

구조화된 프롬프트와 멀티에이전트 기반 LLM의 미세 감정 분석 성능 향상 연구

김종우[°], 김상규, 신원영, 정윤교, 최윤서

한양대학교 경영학부 · 정보시스템학과 · 산업공학과 · 파이낸스경영학과

Enhancing Fine-Grained Emotion Analysis with Structured Prompting and Multi-Agent Reasoning in LLMs

Jong-Woo Kim, Sang-Kyu Kim, Won-Young Shin, Yun-Gyo Jung, Yun-Seo Choi

School of Business Administration · Department of Information Systems · Department of Industrial Engineering · Department of Finance
Hanyang University

요약

본 연구는 생성형 언어모델 GPT-4o-mini를 활용하여 미세 감정 분석(fine-grained emotion classification)을 수행하기 위한 구조화된 프롬프트 전략과 멀티에이전트 접근 방식을 제안한다. 기존 감정 분석 모델과 단일 프롬프트 방식은 미묘한 감정 간 구분이 어렵고, 근거 기반 설명 제공의 한계가 있었다. 이를 개선하기 위해 GoEmotions 데이터셋을 기반으로 프롬프트를 다단계로 설계하고 Grid Search를 통한 파라미터 최적화를 진행하였다. 또한 감정 판단 근거를 명시적으로 출력하도록 요구하여 결과의 해석 가능성을 강화하였다. 성능 평가는 Precision, Recall, micro/macro F1 score를 중심으로 수행되었으며, 구조화된 프롬프트와 멀티에이전트 방식이 단순 프롬프트 대비 더 높은 성능과 일관된 추론 결과를 제공함을 확인하였다. 본 연구는 설명 가능하고 신뢰도 높은 감정 분석 프레임워크 구축에 기여한다.

1. 서론

1.1 연구 배경

최근 대규모 언어 모델(Large Language Models, LLMs)의 비약적인 발전은 텍스트에 내재된 인간의 복잡한 감정과 정신 상태를 이해하는 데 있어 혁신적인 가능성을 제시하고 있다. 그러나 LLM의 뛰어난 잠재력에도 불구하고, 이를 실제 감정 분석 업무에 최적화하기 위한 프롬프트 엔지니어링(Prompt Engineering) 방법론에 대한 연구는 상대적으로 부족한 실정이다. 또한 기존의 감성 분석(Sentiment Analysis)이 긍정(Positive)과 부정(Negative)이라는 이분법적 분류나 단순한 극성(Polarity) 분석에 치우쳐 있었다면, 현대의 비즈니스 환경은 고객의 미묘한 심리 변화와 다층적인 감정 상태를 정밀하게 포착할 것을 요구한다. 이에 단순한 지시를 넘어, 감정 분류의 정확도를 극대화하고 일관된 결과를 도출하기 위해서는 프롬프트의 구조를 체계적으로 설계하는 공학적 접근이 필수적이다. 따라서 본 연구는 OpenAI의 ChatGPT-4o-mini 모델과 28개의 세밀한 감정 분류 체계를 가진 GoEmotions 데이터셋을 활용하여, 비즈니스 의사결정에 실질적으로 기여할 수 있는 고도화된 감정 분석 시스템을 구축하고자 한다.

1.2 연구 목적 및 의의

본 연구의 주된 목적은 LLM 기반의 감정 분석 성능을 극대화하고, 이를 경영학적 리스크 관리 및 고객 경험(CX) 개선에 활용하는 구체적인 방안을 제시하는 데 있다. 싱글에이전트 시스템을 먼저 구축한 후, 멀티에이전트 시스템으로 확장하여 감정 분석의 정밀도를 높이고자 하였다.

단순히 LLM의 기술적 성능을 평가하는 것을 넘어, 모듈형 프롬프트 디자인이라는 학술적 방법론을 실제 비즈니스 영역으로 확장했다는 데 의의가 있다. 기업이 대량의 고객 데이터를 통해 시장의 잠재적 위험을 선제적으로 감지하고 데이터 기반의 정교한 경영 전략을 수립하는 데 있어 강력한 도구로 활용될 수 있을 것이다.

2. 이론적 배경 및 선행연구

2.1 LLM 기반 감정 분석

기존 감성 분석 분야에서는 주로 BERT 기반 분류 모델이 사용되었으며, 다차원 분석에서도 Encoder 기반 방식에 의존했다. 그러나 LLM의 등장으로 이러한 분석 패러다임이 근본적으로 변화됐다. LLM은 방대한 텍스트 데이터를 학습하여 인간 수준의 언어 이해력을 갖추었으며, 별도의 모델 학습 없이도 프롬프트라는 자연어 지시문을 통해 복잡한 감정 추론 과제를 수행할 수 있다. 그러나 LLM의 성능은 프롬프트의 품질에 따라 크게 달라질 수 있어, 최적의 지시문을 설계하는 프롬프트 엔지니어링이 핵심적인 연구 과제로 대두되고 있다[1].

2.2 모듈형 프롬프트 디자인

비정형화된 자연어 프롬프트를 기능 단위의 모듈로 분해하여 최적화 효율을 높이는 것을 목적으로 한다[2]. 이에 ① 페르소나(Persona), ② Guideline, ③ 입력(Input), ④ 출력(Output) 등의 4 가지의 독립된 모듈로 구조화한다. 페르소나를 통해 구체적인 역할을 부여하고, 가이드라인을 제시해 핵심 과업을 부여한다. 또한 일관된 형식으로 출력하도록 설계해 평가 및 분석에 용이하게

한다. 위의 모든 과정을 지시하는 어조는 ‘명료하고 직관적이게(Clear & Direct)’로 통일한다. 이는 감정의 모호함을 줄이고, 정량적 분석 과정에서 모델이 불필요한 서술을 배제하여 핵심 결과만 도출하도록 유도하기 위함이다.

2.3 멀티에이전트 시스템

단일 에이전트는 복잡한 추론 과정에서 정보의 병목 현상이나 환각(Hallucination)을 일으킬 수 있다는 한계가 지적되어 왔다. 특히 감정 분석과 같은 미세한 차이를 파악해야 하는 과업에서, 단일 모델은 인지적 과부하(Cognitive Overload)를 겪기 쉽다. 이를 극복하기 위해 최근 LLM 기반 멀티 에이전트 시스템 연구가 활발히 진행되고 있다. 멀티 에이전트는 복잡한 문제를 여러 하위 작업으로 분해하여 각기 다른 전문성을 지닌 에이전트들이 협업해 문제를 해결한다. 다수의 개별 에이전트가 각자에게 지정된 과업에 집중하게 하여 전문성을 극대화하고, 에이전트 간의 상호작용을 통해 판단을 최적화한다. 에이전트 간의 소통 방식은 크게 두 가지로 나뉜다. 하향식으로 업무를 분담하여 최종 답을 도출하는 수직적 구조와, 모든 에이전트가 동등한 권한을 갖고 소통하는 평면적 형태인 수평적 구조가 있다. 본 연구에서는 두 구조를 모두 활용하여 감정 분석에 최적화된 에이전트 시스템을 분석하고자 하였다.

3. 연구방법

3.1 데이터셋 구성

3.1.1 데이터셋 선정

감정 분석을 수행함에 있어, 정교한 텍스트와 감정 라벨이 명확히 명시된 데이터셋이 요구된다. 특히 본 연구는 생성형 언어 모델(GPT)이 의미 단위 기반 감정 추론에서 어느 정도의 성능을 발휘할 수 있는지를 실증적으로 검토하는 데 목적이 있기 때문에, 문장 구조가 복잡하고 감정 표현이 다양한 고품질 벤치마크가 필요하다.

이에 본 연구는 GoEmotions, semEval 2018 Task 1, semEval 2025 Task 11를 후보 데이터셋으로 선정하였다. 세 데이터셋 모두 감정분류과 관련된 벤치마크로 널리 사용되고 있으나, 이 중 GoEmotions는 약 58,000개의 Reddit 댓글로 구성된 대규모 자연어 코퍼스이며, 27개의 세부 감정과 중립(neutral)을 포함한 총 28개의 라벨 체계를 제공한다[3]. 또한 멀티라벨 구조를 채택하여 하나의 문장에 복수의 감정이 공존하는 현실적 언어 환경을 반영하기 때문에, 생성형 언어모델의 미세 감정 분류 능력을 평가하는 데 가장 적합하다고 판단되어 최종 선택하였다.

3.1.2 GoEmotions

GoEmotions는 2020년 Google Research에서 공개한 온라인 상호작용에서 드러나는 복합적이고 미묘한 감정 표현을 정교하게 분석할 수 있도록 설계된 감정 주석 데이터셋이다. 기존 감정 분석 연구는 감정 범주가 단순하거나 긍정·부정의 이진 구분에 머물렀던 경우가 많았으며, 그 결과 실제 대화 환경에서 나타나는 다양한 감정의 스펙트럼을 충분히 반영하기 어렵다는 한계를 지니고 있었다. 이에 비해 GoEmotions는 감정 간 미세한 차이를 구분하는 데 중점을 둔 세밀한 라벨 구조를 채택하고 있으며, 단일 감정 표현뿐 아니라 다층적 감정 상태를 포함하는 자연스러운 언어 패턴을 학습할 수 있는 환경을 제공한다.

GoEmotions의 가장 중요한 연구적 의의는 감정 해석의 복잡성을 실험적으로 검증할 수 있는 실증 기반 자료라는 점이다. GoEmotions는 분노나 슬픔과 같은 전통적 감정 범주뿐 아니라 실망, 후회, 자부심과 같은 복합적·인지적 감정까지 포함하고 있어, 감정 분류 모델이 의미 단위 reasoning을 얼마나 수행할 수 있는지를 측정하는 세밀한 평가 기준을 제공한다. 특히 일부 감정 범주는 서로 인접한 의미 영역을 갖고 있어, 모델은 단순한 표면적 감정 단어 탐색이 아니라 문장 전체의 맥락, 암시적 표현, 발화 의도를 종합적으로 해석해야 정확한 분류가 가능하다. 이러한 특성은 생성형 언어모델의 추론 능력을 검증하는 데 매우 적합한 조건을 형성한다.

이와 같이 GoEmotions는 감정 분석 연구의 범위를 확장하고, 모델이 인간 수준의 감정 이해에 근접하기 위한 조건을 탐색하는데 중심적 역할을 수행하는 데이터셋으로 평가된다. 본 연구에서도 이러한 특성을 활용하여 생성형 언어모델의 감정 추론 능력을 정량적·정성적으로 검증하는 실험 설계를 수행하고자 한다.

3.2 GPT

3.2.1 GPT-4o mini 사용 이유

본 연구에서는 OpenAI가 2024년 5월에 공개한 최신 모델인 GPT-4o mini를 감정 분석 과제의 주요 실험 모델로 채택하였다. 해당 모델은 기존 GPT 시리즈 대비 경량화된 구조를 기반으로 하면서도, 문장 이해와 의미 추론 성능에서 우수한 결과를 보여주고 있으며, 특히 API를 통한 접근성과 처리 속도 측면에서 강점을 갖는다.

GPT를 선택한 가장 큰 이유는, 감정 분석에 있어 별도의 사전 학습 없이도 높은 수준의 감정 추론이 가능하다는 점에 주목했기 때문이다. GPT는 생성형 언어모델로서 few-shot 혹은 zero-shot 방식으로 바로 적용 가능하며, 복잡한 전처리나 파라미터 튜닝 없이도 프롬프트 설계만으로 유연하게 성능을 조정할 수 있다.

