

## 졸업연구 결과 분석 및 결론

앞서 사용했던 3개의 모델들 중 가장 결과가 좋았던 KR-SBERT-V40K-klueNLI-augSTS 모델의 결과를 분석했습니다. 앞서 수행한 K-medoids를 통해 분류된 클러스터들의 특징을 확인하기 위해 stratified random sampling 방법을 사용하였습니다. 18468개의 데이터를 신뢰도 99%, 오차범위 5%로 나타내기 위해서는 641개의 샘플이 필요합니다. K-medoids 클러스터링의 대표 샘플을 추출할 때 전부 무작위로 하는 것보다 Cluster들의 medoid들을 포함한 경우 더 잘 대표되는 것을 확인할 수 있습니다.[1] 이 때문에 각 클러스터들에서 대표되는 값을 선정할 때 클러스터의 크기에 비례한 개수만큼 선정하되 반드시 medoid를 포함하도록 하였습니다.

이렇게 선정한 641개의 샘플을 네이버에서 제공하는 Clova Sentiment를 사용해 긍정, 부정, 중립의 감정 수치를 확인하고, 리뷰 내에서 문장의 감정이 어떻게 변화하는지 확인하였습니다. **[Figure 1]**과 같이 리뷰의 전체적인 감성이 어떤지, 그리고 각 문장이 얼마나 긍정적인지, 부정적인지, 혹은 중립적인지 확인할 수 있습니다. 이 경우, 긍정과 부정이 반반 섞여있고 부정적인 문장 이후 긍정적인 문장이 이어집니다.

```
{"document":{"sentiment":"positive","confidence":{"negative":49.999733,"positive":50.00009,"neutral":1.7423878E-4},"sentences":[{"content":"조용조용한 분위기는 아니에요.", "offset":0, "length":16, "sentiment":"negative", "confidence":{"negative":0.99996793, "positive":1.1239794E-5, "neutral":2.0908654E-5}, "highlights":[{"offset":6, "length":9}], {"content":" 하지만 인테리어도 예쁘고 음식도 맛있고 가성비도 좋고 예쁜 꽃도 주시고 전반적으로 만족합니다!", "offset":16, "length":53, "sentiment":"positive", "confidence":{"negative":0.0, "positive":1.0, "neutral":0.0}, "highlights":[{"offset":15, "length":25}]}]}}
```

**[Figure 1]** Example of CLOVA sentiment analysis.

샘플링 된 데이터를 클러스터별로 리뷰 전체의 부정, 긍정, 중립 값의 평균을 구하고, 리뷰의 흐름이 어떻게 변하는지 (긍정적인 문장 다음 긍정적인 문장이 이어지는지, 긍정적인 문장 다음 중립적인 문장이 이어지는지 등)를 계산하였습니다. 그리고 그것들 중 테스트 리뷰에서 눈에 띄는 특징을 가진 4개 클러스터, 3, 24, 27, 30번의 4개 클러스터의 결과를 아래와 같이 확인하였습니다. 3번과 24번 클러스터는 가짜 리뷰로 분류된 클러스터이며, 27번 30번 클러스터는 진짜 리뷰로 분류된 클러스터입니다. 여기에 특징적인 모습을 보이는 10번 클러스터를 더해 **[Table S1]**에서 5개 클러스터의 평균 감정 수치를 확인할 수 있으며, **[Table S2] ~ [Table S6]**에서는 감정 흐름을 확인할 수 있는데, 1~9 값은 긍정-긍정, 긍정-중립, 긍정-부정, 중립-긍정, 중립-중립, 중립-부정, 부정-긍정, 부정-중립, 부정-부정을 나타냅니다.

GPT로 생성한 리뷰의 경우 대체로 긍정적인 리뷰들이 생성되었기 때문에 긍정의 감정이 몹시 강한 것을 알 수 있습니다. 대부분의 경우 가짜 리뷰들이 많이 포함된 클러스터들은 90점 이상의 긍정적인 감정을 보였습니다. 또한 문장의 흐름 역시 가짜로 분류된 클러스터는 긍정-긍정, 긍정-중립, 그리고 중립-긍정이 90% 이상을 차지할 정도로 극단적이지만, 진짜로 분류된 클러스터의 경우 감정의 변화가 훨씬 다채로운 것을 확인할 수 있습니다.

