



김성연 개인회고

• 이번 프로젝트에서 나의 목표는 무엇이었는가?

- 1) 새로운 팀원들과 함께 협업이란 어떤 것인지 익히는 시간을 가지고 싶었습니다.
- 2) 딥러닝 엔지니어링 실력을 늘리고 싶었습니다.
- 3) 저번 대회와 마찬가지로 높은 순위를 기록하고 싶었습니다.

• 나는 내 학습목표를 달성하기 위해 무엇을 어떻게 했는가?

우선 팀원들과 함께 어떻게 협업하면 좋을지 틀을 정했습니다. 깃 브랜치 전략, 공부한 것을 공유하는 방식, 두 명씩 짝지어서 모델링하는 방식 등을 진행했습니다. 이를 통해 팀으로써 프로젝트를 진행한다는 것을 느낄 수 있었습니다. 그리고 주도적으로 팀원에게 적절하게 역할 배분을 진행하여 분업을 진행하였습니다.

또한 적절한 딥러닝 모델을 직접 설계해보고 구현도 해봤습니다. 파라미터가 변경되는 것을 breakpoint를 걸어 눈으로 확인해보며 그래디언트 디센트 문제를 방지하는 방법을 고민했습니다.

다음으로 높은 순위를 기록하기 위해 대회 초기부터 4주를 어떻게 진행해야 할지 스케줄링을 했고 남들보다 대회를 참가한 경험을 적극적으로 공유하여 팀원 모두가 성과를 낼 수 있도록 도와주었습니다. 그 결과 1등은 못했지만 public, private 모두 2등인 robust한 모델을 설계할 수 있었습니다.

• 나는 어떤 방식으로 모델을 개선했는가?

우선 데이터가 시계열적인 요소가 있고 개수도 적기 때문에 valid set을 test set과 유사한 환경으로 정의하는 것이 중요했고 이 부분을 많이 신경썼습니다. valid와 test가 score가 다소 차이는 나지만 경향성을 따라가는 것을 확인했고 단순히 public 리더보드 수치가 높다고 좋은 모델이 아니고 valid score가 같이 좋아야한다는 것을 많이 신경썼습니다.

베이스라인에 있는 시퀀셜 모델이 가지는 문제점을 빠르게 파악하고 정형 데이터 모델에 성능이 뛰어난 부스팅 모델을 빠르게 채용하였습니다. 부스팅 모델이 어느정도 성능을 보장해주고 있기 때문에 시간적인 여유를 가지고 다양한 시도를 했습니다.

Optuna 툴을 소개하고 팀원들과 같이 모델 고도화 하는 시간을 가졌던 것도 좋았습니다.

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

팀원과 많이 소통하려고 노력했습니다. 처음에 많이 어색했지만 내가 느낀 것을 적극적으로 공유하고 팀원들에게도 계속 요구했습니다. 초반부에는 잘 안됐지만 진행할 수록 팀원들이 본인의 것을 공유하는데 부끄러움도 사라지고 적극적인 공유가 잘 이뤄진 것 같습니다. 완벽하진 않았지만 서로가 어떤 일을 진행하는지 알 정도는 되었습니다. 더 나아가서 깃허브 레포 내 프로젝트 톨을 이용해 하고있는 일을 기록하였는데 서로를 이해하는데 도움이 많이 되었습니다. 또 팀장으로써 팀원에게 역할 부여를 하는 경험도 새로 해보았는데 어떤 일을 해야 할 지 고민되는 팀원에게 길잡이가 된 것 같아 팀에 긍정적인 효과를 낸 것 같습니다.

• 마주한 한계는 무엇이며, 아쉬웠던 점은 무엇인가?

코드단위의 협업이 되지 못한 것이 아쉽습니다. 타인이 알아볼 수 있는 코드, 다른 팀원의 코드를 읽어보려는 노력이 많이 부족했습니다. 처음에는 괜찮았으나 점점 서로의 코드를 이용하지 못하는 한계가 생겼습니다.

딥러닝 모델링을 생각만큼 많이하지 못한 것도 아쉽습니다. 아예 안한것은 아니나 1등팀의 발표를 들어보니 row level에서 파이토치로 트랜스포머등을 구현했던 것을 보며 성과가 나오지 않고 시간이 없다는 이유로 미뤘던 것이 후회됩니다. 처음 목표와 다르게 등수를 위해 내가 잘해왔던 것만 진행했던 것 같아요.

그리고 나만의 세계에 갇힌 것이 제일 아쉽습니다. 통계학과 출신이기 때문에 다른 사람보단 이론적인 지식이 조금 있습니다. 하지만 자기만의 논리에 갇혀서 확실하지 않은 내용에 대해 다른 레퍼런스를 찾아보거나 조교님나 멘토님에게 물어보지 않고 혼자 결정지은 부분이 너무나도 아쉽습니다. 구체적인 예시를 들면 AUC 평가 메트릭에서 앙상블을 시도할 때 AUC가 랭킹을 잘 매기는 것이라는건 잘 알고 있었는데 이를 해결하는 방안이 명확한 증명이나 레퍼런스 없이 본인의 직관만을 이용해 0-1 Min-Max scaler을 적용한 것 입니다. 아예 말이지 되지 않는 것은 아니나 마스터클래스에서 마스터님이 더 좋은 방안이 있다고 했을 때 정신이 번쩍 들었습니다. 학습 하는데 중요한 결점이라고 생각합니다.

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

실제로 같이 일하는 팀원들이다 생각하고 모든 코드를 누군가 읽는다고 생각하고 쓰는 것을 시도할 것 입니다. 주석도 달고 변수명도 신경쓰고 원하는 변경사항이 있을 때 어려움 없이 자기가 쓴 코드인 것 처럼 생각들게 하는 것이 목표입니다. 이번 프로젝트가 협업의 연습이었다면 다음 프로젝트는 협업의 완성이 되고 싶습니다.

다음으로 등수에 너무 연연하지 않고 완성도 있는 딥러닝 모델을 하나 꼭 만들고 싶습니다. 이 과정에서 자기의 주관을 가지는 것도 중요하지만 이건 이미 잘 하고 있고 대회와 맞는 좋은 레퍼런스를 찾는 노력을 할 것 입니다.