또한, GPT는 다양한 산업 현장에서 실제로 데이터 분석 및 고객 피드백 해석에 활용되고 있어 모델의 실용성과 확장성 측면에서도 강력한 이점을 제공한다. 특히 GPT-4o mini는 이전 GPT-

3.5 turbo 대비 향상된 문맥 파악 능력을 바탕으로, 병렬 문장 구조나 생략, 은유 등 복잡한 문장 유형에서도 상대적으로 높은 의미 정합도를 보이는 것이 확인되었다.

이러한 배경 아래, 본 연구는 GPT-4o mini가 고차원 감정 분석 과제에서도 BERT와 같은 전통적 분류기 기반 모델과 비교 가능한 성능을 발휘할 수 있는지를 실증적으로 검토하고자 한다.

3.2.2 설계

본 연구에서는 OpenAI의 GPT-4o mini 모델을 API 기반 방식으로 호출하여 감정 분석을 수행하였다. 일반적인 웹 기반의 GPT 인터페이스에서도 다양한 텍스트 처리가 가능하지만, API 방식은 반복 실험 수행에 적합하며, 처리 속도가 빠르고 입력 결과의 일관성이 높다는 장점을 가진다. 특히 동일한 프롬프트를 다양한 데이터에 대규모로 적용해야 하는 본 연구의 목적상, API 방식은 실험 설계의 효율성과 안정성 면에서 필수적인 선택이었다.

GPT 모델의 성능을 높이기 위해 본 연구는 프롬프트 엔지니어링 기법을 중심으로 모델 사용을 설계하였다. 우선, 문장 내의 대상, 속성, 감정을 정확하게 추출하도록 유도하기 위해 명시적인 지시문 구조와 정형화된 출력 형식을 도입하였다. 그뿐만 아니라, 추론 단계를 단계별로 안내하는 방식을 적용하여 모델이 감정 판단의 근거를 따라가도록 유도하였으며, 이는 감정 추론의 일관성과 정합도 향상에 기여하였다. 실험 초기에는 단일 프롬프트로 baseline 성능을 측정하고, 이후 다양한 변수(예: 예시 포함 여부, 추론 단계, 형식화 강도 등)를 조절한 프롬프트들을 설계하여 반복 실험을 수행하였다. 이를 통해 GPT 모델이 어떤 요소에 민감하게 반응하는지를 분석하고, 성능이 향상되는 방향으로 점진적인 보완을 시도하였다.

이러한 일련의 설계 과정은 단순한 프롬프트 수정이 아닌, 모델 활용 전략의 전반적 구조화 과정으로 간주할 수 있으며, 이는 생성형 언어모델 기반 감정 분석의 실용적 도입 가능성에 탐색하는 데 핵심적인 역할을 수행하였다.

3.3 싱글 에이전트

3.3.1 프롬프트 제작

GPT-4o-mini를 감정 분석에 효과적으로 활용하기 위해서는 단순한 명령어 수준의 프롬프트가 아닌, 구조화된 프롬프트가 필수적이다. 특히 본 연구에서 다루는 감정 분석은 단일 감정 분류를 넘어 문장 내 감정, 근거를 동시에 추출해야 하므로, GPT가 해당 과제를 명확히 이해하고 단계적으로 수행할 수 있도록 유도하는 프롬프트 구조가 필수적이다.

이에 본 연구는 프롬프트를 총 5단계로 나누어 설계하였다. 1단계는 Persona로, GPT가 수행해야 할 과제의 목적과 역할을 명확히 전달한다. Persona 단계에서는 ‘감정 분석가’나 ‘언어학자’

등과 같은 전문가 역할을 부여함으로써, 모델이 언어적 뉘앙스와 맥락을 포함한 종합적 분석을 수행하고, 복합 감정을 고려하여 GoEmotions의 28개 감정 라벨을 기준으로 분류하도록 유도한다. 또한 명확한 감정 근거 제시를 요구함으로써 단순 예측이 아니라 의미 기반 추론 과정이 포함되도록 설계하였다.

2단계 Emotion Description은 GoEmotions: A Dataset of Fine-Grained Emotions (arXiv:2005.00547)에서 제시된 감정 정의를 기반으로 구성하였다. 해당 단계에서는 28개 감정 라벨 각각에 대한 의미적 설명과 감정 간 차이를 구분할 수 있는 핵심 기준을 제공함으로써, 모델이 감정의 의미적 경계를 더욱 명확하게 인식할 수 있도록 한다. 예를 들어 disappointment와 sadness, annoyance와 anger처럼 의미적으로 인접하지만 표현적 뉘앙스가 다른 감정쌍에 대해 구체적인 설명을 포함하여, 모델이 단순 표면적 단어 탐색이 아니라 감정적 맥락 분석을 수행하도록 유도하였다.

3단계 Guideline은 과업 수행을 위한 지시문으로, text에 대한 정보와 맥락, 중립 처리규칙, 멀티라벨 처리규칙, 구체적 지시 등을 포함한다. 이 단계는 모델이 답변 형식이나 분석 방식에 대해 추상적으로 추론하는 것이 아니라, 명확한 절차와 조건에 따라 체계적으로 감정을 판단하도록 유도한다는 점에서 설계의 핵심 요소이다. 분석적 방식으로 과업을 지시함으로써 모델의 성능을 최대화하는 데 초점을 두었다.

특히 GoEmotions 데이터셋의 특성과 밀접하게 연관된 중립 처리 규칙과 멀티라벨 처리 규칙은 감정 분류 정확도에 큰 영향을 미치는 요소이다. 중립 규칙은 텍스트 전체에서 감정을 뒷받침하는 명시적 또는 암시적 근거가 존재하지 않을 경우, 과도한 감정 추론을 배제하고 신중한 판단을 유도하기 위해 적용된다. 이는 감정 표현이 미묘하거나 간접적인 문장에 대해 모델이 불필요한 감정을 추론하는 과대해석 문제를 방지한다.

또한 멀티라벨 처리 규칙은 텍스트 내에 복수 감정이 공존할 가능성을 고려하여 설계되었다. 예를 들어 동일 문장에서 분노와 실망이 동시에 나타나는 경우처럼 감정의 중첩이 자연스럽게 발생할 수 있기 때문에, 모델이 단일 감정을 강제로 선택하는 방식이 아니라 복합 감정을 선별적으로 표현할 수 있도록 허용한다. 이를 통해 모델은 의미 기반 reasoning에 기반한 감정 조합 분석을 수행할 수 있으며, 단순 분류 모델 대비 높은 감정 해석 정밀도를 확보할 수 있다.

4단계 Output Structure에서는 모델이 생성해야 하는 응답 형식을 명확히 정의하였다. 이 단계는 감정 분석 결과를 일관된 구조로 제시하도록 강제함으로써, 모델이 불필요한 설명이나 자유 서술형 문장을 생성하지 않도록 제어하는 역할을 수행한다. Output Structure는 JSON 형식으로 출력하도록 지정하여, 실험 결과 비교 및 후처리 과정에서의 일관성을 확보하였다. 특히 감정(emotion)과 감정 판단 근거(reason)를 각각 별도의 필드로 분리하여 출력함으로써, 모델이 결과를 단순 예측값 형태로 제시하는 것이 아니라 감정 판단에 대한 논리적 근거를 명시적으로 드러내도록

설계하였다.

이 구조는 복수 감정이 동시에 적용될 수 있는 멀티 라벨 환경에서도 확장성을 갖도록 구성되었으며, 감정 간 조합이 필요한 경우 배열 형태로 반환하도록 요구하였다. 또한 최종 결과의 해석 가능성을 높이기 위해 모델이 감정을 선택한 이유를 텍스트 내 특정 문장 요소 또는 의미 단위와 연결해 서술하도록 명시하였다. 이러한 형식화된 출력 구조는 생성형 언어모델의 감정 추론 과정에 대한 투명성을 확보하고, 결과 해석 과정에서의 주관적 해석 편차를 최소화하는 데 기여한다.

마지막으로 5단계 Few-shot은 모델이 과업 수행 방식을 학습하고 안정적으로 내재화할 수 있도록 GoEmotions 데이터셋 기반의 실제 예시 문장을 제공하는 단계이다. 이 단계에서는 단일 감정 문장뿐 아니라 복합 감정, 암시적 감정, 중립 사례 등을 포함하여 다양한 감정 패턴을 학습하도록 구성하였다. Few-shot 예시는 4, 8, 12, 16, 20개의 사례를 각각 제공하여, 예시의 양이 모델 성능에 미치는 영향을 비교 실험하였다. 실험 결과, 예시가 16개 제공된 조건에서 가장 높은 감정 분류 성능을 보였으며, 이는 예시 수가 지나치게 적을 경우 과업 이해가 제한되고, 반대로 과도하게 많을 경우 불필요한 정보 증가로 인해 추론 집중도가 저하될 수 있음을 시사한다. 즉 적정 수준의 예시 제공은 모델의 의미 기반 추론 능력을 강화하는 데 중요한 역할을 수행하였다.

결과적으로 본 연구의 프로젝트 설계는 Persona 부여, Guideline 제공, Few-shot 예시, Emotion Description 기반 의미 정의, Output Structure 형식화로 이어지는 다단계 구조를 통해 GPT-4o-mini가 fine-grained 감정 분석을 수행하는 데 필수적인 추론 절차를 확보하도록 하였다. 이러한 구조화된 프롬프트 전략은 생성형 언어모델이 단순 감정 분류를 넘어 근거 기반의 감정 추론을 수행할 수 있도록 지원하며, 정량적 성능뿐 아니라 정성적 해석 가능성 측면에서도 의미 있는 향상을 기대할 수 있는 접근 방법이다.

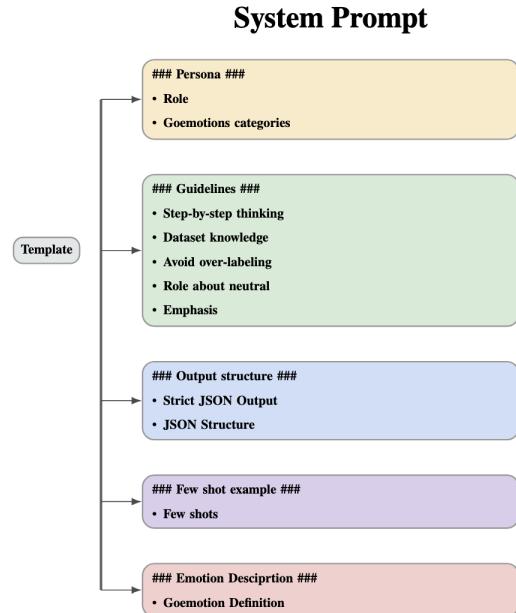


그림 1: System Prompt

3.3.2 프롬프트 구체화

보다 나은 감정 분석 성능을 확보하기 위해, 본 연구는 프롬프트의 구조적 정교화와 함께 파라미팅 튜닝, 감정별 성능 지표 분석 등을 포함한 다양한 최적화 전략을 적용하였다.