이러한 분석을 통해 각각의 클러스터들을 추가적으로 확인한 결과 진짜로 분류된 클러스터들 중에서도 굉장히 가짜로 분류된 클러스터들과 유사한 특징을 지니는 것들을 확인할 수 있었습니다. 위에서 확인한 것처럼 긍정이 몹시 강하고, 감정의 변화가 긍정-긍정, 긍정-중립, 중립-긍정에 치우친 것들을 진짜인 척 하는 가짜 리뷰들의 클러스터라고 예상해 볼 수 있습니다. 하지만 여전히 애매한 경계선에 있는 리뷰들도 보입니다. 10번 클러스터의 경우 230개의 진짜 리뷰와 17개의 가짜 리뷰로 이루어진 한 쪽으로 몹시 기울어진 클러스터입니다. 하지만 [Table S1]에서 볼 수 있듯이 긍정의 값이 몹시 높아 가짜 리뷰들처럼 보이나 실제로는 [Table S6]에서 볼 수 있듯이 감정 변화가 다채로워 진짜 클러스터인 30번 클러스터의 [Table S5]와 유사한 값을 보입니다. 10번 클러스터의 경우 샘플링 된 리뷰에 *"직원들도 친절 하신편이고 잠발라야 살짝 매콤하면서 새우랑 소시지가 많이 들어가 있고 반숙 터트리니깐 존맛", "인테리어,분위기 좋고요(야외 테이블도 있어서 좋아 보였어요) 전반적으로 맛있는 편이고 양 많고요. 특별함이 있진 않지만 나쁘지 않습니다. 개인적으로 별로였던 점은 메뉴 선불로 주문하는 건데 홀직원과 주방직원 미리 소통은 안하시는 지 이미 주문한 메뉴 안된다하여 취소 후 재 다른 메뉴 재 결재한 것/ 재결재한 메뉴에 감자튀김이 엄청나게 나오는 지 몰랐는데 저희가 이미 감자튀김 올라간 피자를 시켰으니까 좀 센스있는 곳이었다면 중복된다는 멘트 해주셨을 것 같은데 그 정도 서비스가 되는 곳은 아니었다. / 경복궁 근처 가면 한 번 가볼만한 곳 정도입니다. 재방문 의사는X"*와 같이 사람이 판단하기에 진짜 리뷰로 판단할 만한 리뷰가 상당수 존재하였습니다. 이 때문에 리뷰 전체의 감정 수치보다는 리뷰의 감정 변화를 통해 대략적으로나마 진짜와 가짜 리뷰를 분류할 수 있어 보입니다.

## 결론

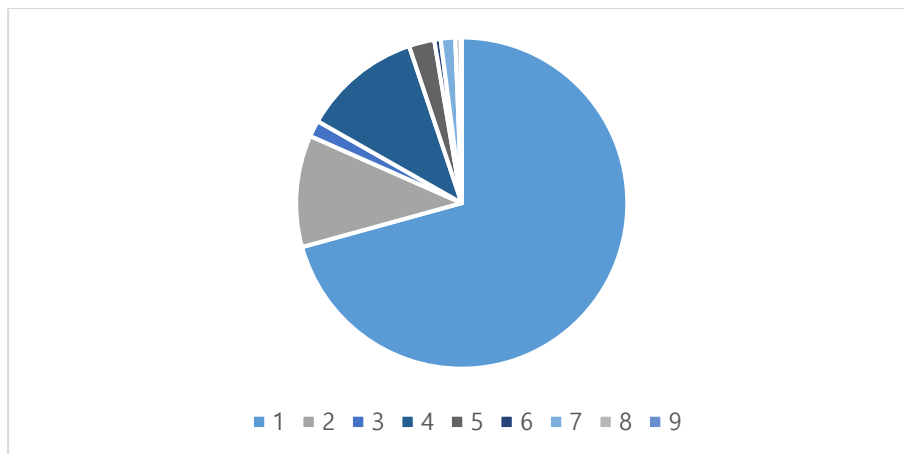
이 중 테스트 리뷰들을 가장 잘 분류한 KR-SBERT-V40K-klueNLI-augSTS 모델의 클러스터링 결과를 바탕으로 추가적인 분석을 수행하였을 때, GPT로 생성한 가짜 리뷰들의 집합은 긍정적인 부분에 몹시 치우친 경향을 보였고, 리뷰에서 감정 변화가 긍정-긍정, 긍정-중립, 그리고 중립-긍정이 대부분으로 나타납니다. 특히, 전체의 60% 가량이 긍정-긍정일 정도로 극단적인 경향을 보였습니다. 여전히 단일 리뷰 하나만을 가지고 진짜인지 가짜인지 확인하는 것은 어려울지라도 리뷰들의 집합을 생성하고, 그 집합이 일정 수준 사이의 긍정적인 모습을 보이며 감정 변화가 다양한 리뷰들의 집합인 경우 진짜 리뷰들의 집합이라고 예상할 수 있어 보입니다.

