공부할 수 있었던 것 같습니다.

6. 프로젝트 개발 중 어려움을 겪은 경험이 있나요? 어떻게 해결했으며, 그 과정에서 얻은 교훈은 무엇인가요?

AI 분야 특성상 분업이 어렵습니다. 궁극적으로 하나의 결과를 내는 모델을 만들어야 하고, 이를 위한 세부 모델을 개발해야 하는데, 모델마다 개발 난이도가 상이합니다. 이번 프로젝트에서는 건형님이 난이도가 높은 graph attention network 개발을 맡아주셨고, 저는 비교적 쉬운 난이도인 multi-layer perceptron 부분을 담당했습니다. 난이도가 쉬우니 개발 속도에 차이가 있었고, 제 부분을 완료한 다음 붓는 시간이 있었습니다. 건형님이 담당한 모델을 함께 개발하는 것은 효율적이지 않아서, 그 남은 시간을 프로젝트 개발에 활용하지 못했습니다.

지금 생각해보면 그 시간 동안 MLP 단일 모델에 대해서 최적화를 수행할 수도 있었을 것 같고, 모델 성능 향상을 위한 다른 메소드를 탐색할 수도 있었을 것 같습니다. 하나의 세부 목표를 완료하였으나 그 다음 목표를 정해두지 않아서 생긴 문제였던 것 같아요.

그래서 다른 개발자 분들과 동일하게 AI 개발자에게도 개발 전 세부 목표와 계획 수립이 중요하다는 것을 느끼게 되었습니다. 프로젝트 중 계획이 변경될 수 있다고 하더라도, 미리 세부 목표를 설정해야겠다는 생각을 하게 됐습니다.