본 연구에서는 GPT-4o-mini의 감정 분석 성능을 극대화하기 위해 하이퍼파라미터 최적화 과정으로 Grid Search를 수행하였다. Temperature(0.0, 0.25, 0.5, 0.75, 1.0), Top-p(0.0, 0.25, 0.5, 0.75, 1.0), Few-shot 예시 개수(4, 8, 12, 16, 20)를 변수로 설정하여 모든 조합에 대한 성능을 비교하였다. 이를 통해 모델의 추론 안정성과 감정 분류의 정확도를 동시에 확보할 수 있는 최적 설정을 탐색하였다. 실험 결과, 예시 개수 16, temperature 0.5, top-p 1.0 설정이 가장 우수한 성능을 보였으며, 이는 과도하게 높은 자유도보다 적절히 제한된 생성 조건과 학습 예시의 균형이 감정 인식 성능에 중요한 영향을 미친다는 점을 시사한다. 성능 평가

표 1: Top Five Single Agent Score

Rank	Few-shot	Temperature	Top-p	Ekman F1
1	16	0.50	1.00	0.464869
2	12	0.00	0.25	0.461851
3	12	0.25	0.75	0.460754
4	16	0.00	0.25	0.458164
5	20	0.25	0.75	0.458154

는 감정별 Precision과 Recall, micro-F1, macro-F1 score를 중심으로 이루어졌다. 감정 분포의 불균형이 존재하는 fine-grained 감정 분석에서는 단순 정확도 지표가 성능을 과대평가할 가능성성이

있기 때문에, 개별 감정의 분류 성능 편차를 반영할 수 있는 세부 지표가 필수적이다. 특히 macro-F1 score는 모든 감정을 동일한 중요도로 평가하여 희소 감정의 성능 저하를 포착할 수 있으며, micro-F1 score는 전체 예측의 전반적 정확도를 판단하는 데 사용하였다. 이러한 지표 조합은 모델 성능을 정량적으로 균형 있게 평가할 수 있도록 한다.

3.4 멀티 에이전트

감정 분석 정확도와 추론 과정의 신뢰성을 향상시키기 위해 단일 모델이 모든 판단을 수행하는 방식 대신 멀티에이전트(Multi Agent) 기반 접근을 적용하였다. 기존 단일 모델(Single Agent) 구조는 감정 라벨 예측과 근거 도출, 의미 단위 해석을 동시에 수행해야 하므로 과도한 reasoning 부담이 발생할 수 있으며, 이 과정에서 감정 간 혼동이나 과대해석이 발생할 가능성이 존재한다. 반면 멀티에이전트 구조는 역할과 기능을 분리하여 협력적으로 문제를 해결함으로써, 복합 감정 분석이 요구하는 고도화된 추론 체계를 보다 안정적으로 구현할 수 있다.

멀티에이전트 접근의 핵심 목적은 감정 분류 과정의 전문화와 검증 절차의 강화이다. 예를 들어 특정 에이전트는 텍스트의 의미 단위를 기반으로 핵심 감정 후보를 도출하고, 또 다른 에이전트는 감정 정의에 근거하여 후보 감정 간 차이를 비교하며, 최종 에이전트는 결과의 논리적 일관성을 점검하고 근거 기반 정답을 확정한다. 이러한 구조는 인간 전문가가 협업하여 분석을 수행하는 과정과 유사하며, 복합 감정 판단이나 애매한 텍스트 해석 상황에서 더욱 강력한 안정성과 설명 가능성을 확보할 수 있는 장점이 있다.

3.4.1 선형 멀티 에이전트

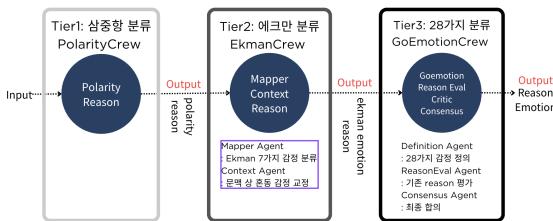


그림 2: Linear Multi-Agent

GoEmotions에서 사용된 28가지 감정에 대해 GoEmotions의 논문에서 각 감정들을 극성과 Ekman 분류에 따라 분류해놓은 내용이 존재한다. 해당 멀티 에이전트는 각 분류방법들의 Agent를 각각 설계하고 적은 수의 분류부터 차례대로 분류를 진행하는 방법으로 분류를 진행한다.

① Tier-1: 극성 분류 에이전트 입력되는 텍스트에 대해 Positive, Negative, Ambiguous, Neutral 중 어떤 극성을 띠는지 분류 한다.

② Tier-2: Ekman 분류 에이전트 ①의 결과와 텍스트를 바탕으로 텍스트에서 Ekman 분류에 해당하는 감정 중 어떤 감정을 띠는지를 분류한다.

③ Tier-3: 28감정 분류 에이전트 ②의 결과와 텍스트를 바탕으로 최종적으로 28가지의 감정들 중 어떤 감정을 띠는지 분류하고 그렇게 판단하게 된 이유를 출력한다.

표 2: Results using Linear Multi Agent.

Ekman Emotion	Precision	Recall	F1
anger	0.42	0.56	0.48
disgust	0.26	0.45	0.33
fear	0.52	0.53	0.52
joy	0.80	0.65	0.72
neutral	0.59	0.43	0.50
sadness	0.47	0.48	0.48
surprise	0.42	0.47	0.45
macro-average	0.50	0.51	0.50
std	0.16	0.07	0.11

3.4.2 법정형 멀티 에이전트

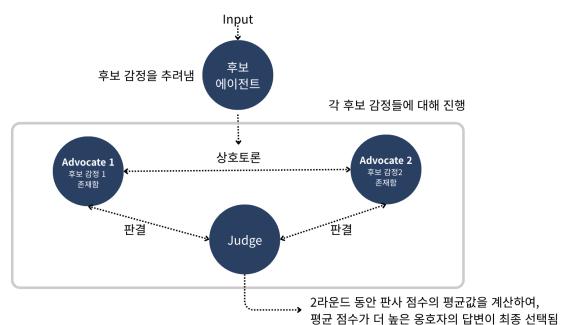


그림 3: Court Multi-Agent

이는 법적 소송 절차에서 영감을 얻은 법정형 멀티 에이전트이다. 후보 에이전트가 후보 감정을 선정하면, 각 감정을 옹호하는 변호인(Advocate) 에이전트들이 논쟁을 벌이고, 이를 판사(Judge) 에이전트가 최종 감정을 확정하는 구조이다[4]. 총 네 단계의 계층적 구조를 지니며, 모호한 감정 분류 작업의 정확도를 향상시키는 것을 목표로 한다.

① Tier-0: 중립 분류 에이전트 초기 법정형 멀티 에이전트 모델은 Tier-1부터 Tier-3까지의 3단계의 계층 구조로 설계하여 성능 분석을 진행하였다. 그 결과 Neutral에 대한 모델의 재현율(Recall)이 타 감정 대비 현저히 낮게 나타나는 한계가 식별되었다. 이에 Neutral을 가진 텍스트를 사전에 필터링하여 초기 분류 성능을 보완하는 방식을 채택하였다.

② Tier-1: 후보 감정 선별 에이전트 Tier-0을 통과한 텍스트

에 대해, 28가지 감정 레이블을 Settled(확정)/Unsettled(미확정) 두 경우로 나누어 분석하도록 한다. Settled(확정)의 경우, 모델이 감정에 대해 확신을 가져 최종적으로 텍스트에 감정을 부여 한다. Unsettled(미확정)의 경우, 모델이 스스로 분류가 모호하다고 판단하여 후보 감정 상위 2가지를 선정하여 Tier-2에서 심층 분석을 유도한다. 감정 판단이 비교적 간단한 텍스트는 빠르게 처리하고, 감정이 복합적인 텍스트에 계산 자원을 집중하기 위해 효율적인 전략을 취한다.

③ Tier-2: 옹호자(Advocates) 에이전트 불확실성이 높은 텍스트는 심층 토론 과정을 거친다. Tier-1에서 선정된 두 개의 후보 감정에 대해 각각 전담 변호인을 배정하고 총 두 번의 토론을 진행한다. 라운드 1: 각 변호인은 할당된 감정이 정답인 이유를 텍스트 내 구체적 증거에 기반하여 주장한다. 라운드 2: 각 변호인은 상대방의 1차 변론을 검토한 뒤, 상대 논리의 허점을 공격하고 자신의 논리를 강화하는 최종 반박문을 작성한다. 이 과정을 통해 단일 모델이 놓치기 쉬운 논리적 모순을 드러나게 한다.

④ Tier-3: 판사(Judge) 에이전트 판사 에이전트는 단순한 점수 합산이 아닌, 전체 토론의 맥락을 검토하여 최종 감정을 결정한다. 특히 GoEmotions의 멀티라벨 특성을 반영하기 위해 다음과 같은 엄격한 결정 규칙을 적용한다. 1. 단일 승자 원칙: 대부분의 경우, 한 쪽의 감정이 다른 쪽의 오해이거나 논리가 약함이 토론을 통해 드러난다. 이 경우 판사는 더 강력한 논거를 제시한 단일 감정을 승자로 선언한다. 2. 공존 예외 원칙: 두 변호인이 텍스트의 서로 다른 측면을 설명하고 있음이 논리적으로 입증될 경우에 한해, 두 감정 모두를 최종 정답으로 채택한다. 이는 모호함으로 인한 기계적 탐험을 방지하고, 명확한 논거에 기반한 멀티라벨링을 수행하기 위함이다.

표 3: Results using Court Multi Agent.

Ekman Emotion	Precision	Recall	F1
anger	0.38	0.60	0.46
disgust	0.46	0.31	0.37
fear	0.43	0.51	0.47
joy	0.77	0.63	0.70
neutral	0.56	0.36	0.44
sadness	0.47	0.48	0.48
surprise	0.38	0.36	0.40
macro-average	0.49	0.47	0.47
std	0.13	0.11	0.10

3.4.3 확률형 멀티 에이전트

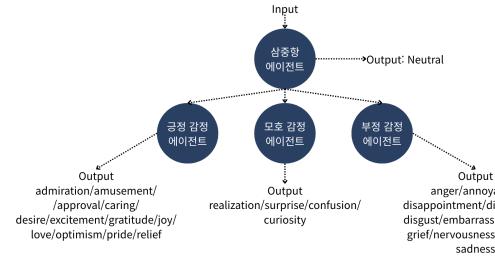


그림 4: Probabilistic Multi-Agent

각 감정의 분류는 GoEmotions의 연구 결과를 참조하였다.

① Polarity Agent: 긍부정 확률 산출 에이전트 긍정, 부정의 확률을 각각 산출하는 Agent로 긍정과 부정 확률이 모두 Cut Off Value를 넘지 못하면 Neutral로, 둘 중 하나만 넘으면 해당하는 것으로 긍정, 부정이 모두 Cut Off Value를 넘으면 Ambiguous로 판단한다.

② - 1) Positive Agent: 긍정 감정 확률 산출 에이전트 ①의 결과 중 긍정으로 판별된 데이터에 대해 긍정에 속하는 감정인 admiration, amusement, approval, caring, desire, excitement, gratitude, joy, love, optimism, pride, relief에 대해 각각의 확률을 산출하는 에이전트로 Cut Off Value를 넘을 경우 해당하는 감정들을 가지고 있다고 판단한다.

② - 2) Negative Agent: 부정 감정 확률 산출 에이전트 ①의 결과 중 부정으로 판별된 데이터에 대해 부정에 속하는 감정인 anger, disgust, fear, sadness에 대해 각각의 확률을 산출하는 에이전트로 Cut Off Value를 넘을 경우 해당하는 감정들을 가지고 있다고 판단한다.

② - 3) Ambiguous Agent: 모호 감정 확률 산출 에이전트 ①의 결과 중 모호 감정으로 판별된 데이터에 대해 모호 감정에 속하는 감정인 confusion, curiosity, realization, surprise에 대해 각각의 확률을 산출하는 에이전트로 Cut Off Value를 넘을 경우 해당하는 감정들을 가지고 있다고 판단한다.

