

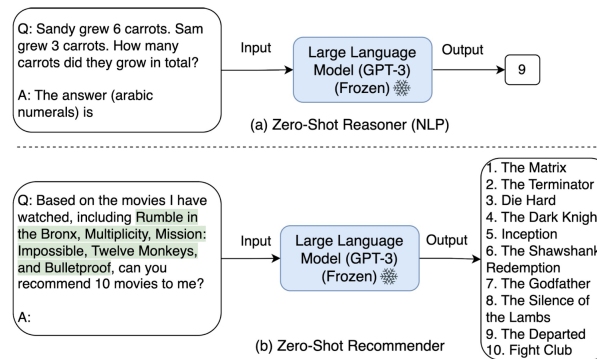
## Motivation

요즘 연구에서 LLMs은 다양한 NLP 문제들에 대해 추가적인 학습 또는 추가적인 training sample 없이 즉, zero-shot setting에서 task-specific prompts를 통하여, 이 문제들을 해결하는 것에 대해 높은 성능을 보여주고 있다. 과연 zero-shot setting에서 LLMs이 recommender로 이와 같은 좋은 성능을 보여줄 수 있을까? Zero-shot setting이 중요할 수 있는 이유는, 모델이 매우 커서 training이 어려운 경우가 존재하거나, 모델이 api로만 접근될 수 있는 경우가 존재하기 때문이다.

## 1. Challenges

저자들은 zero-shot setting에서 LLMs의 NIR 성능을 향상시키려는 연구를 진행했고, LLMs이 recommender로 동작할 수 있도록, 해결해야할(address) 2가지 major challenges를 정의(identify)했다.

1. The recommendation space can be extremely large for LLMs,
2. and LLMs do not know about the target user's past interacted items and preferences.



**Figure 1: Example inputs and outputs of GPT-3 with zero-shot prompting for a NLP task and a recommendation task.**

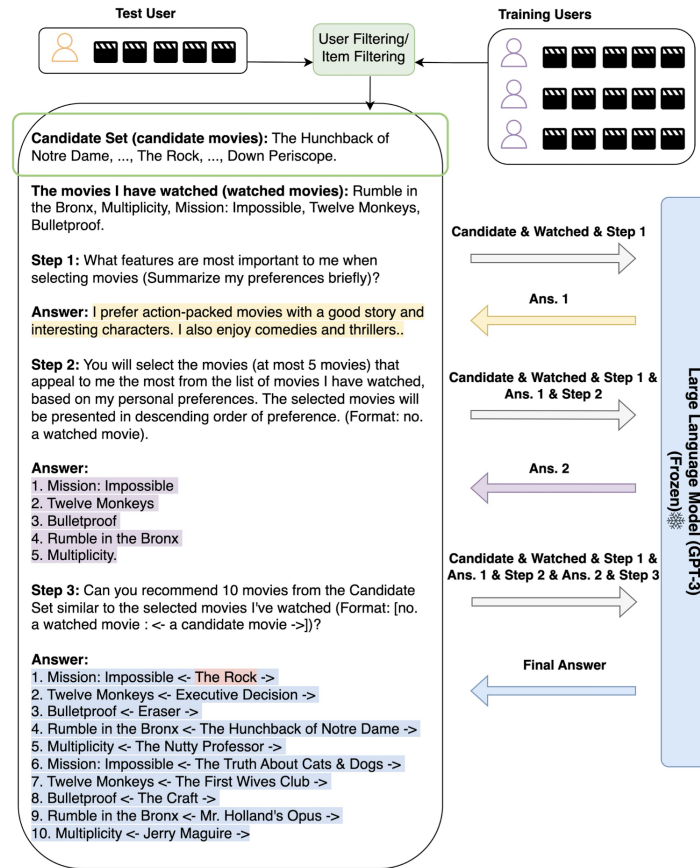
LLMs은 추천되어야 할 universe of items을 모두 알아야 한다.

## 2. Solutions

논문에서는 zero-shot setting에서 NIR 성능을 올리기 위해 아래의 3가지 methodology를 활용한다.

1. Recommendation space를 줄이기 위해, user 또는 item filtering을 활용한 candidate item set을 생성한다. (External module)
2. GPT-3가 3가지 subtasks를 수행할 수 있도록 가이드 하는 3-step prompting 기법을 적용한다.
  - (a) user preference를 capture 하는 것,
  - (b) user가 과거에 시청했던 영화중에서 representative를 select 하는 것,
  - (c) 10개의 영화 리스트를 recommend 하는 것
3. 쉽게 결과를 extract 하기 위하여, formatting technique 사용

## Methodology



**Figure 2: Zero-Shot NIR prompts. The ground truth movie (i.e., The Rock) has been highlighted in red.**

Zero-shot setting에서는 추가적인 training examples를 제공할 수 없기 때문에, recommendation space에 대한 정보를 prompt로 제공해야 한다. Recommendation space 정보를 모두 제공할 필요 없이, target user에 적합한 candidate item set으로 recommendation space를 구성하여 prompt에 제공하는 방법이 있다. 모델 구조를 살펴보면, candidate set을 prompting 하기 때문에, item의 수가 많으면 안된다.

In zero-shot setting of MovieLens 100K dataset, recommendation space를 구성하는 items은 1,683개이며 prompt에 사용하기에는 매우 큰 숫자이다. 따라서 item filtering과 user filtering을 활용하여 target user에 대한 candidate set을 추출한다.

- User filtering : User의 interacted items multi-hot vector를 이용하여, target user와 매우 비슷한 users를 찾는다. 매우 비슷한 users의 interacted items 중 most popular items를 선택한다.
- Item filtering : Item의 interacted users multi-hot vector를 이용하여, target user interacted items와 매우 비슷한 items를 찾는다. 매우 비슷한 items 중 most popular items를 선택한다.

## Weakness

이 논문에서 이야기하고자 하는 결론은, LLMs을 NIR에 활용할 때, 적절한 prompting strategy를 사용한다면 zero-shot setting에서 꽤 좋은(reasonable) 성능을 낼 수 있다는 것이다.

Target user의 interaction history를 기반으로, external module을 이용하여 얻은 candidate set을 LLMs이 reranking 하는 methodology 이다. LLMs 자체의 성능이 아닌, external module 즉 candidate set에 매우 큰 종속성을 가지게 된다. Candidate set 자체에 preferred next item이 존재하지 않는다면? LLMs의 reranking은 아무런 효과가 없다.