

## 2-1. T4056\_김찬호

- 나는 내 학습목표를 달성하기 위해 무엇을 어떻게 했는가?

- 우리 팀과 나의 학습목표는 무엇이었나?

우리 팀의 학습목표는 노션과 깃허브와 같은 툴을 보다 적극적으로 활용하여 원활하게 협업하는 것이었고, 나의 학습목표는 AI 모델을 직접 구현해보는 것과 데이터 기반의 접근방법을 적용하는 것이었다.

- 개인학습측면

나의 학습목표를 달성하기 위해, 가장 먼저 주어진 베이스라인들을 확인하고 인퍼런스와 같은 필요한 기능들을 추가했다. 이후에 좋은 성능을 보여준 AE 계열의 모델들을 탐색하고 EASE 등의 모델들을 구현했다. 데이터 측면에서는 다양한 EDA와 Feature Engineering을 진행하지 않아 아쉬움이 남는다.

- 공동학습측면

우리 팀은 팀 차원에서의 학습목표를 달성하기 위해, 먼저 github를 어떻게 활용할지 설정했다. github commit convention과 issue 활용방법, github 코드 관리방법 등을 토의를 통해 정하여 이번 프로젝트를 진행하면서 깃허브 충돌 문제를 겪지 않았고, 다른 팀원들이 작업 중인 내용을 보다 쉽게 들여다 볼 수 있었다.

- 나는 어떤 방식으로 모델을 개선했는가?

- 사용한 지식과 기술

부스트캠프 내부 경진대회 중 처음으로 경험하는 랭킹 문제이기 때문에 여러모로 어려움을 겪었다. 가장 먼저 인퍼런스를 진행하는 방법이 prediction 값이나 classification 결과를 submit으로 제출했던 이전 대회들과 달리 모델의 출력값을 활용하여 아이템의 랭킹을 매겨서 submit을 생성하는 식으로 인퍼런스를 구현했다. 또, 시퀀셜 모델의 경우 시퀀스 데이터를 수정하여 유저간 데이터 편차를 줄이고 가능한 많은 데이터를 활용할 수 있도록 하는 것이 모델의 성능 향상에 큰 영향을 끼친다는 지난 dkt 대회의 경험을 토대로 시퀀스 데이터를 수정하여 모델 성능 향상을 꾀했다.

- 내가 한 행동의 결과로 어떤 지점을 달성하고, 어떠한 깨달음을 얻었는가?

각각의 모델에 대해서 개별적으로 인퍼런스 등의 필요한 기능을 구현하면서, 여러 모델을 실험하기 위해서는 동일한 기능을 가지는 부분을 블록화하여 재사용할 수 있도록 구성하여 리소스를 줄이는 시스템이 필요하다는 생각이 들었다. 또, 명확한 논리의 흐름에 따라 프로젝트를 진행하지 않으면 중간 단계에서 방향을 잃을 위험성이 있다는 것을 느꼈다.

- 전과 비교해서, 내가 새롭게 시도한 변화는 무엇이고, 어떤 효과가 있었는가?

깃허브 관리방식을 변경했고, 모델의 인퍼런스 등 필요한 기능들을 나아서 구현했다. 깃허브 충돌을 방지했고, 협업을 위한 툴을 마련할 수 있었다.

- **마주한 한계는 무엇이며, 아쉬웠던 점은 무엇인가?**

각각의 모델에서 단일 모델 성능을 보다 높일 수 있는 요소들이 있었음에도 불구하고 제대로 적용하지 못한 점이 아쉽다. 또, 앙상블의 경우 보다 전략적으로 접근했으면 더 많은 가설들을 검증하고 좋은 결과를 낼 수 있었을 것 같다.

- **한계/교훈을 바탕으로 다음 프로젝트에서 스스로 새롭게 시도해볼 것은 무엇일까?**

가장 먼저 시도해보고 싶은 것은 RecBole과 같은 라이브러리를 활용하여 최소한의 리소스로 여러 모델을 실험할 수 있는 베이스라인을 구축하는 것이다. 또, 문제 해결 과정에서 논리의 흐름대로 프로젝트의 파이프라인을 구성하여 놓치는 부분 없이 태스크를 완수하고 싶다. 추가로 깃허브 외에도 다양한 툴과 방법을 활용하여 협업적인 부분도 발전시킬 것이다.