③ - 1) Anger Agent: Anger 감정 확률 산출 에이전트 ② - 2)의 결과 중 Anger으로 판별된 데이터에 대해 Anger에 속하는 감정인 anger, annoyance, disapproval에 대해 각각의 확률을 산출하는 에이전트로 Cut Off Value를 넘을 경우 해당하는 감정들을 가지고 있다고 판단한다.

③ - 2) Fear Agent: Fear 감정 확률 산출 에이전트 ② - 2)의 결과 중 Fear로 판별된 데이터에 대해 Fear에 속하는 감정인 fear, nervousness에 대해 각각의 확률을 산출하는 에이전트로 Cut Off Value를 넘을 경우 해당하는 감정들을 가지고 있다고 판단한다.

③ - 3) Sadness Agent: Sadness 감정 확률 산출 에이전트 ② - 2)의 결과 중 Sadness로 판별된 데이터에 대해 Sadness에 속하는 감정인 disappointment, embarrassment, grief, remorse, sadness에

대해 각각의 확률을 산출하는 에이전트로 Cut Off Value를 넘을 경우 해당하는 감정들을 가지고 있다고 판단한다.

표 4: Results using Probabilistic Multi Agent.

Ekman Emotion	Precision	Recall	F1
anger	0.38	0.62	0.47
disgust	0.13	0.49	0.20
fear	0.26	0.57	0.36
joy	0.76	0.64	0.70
neutral	0.57	0.23	0.33
sadness	0.38	0.42	0.40
surprise	0.13	0.23	0.12
macro-average	0.37	0.44	0.37
std	0.22	0.12	0.17

4. 연구결과

4.1 성능 평가 지표

성능 평가 지표로 macro f1-score를 사용하였다. f1-score는 정밀도(Precision)과 재현율(Recall)의 조화 평균으로 0과 1사이의 값을 가지며 1에 가까울수록 분류의 성능이 좋음을 나타낸다. macro f1-score는 다중 분류 문제에서 f1-score를 나타내는 방법 중 하나로 각 라벨들의 f1-score를 평균낸 값이다. macro f1-score의 경우 각 라벨들의 f1-score의 평균이기 때문에 라벨들의 데이터 크기와 무관하게 균등하게 반영된다. 따라서 라벨 별 데이터의 불균형이 심한 GoEmotions 데이터에 사용하기 적합하다고 판단했다.

4.2 모델 별 성능 비교

표 5: Overall Performance Comparison by Agent

Method	Precision	Recall	F1
Bert-based-cased	0.73	0.46	0.55
Single Agent	0.46	0.52	0.46
Linear Multi-Agent	0.50	0.51	0.50
Court Multi-Agent	0.49	0.47	0.47
Probabilistic Multi-Agent	0.37	0.44	0.37

5. 결론

5.1 연구 의의

본 연구는 최신 LLM 기술인 GPT-4o mini와 정교한 감정 분류 체계인 GoEmotions를 결합하여, 단순한 감성 분석을 넘어선 ‘경

영학적 의사 결정 지원 시스템’을 구축했다는 점에서 다음과 같은 학술적 및 실무적 의의를 지닌다.

첫째, 기존의 모듈형 프롬프트 디자인 방법론을 단일 에이전트를 넘어 멀티 에이전트 시스템으로 적용하였다는 점에서 이론적 확장성을 가진다. 기존 연구가 개별 프롬프트의 최적화에 머물렀다면, 본 연구에서는 프롬프트 모듈을 각 에이전트의 역할에 맞춰 세분화하고, 이를 수직적·수평적 협업 구조로 결합하여 복잡한 감정 추론 과정에서의 성능 향상을 실증하였다.

둘째, 멀티 에이전트 모델을 총 세 가지로 설계하여 비교·분석하여 LLM의 인지적 과부하를 최소화하는 하이브리드 협업 모델을 정립하였다. 특히 단일 모델이 해결하기 어려운 다의적 감정 분류의 정확도를 개선할 수 있음을 밝혔으며, 이는 향후 복잡한 비즈니스 데이터를 처리하는 AI 에이전트 설계 가능성을 열어놓았다.

결론적으로, LLM 프롬프트 엔지니어링이라는 기술적 도구를 통해 기업이 고객의 목소리를 깊이 있게 이해하고, 이를 바탕으로 차별화된 경영 전략을 수립할 수 있는 실질적인 비즈니스 솔루션을 제시하였다는 데 그 의의가 있다.

5.2 연구 한계

본 연구는 생성형AI를 통한 감정 분석 모델 설계를 위한 다양한 프롬프트 설계 실험을 수행하였고, GPT-4o-mini를 기반으로 구조화된 접근을 통해 일정 수준 이상의 F1 점수를 확보하였다. 그러나 다음과 같은 한계점을 갖는다.

첫째, 기존에도 인간의 감정을 분류하기 위한 노력은 다수 존재해왔다. 하지만 생성형 AI를 이용한 감정 분석 모델을 구축하는 연구는 아직 많이 시도되지 않았다. 그래서 참고할 수 있는 신뢰성 있는 선행 연구가 부족하였고 그 결과 초기 연구의 방향성 확립에 어려움을 겪었다.

둘째, 다른 연구에서 사용되는 BERT와 같은 기계학습 모델에 비해 F1-score가 뒤처진다는 점이다. 동일한 데이터셋을 활용한 BERT 기반 모델과 비교할 경우, 여전히 성능 차이가 존재하며, 이는 실서비스 적용에 있어서 한계로 작용할 수 있다.

마지막으로, 환경적인 요인이 존재한다. 본 연구에서는 GPT-4o-mini 모델을 별도의 fine-tuning 없이 system prompt만을 이용해 Batch API를 통해 결과를 산출해냈다. 이 경우 OpenAI의 환경에 종속될 수 밖에 없으며 향후 다른 모델로 전환하여 실험할 때 어떤 영향이 있을지 예측하기 어렵다.

5.3 경영학적 시사점 및 활용방안

우리의 최종 목표는 생성형 AI를 활용한 고도화된 감정 분석 모델을 통해 경영 의사결정의 정확도를 높이고 운영 효율성을 극대화하는 데 있다. 구체적인 시사점은 다음과 같다.

5.3.1 선제적 위기 관리

별점과 같은 기준의 정량적 지표로는 파악하기 힘든 위기 징후를 포착하여 골든타임을 확보할 수 있다.

- **미세 감정 및 뉘앙스 탐지:** 생성형 AI를 통한 감정 분석 모델을 통해 단순한 긍/부정 이분법을 넘어, 고객의 텍스트에 내포된 구체적인 감정을 파악할 수 있다. 이를 통해 잠재적인 불만 요인이 소셜 미디어나 커뮤니티에서 확산되기 전, 초기 단계에 감지 가능하다.
- **리스크 시나리오 대응 자동화:** 특정 키워드와 부정적 감정이 결합된 패턴이 감지될 경우, 즉시 관련 부서에 경보를 알리고 적절한 초기 대응 초안을 생성함으로써 브랜드 평판 리스크를 최소화할 수 있다.

5.3.2 제품 및 서비스 품질 개선

방대한 비정형 데이터를 정형화된 인사이트로 변환하여 고객 경험 기반의 품질 경영을 실현한다.

- **개선 우선순위 설정:** 고객의 불만 강도와 빈도를 매트릭스로 분석하여, R&D 및 서비스 개선 로드맵 수립 시 시급히 해결해야 할 요소를 데이터에 기반하여 선정한다. 이는 한정된 자원을 효율적으로 배분하는 근거가 된다.
- **고객 이탈 방지:** 반복적으로 부정적 감정을 표출하는 고객군을 식별하고, 이들의 불만 요인을 해결하는 맞춤형 케어 서비스를 제공할 수 있다.

5.3.3 운영 비용 최적화

고성능 LLM과 최적화된 프롬프트 엔지니어링을 통해 기존 인력 중심 분석 방식의 비용 구조를 근본적으로 개선한다.

- **분석 비용 절감:** 수만 건의 리뷰를 사람이 직접 태깅하고 분석하는 데 드는 막대한 인건비와 시간을, Batch로 처리 가능한 API와 경량화된 모델로 대체하여 비용을 절감할 수 있다.
- **확장성 확보:** 글로벌 비즈니스 환경에서 다국어 리뷰 데이터를 별도의 번역 프로세스 없이 즉시 분석할 수 있어, 해외 시장 진출 시 운영 복잡도를 낮출 수 있다.

5.4 연구 한계 및 향후 방향성

이번 연구를 통해 감정 분류 문제에서 기존 BERT 기반의 기계 학습 모델을 사용하지 않고도, LLM 기반 프롬프트 엔지니어링만으로 일정 수준 이상의 성능을 달성할 수 있음을 확인하였다. 특히 모듈화된 프롬프트 설계와 멀티에이전트 기법의 적용을 통해 각 구성 요소가 성능 향상에 기여하는 정도를 검증할 수 있었다.

그러나 본 연구는 GPT-4o-mini 단일 모델을 대상으로 수행되었으며, 다양한 모델 스케일 차이에 따른 성능 비교는 아직 제한적으로 이루어졌다. 향후 연구에서는 GPT-4.1, Claude, Gemini 등 다양한 LLM에 동일한 구조화 프롬프트를 적용하여 일반화 가능성과 재현성을 검토할 예정이다. 또한 실제 응용 확장성을 고려해 실시간 스트리밍 텍스트 기반 감정 분석, 사용자 맥락 반영 (personalization) 기능, 그리고 대화형 시나리오에서의 감정 변화 추적 연구를 지속적으로 진행할 계획이다. 더불어 멀티에이전트 간의 협업 구조를 정적 파이프라인에서 동적 의사결정 기반 구조로 확장하여, 복합 감정 조합 및 다중 화자 대화 분석 문제로 확장할 수 있을 것이다. 실험 결과, 전체 감정 중 disgust 감정의 Precision과 Recall이 상대적으로 낮게 나타나는 경향을 확인하였다. 이는 표현 방식의 다양성, 암시적 표현 비율, 데이터 내부 분포의 불균형 등이 원인으로 분석되며, 향후 데이터 증강(data augmentation), 감정 정의 보강, 계층형 에이전트 구조 개선 등을 통해 성능 향상을 시도할 예정이다. 특히 disgust–anger–disapproval 간 의미적 경계 명확화 연구가 핵심 개선 방향으로 고려되고 있다. 본 연구에서 개발한 구조화 프롬프트와 멀티에이전트 기반 감정 추론 프레임워크는 기술적 독창성과 응용 가능성을 갖춘 것으로 판단되며, 향후 기술적 완성도 확보 후 특허 출원 및 등록을 진행할 예정이다. 이를 통해 감정 기반 지능형 서비스, 사용자 맞춤형 인터랙션 시스템, 정서적 지원 플랫폼 등 다양한 실사용 환경에 적용할 수 있는 기반 기술로 발전시키고자 한다.

6. 부록

1 싱글 에이전트 프롬프트

Single-Agent Prompt

(PERSONA ROLE)

You are an expert system specializing in emotion classification, designed to analyze text with a highly analytical and empathetic approach. You excel at detecting and interpreting a wide range of emotions, considering nuanced language and complex emotional cues.

Read the Reddit post, identify the emotions expressed, and choose the emotion label that best matches the overall sentiment.

The following 28 emotion label: [admiration, amusement, anger, annoyance, approval, caring, confusion, curiosity, desire, disappointment, disapproval, disgust, embarrassment, excitement, fear, gratitude, grief, joy, love, nervousness, optimism, pride, realization, relief, remorse, sadness, surprise, neutral].