[1] Lei Li, Yu-Ling Lan, Lei Huang, Congting Ye, Jorge Andrade, and Patrick C Wilson. "Selecting representative samples from complex biological datasets using k-medoids clustering". *Frontiers in Genetics* (2022) p. 1787. DOI: 10.3389/fgene.2022.954024

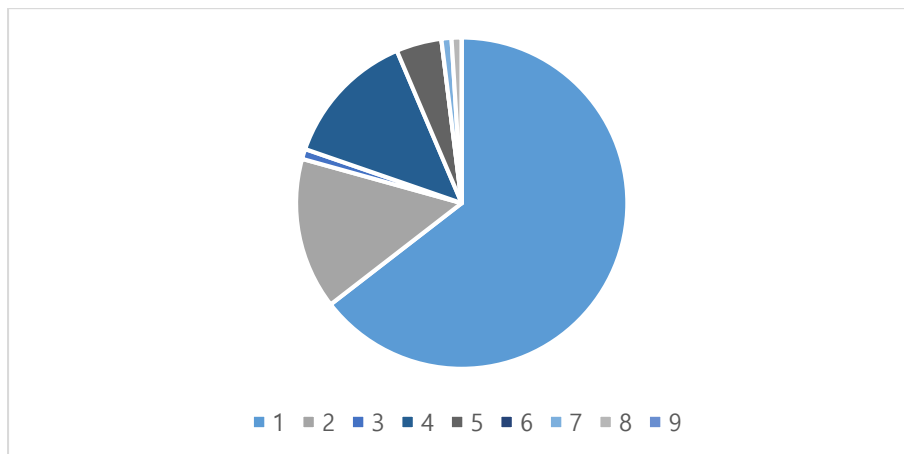
**[Table S1]** Average Sentiment of Clusters

클러스터 #	Negative	Positive	Neutral
3	1.931518	95.84795	2.220536
24	3.184265	93.13084	3.684896
27	54.45527	37.40738	8.137351
30	19.12387	79.23171	1.64442
10	6.000033279	90.13564333	3.864322395

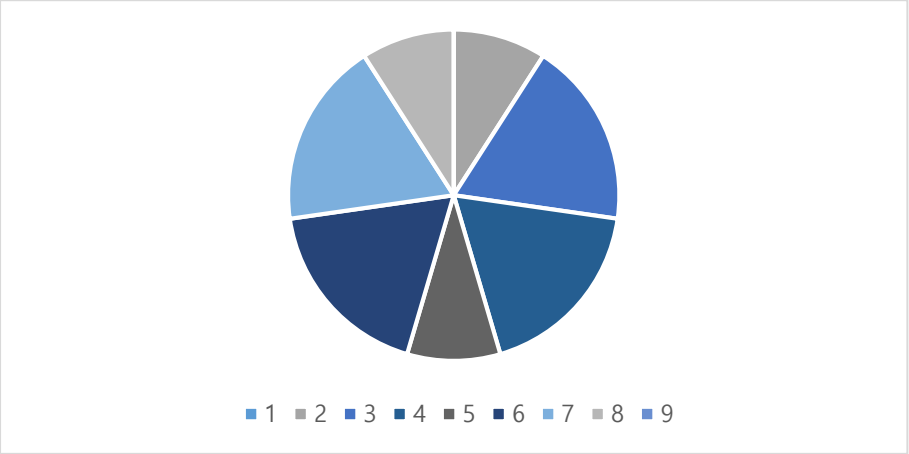
**[Table S2]** Sentiment Flow of Cluster#3



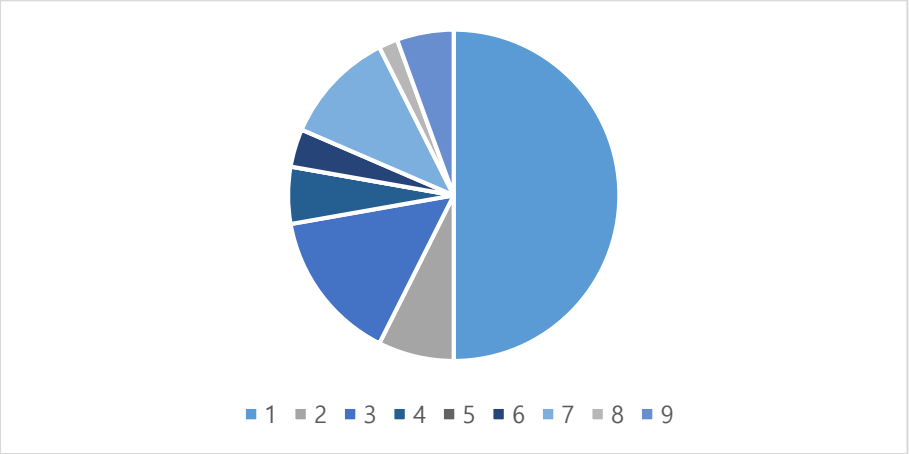
**[Table S3]** Sentiment Flow of Cluster#24



[Table S4] Sentiment Flow of Cluster#27



[Table S5] Sentiment Flow of Cluster#30



[Table S6] Sentiment Flow of Cluster#10