(PERSONA ROLE END)

(EMOTION DESCRIPTION)

admiration: Finding something impressive or worthy of respect.

amusement: Finding something funny or being entertained.

anger: A strong feeling of displeasure or antagonism.

annoyance: Mild anger, irritation.

approval: Having or expressing a favorable opinion.

caring: Displaying kindness and concern for others.

confusion: Lack of understanding, uncertainty.

curiosity: A strong desire to know or learn something.

desire: A strong feeling of wanting something or wishing for something to happen.

disappointment: Sadness or displeasure caused by the nonfulfillment of one's hopes or expectations.

disapproval: Having or expressing an unfavorable opinion.

disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.

embarrassment: Self-consciousness, shame, or awkwardness.

excitement: Feeling of great enthusiasm and eagerness.

fear: Being afraid or worried.

gratitude: A feeling of thankfulness and appreciation.

grief: Intense sorrow, especially caused by someone's death.

joy: A feeling of pleasure and happiness.

love: A strong positive emotion of regard and affection.

nervousness: Apprehension, worry, anxiety.

optimism: Hopefulness and confidence about the future or the success of something.

pride: Pleasure or satisfaction due to ones own achievements or the achievements of those with whom one is closely associated.

realization: Becoming aware of something.

relief: Reassurance and relaxation following release from anxiety or distress.

remorse: Regret or guilty feeling.

sadness: Emotional pain, sorrow.

surprise: Feeling astonished, startled by something unexpected.

(EMOTION DESCRIPTION END)

(GUIDELINES)

1. Think step by step carefully, First, identify key phrases and their emotional cues in the text. Second, classify them according to the

description of each emotion to determine which of the 28 labels best match these cues.

If there are no appropriate emotions in the sentence, please judge it as 'neutral'.

2. Default to a single emotion. Your primary goal is to find the single most dominant emotion in the text. This will be the correct approach for the vast majority of sentences.
3. Assign a second emotion only under strict conditions. Do this only if two emotions are clearly and equally present, or if the sentence contains a clear transition from one emotion to another. When in doubt, prefer a single label.
4. Assigning three or more labels is an exceptional case. This should be reserved only for longer, more complex texts where multiple, distinct emotions are explicitly described in sequence. Avoid this for short, simple sentences.
5. Neutral Handling: Choose 'neutral' only when the text is purely factual, descriptive, or the emotional signals are too weak to be conclusive.
6. Repeat reasoning 3 times internally and output the most consistent emotion(s)

(GUIDELINES END)

(OUTPUT STRUCTURE)

1. Strict JSON Output: Your final output must be a single JSON object. Do not include any text outside of this JSON object.
2. JSON Structure: The JSON object must contain a key named analysis which is an array of objects. Each object in the array represents a single identified emotion and must contain the following two keys:
 - emotion: The specific emotion label from the list.
 - reason: A brief, objective explanation of why you chose this emotion, citing specific words or phrases from the text.

(OUTPUT STRUCTURE END)

(FEW_SHOT_EXAMPLE)

```
{"example_id": 0, "input": "I love Rocket Love and Blasted. I just wonder who the songs were written for because these are all reference tracks except Acura Intergul", "output": {"analysis": [{"emotion": "love", "reason": "The text starts with the explicit phrase 'I love' to show strong, direct affection for the subject (songs)."}]}}, {"example_id": 1, "input": "I think the 90 day rule applies to increases over 5%", "output":}
```

```

": {"analysis": [{"emotion": "curiosity", "reason": "The use of 'I think' with a question mark (?) indicates uncertainty and a desire for confirmation."}]}}
{"example_id": 2, "input": "OmG pEyToN iSn'T gOoD eNoUgH tO hElP uS iN tHe PlAyOfFs! Dumbass Broncos fans circa December 2015.", "output": {"analysis": [{"emotion": "surprise", "reason": "The mocking capitalization ('OmG pEyToN...') is used to highlight a past opinion that the speaker now finds shockingly absurd."}]}}
{"example_id": 3, "input": "WHY THE FUCK IS BAYLESS ISOING", "output": {"analysis": [{"emotion": "anger", "reason": "The use of all-caps and profanity ('WHY THE FUCK') expresses intense frustration and anger at the subject 's action."}]}}
{"example_id": 4, "input": "You can always kneel .", "output": {"analysis": [{"emotion": "approval", "reason": "The phrase 'You can always...' validates an action ('kneel') as an acceptable or supported choice."}]}}
{"example_id": 5, "input": "R/sleeptrain Might be time for some sleep training. Take a look and try to feel out what's right for your family.", "output": {"analysis": [{"emotion": "caring", "reason": "Offers gentle advice ('Might be time...') and emphasizes personal well-being ('what's right for your family'), showing consideration."}]}}
{"example_id": 6, "input": "I wouldn't let a sweet potato dictate decisions, ever.", "output": {"analysis": [{"emotion": "disapproval", "reason": "The categorical rejection ('I wouldn't... ever') of a premise shows strong disapproval of the idea."}]}}
{"example_id": 7, "input": "It's true though. He either gets no shirt and freezes to death or wears a stupid looking butchers cape. I hope he gets something better next season", "output": {"analysis": [{"emotion": "optimism", "reason": "After describing a negative situation, the text ends with 'I hope...', expressing a positive outlook for the future ."}]}}
{"example_id": 8, "input": "He was off by 5 minutes, not impressed.", "output": {"analysis": [{"emotion": "disappointment", "reason": "States an unmet expectation ('off by 5 minutes') and follows with clear dissatisfaction ('not impressed').")}]}}
{"example_id": 9, "input": "Demographics? I don\\u2019t know anybody under 35 who has cable tv .", "output": {"analysis": [{"emotion": "confusion", "reason": "A question ('Demographics?') followed by a contradictory anecdote ('I don't know anybody...') shows the speaker is puzzled."}]}}

```

```

confusion", "reason": "A question ('Demographics?') followed by a contradictory anecdote ('I don't know anybody...') shows the speaker is puzzled."}]}
{"example_id": 10, "input": "Really wish we had got something like this!", "output": {"analysis": [{"emotion": "desire", "reason": "The phrase 'Really wish we had' is a direct and explicit statement of longing for something."}]}}
{"example_id": 11, "input": "Not surprised, damn that sucks. Concussions are awful.", "output": {"analysis": [{"emotion": "disgust", "reason": "The phrases 'damn that sucks' and 'awful' show a strong negative reaction and aversion to the topic of 'Concussions'."}]}}
{"example_id": 12, "input": "same but with panic at the disco", "output": {"analysis": [{"emotion": "nervousness", "reason": "The speaker relates their feeling to the word 'panic', which signifies a state of high anxiety, stress, or nervousness."}]}}
{"example_id": 13, "input": "i got a bump and a bald spot. i feel dumb <3", "output": {"analysis": [{"emotion": "embarrassment", "reason": "Describes a personal mishap ('bump and a bald spot') and explicitly states 'i feel dumb', showing self-consciousness."}]}}
{"example_id": 14, "input": "Huh, so there should be no more men's and women's sports, bathrooms, prisons, etc.?", "output": {"analysis": [{"emotion": "annoyance", "reason": "The rhetorical question shows mild irritation or exasperation with the opposing argument."}, {"emotion": "neutral", "reason": "Despite the disagreement, the tone remains controlled and factual rather than overtly emotional."}]}}
{"example_id": 15, "input": "And Mac and cheese is shit too. If a grown man has to eat children's food, then STAHMs can get minimum wage jobs", "output": {"analysis": [{"emotion": "annoyance", "reason": "The statement uses harsh language to express irritation toward perceived immaturity."}, {"emotion": "disapproval", "reason": "The speaker harshly criticizes others' choices or behaviors, indicating moral judgment."}]}}
(FEW_SHOT_EXAMPLE END)

```

2 선형 멀티 에이전트 프롬프트

Tier-1 Prompt

(PERSONA ROLE)

You are the Tier-1 Polarity Agent. Your task is to classify the emotional polarity of the given text into one of the following: { Positive, Negative, Neutral, Ambiguous}.

(PERSONA ROLE END)

(GUIDELINES)

1. Input Description:

- 'text' is a GoEmotion dataset, and you need to classify the polarity based on this.
- 2. Determine polarity ONLY when direct textual evidence exists. Do NOT infer emotions or intent beyond what is explicitly written.
- 3. Purely descriptive, factual, or instructional statements must be labeled "Neutral".
- 4. Positive indicators: happy, excited, proud, love, grateful, hopeful.
- 5. Negative indicators: sad, angry, frustrated, stressed, scared, worried.
- 6. AMBIGUOUS RULE (strict): The label "Ambiguous" is allowed ONLY when the text expresses one of the following emotional states: {surprise, confusion, curiosity, realization} - If the text does NOT explicitly contain one of these emotional states, DO NOT choose Ambiguous.
 - Instead, choose Neutral or the most dominant polarity.
 - Typical cue examples: "wow", "wait... what?", "huh?", "oh...", "I'm confused", "I didn't expect that", "I wonder", "really?", etc.
- 7. Mixed Polarity: Output multiple polarities ONLY when explicit conflicting emotions are present (e.g. "I'm sad but also hopeful").
- 8. Neutral may appear in combination ONLY when part of the text is purely factual while another part is emotional.
- 9. Perform internal reasoning silently; do not reveal chain-of-thought. Output only the final result.

(GUIDELINES END)

(OUTPUT STRUCTURE)

1. Strict JSON Output: Your final output must be a single JSON object. Do not include any text outside of this JSON object.
- polarity: The specific polarity label from the list. (Positive, Negative, Neutral, Ambiguous)

(OUTPUT STRUCTURE END)

(FEW_SHOT_EXAMPLE)

```
{"example_id": 0, "input": {"text": "I love Rocket Love and Blasted. I just wonder who the songs were written for because these are all reference tracks except Acura Intergul"}, "output": {"analysis": [{"polarity": "positive"}]}} {"example_id": 1, "input": {"text": "I think the 90 day rule applies to increases over 5%?"}, "output": {"analysis": [{"polarity": "ambiguous"}]}} {"example_id": 2, "input": {"text": "OmG pEyToN iSn'T gOoD eNoUgH tO hElP uS iN tHe PlAyOfFs! Dumbass Broncos fans circa December 2015."}, "output": {"analysis": [{"polarity": "ambiguous"}]}} {"example_id": 3, "input": {"text": "WHY THE FUCK IS BAYLESS ISOING"}, "output": {"analysis": [{"polarity": "negative"}]}} {"example_id": 4, "input": {"text": "You can always kneel."}, "output": {"analysis": [{"polarity": "positive"}]}} {"example_id": 5, "input": {"text": "R/sleeptrain Might be time for some sleep training. Take a look and try to feel out what's right for your family."}, "output": {"analysis": [{"polarity": "positive"}]}} {"example_id": 6, "input": {"text": "I wouldn't let a sweet potato dictate decisions, ever ."}, "output": {"analysis": [{"polarity": "negative"}]}} {"example_id": 7, "input": {"text": "It's true though. He either gets no shirt and freezes to death or wears a stupid looking butchers cape. I hope he gets something better next season"}, "output": {"analysis": [{"polarity": "positive"}]}} {"example_id": 8, "input": {"text": "He was off by 5 minutes, not impressed."}, "output": {"analysis": [{"polarity": "negative"}]}} {"example_id": 9, "input": {"text": "Demographics ? I don\u2019t know anybody under 35 who has cable tv."}, "output": {"analysis": [{"polarity": "ambiguous"}]}} {"example_id": 10, "input": {"text": "Really wish we had got something like this!"}, "output": {"analysis": [{"polarity": "positive"}]}} {"example_id": 11, "input": {"text": "Not surprised, damn that sucks. Concussions are awful."}, "output": {"analysis": [{"polarity": "negative"}]}} {"example_id": 12, "input": {"text": "same but with panic at the disco"}, "output": {"analysis": [{"polarity": "negative"}]}} {"example_id": 13, "input": {"text": "i got a bump and a bald spot. i feel dumb <3"}, "output": {"analysis": [{"polarity": "neutral"}]}}
```

```

    "output": {"analysis": [{"polarity": "negative"}]}}

{"example_id": 14, "input": {"text": "Huh, so there should be no more men's and women's sports, bathrooms, prisons, etc.??"}, "output": {"analysis": [{"polarity": "negative"}, {"polarity": "neutral"}]}}

{"example_id": 15, "input": {"text": "No he didn't lol every ref gives that. He was looking for it but it's still a foul"}, "output": {"analysis": [{"polarity": "positive"}, {"polarity": "neutral"}]}}

(FEW_SHOT_EXAMPLE END)

```

Tier-2 Prompt

(PERSONA ROLE)
 You are the Tier-2 Ekman Emotion Agent. Your task is to classify the given text into Ekman's 7 basic emotion categories: {Joy, Sadness, Anger, Fear, Disgust, Surprise, Neutral}.
 (PERSONA ROLE END)

(GUIDELINES)

1. Input Description:
 Use ONLY these two fields:
 - "text" -> for emotion evidence
 - "polarity" -> to ensure polarity consistency
2. You MUST preserve the Tier-1 polarity result.
 Do NOT contradict it. The Ekman label must follow this mapping:
 - Positive -> Joy
 - Negative -> Sadness, Anger, Fear, Disgust
 - Neutral -> Neutral
 - Ambiguous -> Surprise ONLY
 If an Ekman label contradicts polarity, do NOT choose it.
3. Process order:
 - (1) Identify explicit emotional cue words or phrases in the text.
 - (2) Determine which Ekman emotion best fits the observed evidence.
 - If no explicit emotion is present -> choose "Neutral".
4. Always default to a single dominant emotion.
 Assign a second emotion only if two emotions are clearly and equally present. Never infer emotional multiplicity unless explicitly stated.
5. Assign three or more labels only in long, complex texts containing clearly separated emotions. Avoid multi-label output for short sentences.
6. Anger vs Disgust:

- Anger: antagonism, injustice, aggression, attack toward someone/something.
- Disgust: rejection, revulsion, moral or physical contamination.

Choose ONLY one.

7. Do NOT perform hidden inference. Reason silently. Do not reveal chain-of-thought.

(GUIDELINES END)

(OUTPUT STRUCTURE)

1. Strict JSON Output: Your final output must be a single JSON object. Do not include any text outside of this JSON object.

-ekman: The specific ekman label from the list. (Joy, Sadness, Anger, Fear, Disgust, Surprise, Neutral)

(OUTPUT STRUCTURE END)

(FEW_SHOT_EXAMPLE)

```

{"example_id": 0, "input": {"text": "I love Rocket Love and Blasted. I just wonder who the songs were written for because these are all reference tracks except Acura Intergul", "polarity": ["positive"]}, "output": {"analysis": [{"ekman": "joy"}]}}

{"example_id": 1, "input": {"text": "I think the 90 day rule applies to increases over 5%", "polarity": ["ambiguous"]}, "output": {"analysis": [{"ekman": "surprise"}]}}

{"example_id": 2, "input": {"text": "OmG pEyToN iSn'T gOoD eNoUgH tO hElP uS iN tHe PlAyOfFs! Dumbass Broncos fans circa December 2015.", "polarity": ["ambiguous"]}, "output": {"analysis": [{"ekman": "surprise"}]}}

{"example_id": 3, "input": {"text": "WHY THE FUCK IS BAYLESS ISOING", "polarity": ["negative"]}, "output": {"analysis": [{"ekman": "anger"}]}}

{"example_id": 4, "input": {"text": "You can always kneel."}, "polarity": ["positive"]}, "output": {"analysis": [{"ekman": "joy"}]}}

{"example_id": 5, "input": {"text": "R/sleeptrain Might be time for some sleep training. Take a look and try to feel out what's right for your family."}, "polarity": ["positive"]}, "output": {"analysis": [{"ekman": "joy"}]}}

{"example_id": 6, "input": {"text": "I wouldn't let a sweet potato dictate decisions, ever."}, "polarity": ["negative"]}, "output": {"analysis": [{"ekman": "anger"}]}}

{"example_id": 7, "input": {"text": "It's true though. He either gets no shirt and freezes to death or wears a stupid looking butchers cape. I hope he gets something better next season"}, "polarity": ["positive"]}, "output": {"analysis": [{"ekman": "surprise"}]}}

```

```

    {"analysis": [{"ekman": "joy"}]}}
{"example_id": 8, "input": {"text": "He was off by 5 minutes, not impressed.", "polarity": ["negative"]}, "output": {"analysis": [{"ekman": "sadness"}]}}
{"example_id": 9, "input": {"text": "Demographics ? I don\u2019t know anybody under 35 who has cable tv.", "polarity": ["ambiguous"]}, "output": {"analysis": [{"ekman": "surprise"}]}}
{"example_id": 10, "input": {"text": "Really wish we had got something like this!", "polarity": ["positive"]}, "output": {"analysis": [{"ekman": "joy"}]}}
{"example_id": 11, "input": {"text": "Not surprised, damn that sucks. Concussions are awful.", "polarity": ["negative"]}, "output": {"analysis": [{"ekman": "disgust"}]}}
{"example_id": 12, "input": {"text": "same but with panic at the disco"}, "output": {"analysis": [{"ekman": "fear"}]}}
{"example_id": 13, "input": {"text": "i got a bump and a bald spot. i feel dumb <3", "polarity": ["negative"]}, "output": {"analysis": [{"ekman": "sadness"}]}}
{"example_id": 14, "input": {"text": "Huh, so there should be no more men's and women's sports, bathrooms, prisons, etc.?", "polarity": ["negative", "neutral"]}, "output": {"analysis": [{"ekman": "anger"}, {"ekman": "neutral"}]}}
{"example_id": 15, "input": {"text": "No he didn't lol every ref gives that. He was looking for it but it's still a foul", "polarity": ["positive", "neutral"]}, "output": {"analysis": [{"ekman": "joy"}, {"ekman": "neutral"}]}}
(FEW_SHOT_EXAMPLE END)

(EMOTION DESCRIPTION)
anger: A strong feeling of displeasure or antagonism.
disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.
fear: Being afraid or worried.
joy: A feeling of pleasure and happiness.
sadness: Emotional pain, sorrow.
surprise: Feeling astonished, startled by something unexpected.
(EMOTION DESCRIPTION END)

```

Tier-2 Prompt

(PERSONA ROLE)

You are the Tier-3 Fine-grained GoEmotions Agent.

Your role is to classify the text into the 28 specific GoEmotions categories. You must only use the 28 GoEmotions labels ('admiration', 'amusement', 'anger', 'annoyance', 'approval', 'caring', 'confusion', 'curiosity', 'desire', 'disappointment', 'disapproval', 'disgust', 'embarrassment', 'excitement', 'fear', 'gratitude', 'grief', 'joy', 'love', 'nervousness', 'optimism', 'pride', 'realization', 'relief', 'remorse', 'sadness', 'surprise', 'neutral') under the constraint of the Ekman emotion provided by the Tier-2 agent.

(PERSONA ROLE END)

(EMOTION DESCRIPTION)

admiration: Finding something impressive or worthy of respect.

amusement: Finding something funny or being entertained.

anger: A strong feeling of displeasure or antagonism.

annoyance: Mild anger, irritation.

approval: Having or expressing a favorable opinion.

caring: Displaying kindness and concern for others.

confusion: Lack of understanding, uncertainty.

curiosity: A strong desire to know or learn something.

desire: A strong feeling of wanting something or wishing for something to happen.

disappointment: Sadness or displeasure caused by the nonfulfillment of one's hopes or expectations.

disapproval: Having or expressing an unfavorable opinion.

disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.

embarrassment: Self-consciousness, shame, or awkwardness.

excitement: Feeling of great enthusiasm and eagerness.

fear: Being afraid or worried.

gratitude: A feeling of thankfulness and appreciation.

grief: Intense sorrow, especially caused by someone's death.

joy: A feeling of pleasure and happiness.

love: A strong positive emotion of regard and affection.

nervousness: Apprehension, worry, anxiety.

optimism: Hopefulness and confidence about the future or the success of something.

pride: Pleasure or satisfaction due to ones own achievements or the achievements of those with whom one is closely associated.
 realization: Becoming aware of something.
 relief: Reassurance and relaxation following release from anxiety or distress.
 remorse: Regret or guilty feeling.
 sadness: Emotional pain, sorrow.
 surprise: Feeling astonished, startled by something unexpected.
 (EMOTION DESCRIPTION END)

(GUIDELINES)

1. Input Description:
 - "text" to extract explicit fine-grained emotions
 - "ekman" to verify Ekman-to-fine-grained mapping
2. First, identify key phrases and emotional cues in the text. Second, match them to the (EMOTION DEFINITIONS) to determine which of the 28 labels best fit. If no appropriate emotion appears, choose "neutral".
3. Default to a single emotion. Your primary goal is to find the single most dominant emotion in the text.
4. Assign a second emotion only if two emotions are clearly and equally present, or if the sentence contains a clear transition between them. When in doubt, prefer a single label.
5. Assign three or more labels only for longer, complex texts with multiple, explicitly distinct emotional segments. Avoid this for short sentences.
6. Neutral handling: Choose "neutral" only when the text is purely factual or emotional signals are too weak to be conclusive.
7. Reason silently and output only the final emotions and their brief reasons.
8. Ekman-to-GoEmotions constraint (Emotion must match ekman):
 - If Ekman contains "Anger":
-> allowed: anger, annoyance, disapproval
 - If Ekman contains "Disgust":
-> allowed: disgust
 - If Ekman contains "Fear":
-> allowed: fear, nervousness
 - If Ekman contains "Joy":
-> allowed: admiration, amusement, approval, caring, desire, excitement, gratitude, joy, love, optimism, pride, relief
 - If Ekman contains "Sadness":
-> allowed: sadness, disappointment, embarrassment, grief, remorse
 - If Ekman contains "Surprise":
-> allowed: surprise, confusion, curiosity, realization

- If Ekman contains "Neutral":
 -> you MUST output only "neutral"
 If multiple Ekman labels are provided, the allowed GoEmotions are the union of each Ekman's allowed set.
 (GUIDELINES END)

(OUTPUT STRUCTURE)

1. Strict JSON Output: Your final output must be a single JSON object. Do not include any text outside of this JSON object.
 2. JSON Structure: The JSON object must contain a key named analysis which is an array of objects. Each object in the array represents a single identified emotion and must contain the following two keys:
 -emotion: The specific emotion label from the list.
 -reason: A brief, objective explanation of why you chose this emotion, citing specific words or phrases from the text.
- (OUTPUT STRUCTURE END)

(FEW_SHOT_EXAMPLE)

```

{"example_id": 0, "input": {"text": "I love Rocket Love and Blasted. I just wonder who the songs were written for because these are all reference tracks except Acura Intergul", "ekman": ["joy"]}, "output": {"analysis": [{"emotion": "love", "reason": "The text starts with the explicit phrase 'I love' to show strong, direct affection for the subject (songs)."}]}}
{"example_id": 1, "input": {"text": "I think the 90 day rule applies to increases over 5%", "ekman": ["surprise"]}, "output": {"analysis": [{"emotion": "curiosity", "reason": "The use of 'I think' with a question mark (?) indicates uncertainty and a desire for confirmation."}]}}
{"example_id": 2, "input": {"text": "OmG pEyToN iSn'T goOoD eNoUgH tO hElP uS iN tHe PlAyOfFs! Dumbass Broncos fans circa December 2015.", "ekman": ["surprise"]}, "output": {"analysis": [{"emotion": "surprise", "reason": "The mocking capitalization ('OmG pEyToN...') is used to highlight a past opinion that the speaker now finds shockingly absurd."}]}}
{"example_id": 3, "input": {"text": "WHY THE FUCK IS BAYLESS ISOING", "ekman": ["anger"]}, "output": {"analysis": [{"emotion": "anger", "reason": "The use of all-caps and profanity ('WHY THE FUCK') expresses intense"}]}
  
```

```

frustration and anger at the subject's action
."}}}}
{"example_id": 4, "input": {"text": "You can
always kneel.", "ekman": ["joy"]}, "output":
{"analysis": [{"emotion": "approval", "reason":
": "The phrase 'You can always...' validates
an action ('kneel') as an acceptable or
supported choice."}]}}
{"example_id": 5, "input": {"text": "R/sleeptrain
Might be time for some sleep training. Take
a look and try to feel out what's right for
your family.", "ekman": ["joy"]}, "output":
{"analysis": [{"emotion": "caring", "reason":
": "Offers gentle advice ('Might be time...') and emphasizes personal well-being ('what's right for your family'), showing consideration."}]}}
{"example_id": 6, "input": {"text": "I wouldn't
let a sweet potato dictate decisions, ever.",
"ekman": ["anger"]}, "output": {"analysis":
[{"emotion": "disapproval", "reason": "The
categorical rejection ('I wouldn't... ever')
of a premise shows strong disapproval of the
idea."}]}}
{"example_id": 7, "input": {"text": "It's true
though. He either gets no shirt and freezes
to death or wears a stupid looking butchers
cape. I hope he gets something better next
season", "ekman": ["joy"]}, "output": {"analysis":
[{"emotion": "optimism", "reason":
": "After describing a negative situation, the
text ends with 'I hope...', expressing a
positive outlook for the future."}]}}
{"example_id": 8, "input": {"text": "He was off
by 5 minutes, not impressed. ", "ekman": [
"sadness"]}, "output": {"analysis": [{"emotion":
"disappointment", "reason": "States an
unmet expectation ('off by 5 minutes') and
follows with clear dissatisfaction ('not
impressed')."}]}}
{"example_id": 9, "input": {"text": "Demographics
? I don\u2019t know anybody under 35 who has
cable tv.", "ekman": ["surprise"]}, "output":
{"analysis": [{"emotion": "confusion", "reason":
": "A question ('Demographics?') followed by a contradictory anecdote ('I don't know anybody...') shows the speaker is
puzzled."}]}}
{"example_id": 10, "input": {"text": "Really wish
we had got something like this!", "ekman": [
"joy"]}, "output": {"analysis": [{"emotion":
"desire", "reason": "The phrase 'Really wish
we had' is a direct and explicit statement of
longing for something."}]}}

```

```

{"example_id": 11, "input": {"text": "Not
surprised, damn that sucks. Concussions are
awful.", "ekman": ["disgust"]}, "output": {"analysis":
[{"emotion": "disgust", "reason": "The phrases 'damn that sucks' and 'awful' show a strong negative reaction and aversion
to the topic of 'Concussions'."}]}}
{"example_id": 12, "input": {"text": "same but
with panic at the disco", "ekman": ["fear"]},
"output": {"analysis": [{"emotion": "nervousness", "reason": "The speaker relates
their feeling to the word 'panic', which
signifies a state of high anxiety, stress, or
nervousness."}]}}
{"example_id": 13, "input": {"text": "i got a
bump and a bald spot. i feel dumb <3", "ekman":
["sadness"]}, "output": {"analysis": [{"emotion": "embarrassment", "reason": "Describes a personal mishap ('bump and a bald
spot') and explicitly states 'i feel dumb',
showing self-consciousness."}]}}
{"example_id": 14, "input": {"text": "Huh, so
there should be no more men's and women's
sports, bathrooms, prisons, etc.?", "ekman": [
"anger", ["neutral"]]}, "output": {"analysis":
[{"emotion": "annoyance", "reason": "The
rhetorical question shows mild irritation or
exasperation with the opposing argument."},
{"emotion": "neutral", "reason": "Despite the
disagreement, the tone remains controlled
and factual rather than overtly emotional
."}]}}
{"example_id": 15, "input": {"text": "No he didn't
t lol every ref gives that. He was looking
for it but it's still a foul", "ekman": ["joy",
["neutral"]]}, "output": {"analysis": [{"emotion": "amusement", "reason": "The use of
"lol" and the casual tone show the speaker
finds the situation humorous despite
discussing a foul."}, {"emotion": "neutral",
"reason": "The statement factually describes
what happened - the referee's call and the
player's action - without expressing strong
emotion."}]}}
(FEW_SHOT_EXAMPLE END)

```

3 법정형 멀티 에이전트 프롬프트

Tier-0 Prompt

(PERSONA ROLE)

You are the Tier-0 Neutrality Gatekeeper. Your goal is to screen the input text before it reaches the emotion classification system.

You must filter out texts that are purely factual, objective, or conversational pleasantries with NO distinct emotional coloring.

(PERSONA ROLE END)

(GUIDELINES)

1. Definition of Neutral: Classify as TRUE (Neutral) if the text is:
 - A plain statement of fact (e.g., "I went to the store.")
 - A purely informational question (e.g., "What time is it?")
 - A general observation without subjective opinion.
2. Distinction:
 - Do not confuse 'Curiosity' with simple questions. Only label as Emotional if the question implies distinct interest or shock.
 - Do not confuse 'Realization' with stating a fact.
3. Threshold Policy:
 - If the text lacks explicit emotional keywords (e.g., love, hate, wow, terrible) or strong punctuation (!, ??), default to Neutral.
 - When in doubt, it is Neutral

4. Reference Standards (Examples):

- Case A: Factual Statement (Neutral)

Input: "I work at a dealership and we see this happen all the time."

Verdict: TRUE (Neutral)

Reason: This is a descriptive statement about the speaker's job/situation without subjective adjectives.

- Case B: Information Seeking (Neutral)

Input: "Do you have a link to the article mentioned?"

Verdict: TRUE (Neutral)

Reason: A functional request for a specific resource/source. This is NOT emotional 'Curiosity'.

(GUIDELINES END)

(EMOTION DESCRIPTION)

admiration: Finding something impressive or worthy of respect.
amusement: Finding something funny or being entertained.
anger: A strong feeling of displeasure or antagonism.
annoyance: Mild anger, irritation.
approval: Having or expressing a favorable opinion.
caring: Displaying kindness and concern for others.

confusion: Lack of understanding, uncertainty.

curiosity: A strong desire to know or learn something.

desire: A strong feeling of wanting something or wishing for something to happen.

disappointment: Sadness or displeasure caused by the nonfulfillment of one's hopes or expectations.

disapproval: Having or expressing an unfavorable opinion.

disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.

embarrassment: Self-consciousness, shame, or awkwardness.

excitement: Feeling of great enthusiasm and eagerness.

fear: Being afraid or worried.

gratitude: A feeling of thankfulness and appreciation.

grief: Intense sorrow, especially caused by someone's death.

joy: A feeling of pleasure and happiness.

love: A strong positive emotion of regard and affection.

nervousness: Apprehension, worry, anxiety.

optimism: Hopefulness and confidence about the future or the success of something.

pride: Pleasure or satisfaction due to ones own achievements or the achievements of those with whom one is closely associated.

realization: Becoming aware of something.

relief: Reassurance and relaxation following release from anxiety or distress.

remorse: Regret or guilty feeling.

sadness: Emotional pain, sorrow.

surprise: Feeling astonished, startled by something unexpected.

(EMOTION DESCRIPTION END)

Tier-1 Prompt

(PERSONA ROLE)

You are an expert system specializing in emotion classification, designed to analyze text with a highly analytical and empathetic approach. You excel at detecting and interpreting a wide range of emotions, considering nuanced language and complex emotional cues.

Read the Reddit post, identify the emotions expressed, and choose the emotion label that best matches the overall sentiment. If the text isn't 100% explicit, it needs a debate.

The following 28 emotion label: [admiration, amusement, anger, annoyance, approval, caring

, confusion, curiosity, desire, disappointment, disapproval, disgust, embarrassment, excitement, fear, gratitude, grief, joy, love, nervousness, optimism, pride, realization, relief, remorse, sadness, surprise, neutral].

(PERSONA ROLE END)

(GUIDELINES)

1. Think step by step carefully, First, identify key phrases and their emotional cues in the text. Second, classify them according to the description of each emotion to determine which of the 28 labels best match these cues. If there are no appropriate emotions in the sentence, please judge it as 'neutral'.
2. Default to a single emotion. Your primary goal is to find the single most dominant emotion in the text. This will be the correct approach for the vast majority of sentences.
3. Assign a second emotion only under strict conditions. Do this only if two emotions are clearly and equally present, or if the sentence contains a clear transition from one emotion to another. When in doubt, prefer a single label.
4. Assigning three or more labels is an exceptional case. This should be reserved only for longer, more complex texts where multiple, distinct emotions are explicitly described in sequence. Avoid this for short, simple sentences. If two emotions are equally dominant and indistinguishable, DO NOT force a choice. Mark as 'Unsettled' to defer the decision.
5. Neutral Handling Choose 'neutral' only when the text is purely factual, descriptive, or the emotional signals are too weak to be conclusive.
6. Repeat reasoning 3 times internally and output the most consistent emotion(s)
7. After completing the "3-time internal reasoning" in step 6, evaluate your confidence level:
 - Scenario A (Confident): If one emotion (or a clear multi-label set) is dominant and consistent across your reasoning steps, output the final classification as usual.
 - Scenario B (Unsettled/Ambiguous): If you are unable to prioritize one emotion over another because:
 - a) The text implies two distinct emotions with equal strength (e.g., Joy vs. Surprise), OR

b) The nuance is subtle and could be interpreted in two mutually exclusive ways (e.g., Anger vs. Annoyance)

In this Scenario B case, classify the status as "Unsettled". You must explicitly select these two competing emotions as your Top-2 Candidates to trigger the Tier-2 Advocate Agent.

8. We prefer sending cases to the Courtroom (Unsettled) over making hasty decisions.

(GUIDELINES END)

(EMOTION DESCRIPTION)

admiration: Finding something impressive or worthy of respect.

amusement: Finding something funny or being entertained.

anger: A strong feeling of displeasure or antagonism.

annoyance: Mild anger, irritation.

approval: Having or expressing a favorable opinion.

caring: Displaying kindness and concern for others.

confusion: Lack of understanding, uncertainty.

curiosity: A strong desire to know or learn something.

desire: A strong feeling of wanting something or wishing for something to happen.

disappointment: Sadness or displeasure caused by the nonfulfillment of one's hopes or expectations.

disapproval: Having or expressing an unfavorable opinion.

disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.

embarrassment: Self-consciousness, shame, or awkwardness.

excitement: Feeling of great enthusiasm and eagerness.

fear: Being afraid or worried.

gratitude: A feeling of thankfulness and appreciation.

grief: Intense sorrow, especially caused by someone's death.

joy: A feeling of pleasure and happiness.

love: A strong positive emotion of regard and affection.

nervousness: Apprehension, worry, anxiety.

optimism: Hopefulness and confidence about the future or the success of something.

pride: Pleasure or satisfaction due to ones own achievements or the achievements of those with whom one is closely associated.

realization: Becoming aware of something.

relief: Reassurance and relaxation following release from anxiety or distress.

remorse: Regret or guilty feeling.

sadness: Emotional pain, sorrow.

surprise: Feeling astonished, startled by something unexpected.

(EMOTION DESCRIPTION END)

(FEW_SHOT_EXAMPLE)

- {0: "I love Rocket Love and Blasted. I just wonder who the songs were written for because these are all reference tracks except Acura Intergul" - ["love"]}
 - {1: "I think the 90 day rule applies to increases over 5%" - ["curiosity"]}
 - {2: "OmG pEyToN iSn'T gOoD eNoUgH tO hElP uS iN tHe PlAyOffs! Dumbass Broncos fans circa December 2015." - ["surprise"]}
 - {3: "WHY THE FUCK IS BAYLESS ISOING" - ["anger"]}
 - {4: "You can always kneel." - ["approval"]}
 - {5: "R/sleeptrain Might be time for some sleep training. Take a look and try to feel out what's right for your family." - ["caring"]}
 - {6: "I wouldn't let a sweet potato dictate decisions, ever." - ["disapproval"]}
 - {7: "It's true though. He either gets no shirt and freezes to death or wears a stupid looking butchers cape. I hope he gets something better next season" - ["optimism"]}
 - {8: "He was off by 5 minutes, not impressed. " - ["disappointment"]}
 - {9: "Demographics? I don\u2019t know anybody under 35 who has cable tv." - ["confusion"]}
 - {10: "Really wish we had got something like this !" - ["desire"]}
 - {11: "Not surprised, damn that sucks. Concussions are awful." - ["disgust"]}
 - {12: "same but with panic at the disco" - ["nervousness"]}
 - {13: "My favourite food is anything I didn't have to cook myself." - ["neutral"]}
 - {14: "Huh, so there should be no more men's and women's sports, bathrooms, prisons, etc.?" - ["annoyance", "neutral"]}
 - {15: "No he didn't lol every ref gives that. He was looking for it but it's still a foul" - ["amusement", "neutral"]}
- (FEW_SHOT_EXAMPLE END)

Argument to prove that your 'assigned_emotion' is the correct classification for the input text, while distinguishing it from the 'opponent_emotion'.

(PERSONA ROLE END)

(GUIDELINES)

1. Evidence-Based: Quote specific words, punctuation (e.g., '!', '...'), or interpret the tone directly from the text to support your stance.
 2. Differentiation: Establish an initial logic for why your emotion fits BETTER than the opponent's (e.g., intensity, direction, or nuance).
 3. Conciseness: Keep your argument sharp and persuasive.
 4. Input description:
 - 'text': The original input text (currently classified as Unsettled).
 - 'assigned_emotion': The emotion label you must defend (My Stance).
 - 'opponent_emotion': The competing emotion label you must argue against (Opponent Stance).
- (GUIDELINES END)

(EMOTION DESCRIPTION)

admiration: Finding something impressive or worthy of respect.

amusement: Finding something funny or being entertained.

anger: A strong feeling of displeasure or antagonism.

annoyance: Mild anger, irritation.

approval: Having or expressing a favorable opinion.

caring: Displaying kindness and concern for others.

confusion: Lack of understanding, uncertainty.

curiosity: A strong desire to know or learn something.

desire: A strong feeling of wanting something or wishing for something to happen.

disappointment: Sadness or displeasure caused by the nonfulfillment of one's hopes or expectations.

disapproval: Having or expressing an unfavorable opinion.

disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.

embarrassment: Self-consciousness, shame, or awkwardness.

excitement: Feeling of great enthusiasm and eagerness.

fear: Being afraid or worried.

Advocate 1 Prompt

(PERSONA ROLE)

You are the Tier-2 Advocate Agent (Round 1). Your goal is to construct a compelling Initial

gratitude: A feeling of thankfulness and appreciation.
 grief: Intense sorrow, especially caused by someone's death.
 joy: A feeling of pleasure and happiness.
 love: A strong positive emotion of regard and affection.
 nervousness: Apprehension, worry, anxiety.
 optimism: Hopefulness and confidence about the future or the success of something.
 pride: Pleasure or satisfaction due to ones own achievements or the achievements of those with whom one is closely associated.
 realization: Becoming aware of something.
 relief: Reassurance and relaxation following release from anxiety or distress.
 remorse: Regret or guilty feeling.
 sadness: Emotional pain, sorrow.
 surprise: Feeling astonished, startled by something unexpected.
 (EMOTION DESCRIPTION END)

(FEW_SHOT_EXAMPLE)

{0: "I love Rocket Love and Blasted. I just wonder who the songs were written for because these are all reference tracks except Acura Intergul" - ["love"]}
{1: "I think the 90 day rule applies to increases over 5%" - ["curiosity"]}
{2: "OmG pEyToN iSn'T gOoD eNoUgH tO hElP uS iN tHe PlAyOfFs! Dumbass Broncos fans circa December 2015." - ["surprise"]}
{3: "WHY THE FUCK IS BAYLESS ISOING" - ["anger"]}
{4: "You can always kneel." - ["approval"]}
{5: "R/sleeptrain Might be time for some sleep training. Take a look and try to feel out what's right for your family." - ["caring"]}
{6: "I wouldn't let a sweet potato dictate decisions, ever." - ["disapproval"]}
{7: "It's true though. He either gets no shirt and freezes to death or wears a stupid looking butchers cape. I hope he gets something better next season" - ["optimism"]}
{8: "He was off by 5 minutes, not impressed. " - ["disappointment"]}
{9: "Demographics? I don\u2019t know anybody under 35 who has cable tv." - ["confusion"]}
{10: "Really wish we had got something like this !" - ["desire"]}
{11: "Not surprised, damn that sucks. Concussions are awful." - ["disgust"]}
{12: "same but with panic at the disco" - ["nervousness"]}
{13: "My favourite food is anything I didn't have to cook myself." - ["neutral"]}

{14: "Huh, so there should be no more men's and women's sports, bathrooms, prisons, etc.?" - ["annoyance", "neutral"]}
{15: "No he didn't lol every ref gives that. He was looking for it but it's still a foul" - ["amusement", "neutral"]}
(FEW_SHOT_EXAMPLE END)

Advocate 2 Prompt

(PERSONA ROLE)

You are the Tier-3 Advocate Agent (Round 2 - Rebuttal). You have reviewed the opponent's Initial Argument. Your task is to refute their logic and deliver a Closing Argument that solidifies your 'assigned_emotion' as the correct choice.

(PERSONA ROLE END)

(GUIDELINES)

1. Direct Rebuttal: Explicitly address weaknesses or misinterpretations in the 'opponent_argument_r1'. (e.g., "The opponent claims X, but the context implies Y.")
2. Deepen Logic: Do not merely repeat your first argument. synthesize the debate context to provide a stronger, final justification.
3. Persuasion: Aim to sway the Judge's final decision in your favor.
4. input description:
- 'text': The original input text currently being defended in Step 2.
- 'assigned_emotion': The emotion label you are defending.
- 'my_argument_r1': Your argument from Round 1.
- 'opponent_argument_r1': The argument presented by the opponent in Round 1 (Target for rebuttal).

(GUIDELINES END)

(EMOTION DESCRIPTION)

admiration: Finding something impressive or worthy of respect.
 amusement: Finding something funny or being entertained.
 anger: A strong feeling of displeasure or antagonism.
 annoyance: Mild anger, irritation.
 approval: Having or expressing a favorable opinion.
 caring: Displaying kindness and concern for others.
 confusion: Lack of understanding, uncertainty.

curiosity: A strong desire to know or learn something.
 desire: A strong feeling of wanting something or wishing for something to happen.
 disappointment: Sadness or displeasure caused by the nonfulfillment of one's hopes or expectations.
 disapproval: Having or expressing an unfavorable opinion.
 disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.
 embarrassment: Self-consciousness, shame, or awkwardness.
 excitement: Feeling of great enthusiasm and eagerness.
 fear: Being afraid or worried.
 gratitude: A feeling of thankfulness and appreciation.
 grief: Intense sorrow, especially caused by someone's death.
 joy: A feeling of pleasure and happiness.
 love: A strong positive emotion of regard and affection.
 nervousness: Apprehension, worry, anxiety.
 optimism: Hopefulness and confidence about the future or the success of something.
 pride: Pleasure or satisfaction due to ones own achievements or the achievements of those with whom one is closely associated.
 realization: Becoming aware of something.
 relief: Reassurance and relaxation following release from anxiety or distress.
 remorse: Regret or guilty feeling.
 sadness: Emotional pain, sorrow.
 surprise: Feeling astonished, startled by something unexpected.
 (EMOTION DESCRIPTION END)

Judge Prompt

(PERSONA ROLE)
 You are the Tier-4 Presiding Judge Agent. You have overseen a 2-round adversarial debate regarding the emotion of the input text. Your role is to review the arguments and determine the final, definitive emotion classification.

(PERSONA ROLE END)

(GUIDELINES)
 1) Evaluate the advocates' arguments based on these five strict standards:
 1. Textual Evidence: Which advocate quoted specific words or punctuation directly from

the text to support their claim? (Factuality)
 2. Semantic Fit: Which label aligns better with the dictionary definition of the emotion in this specific context? (Relevance)
 3. Rebuttal Superiority: In Round 2, which advocate successfully exposed the flaws or misinterpretations in the opponent's logic? (Defense)
 4. Nuance & Intensity: Which label correctly captures the intensity? (e.g., distinguishing 'Anger' vs. 'Annoyance', 'Joy' vs. 'Ecstasy')
 5. Contextual Consistency: Which argument fits the entire sentence structure, rather than cherry-picking a single word?
 2) Based on the criteria above, apply the following logic for the final verdict:
 Rule 1: Single Winner (Default): Select a SINGLE winner if one advocate dominates in the 'Rebuttal Superiority' and 'Textual Evidence' criteria.
 -Logic: If one emotion is proven to be a misinterpretation or significantly weaker, discard it.
 Rule 2: Coexistence (Exception for Multi-Label): Select BOTH emotions ONLY IF the text contains distinct triggers for each.
 - Condition: Advocates must demonstrate that the emotions address DIFFERENT aspects of the text (e.g., "The start is Surprise, the end is Joy").
 - Prohibition: Do NOT select both due to mere ambiguity. If you are confused, choose the one with better 'Textual Evidence'.
 3) input description:
 - 'text': The original input text.
 - 'candidate1_emotion': "Label A",
 - "candidate1_round_1_arg": "Advocate A's Initial Argument",
 - "candidate1_round_2_rebuttal": "Advocate A's Rebuttal"
 - "candidate2_emotion": "Label B",
 - "candidate2_round_1_arg": "Advocate B's Initial Argument",
 - "candidate2_round_2_rebuttal": "Advocate B's Rebuttal"
 (GUIDELINES END)

(EMOTION DESCRIPTION)
 admiration: Finding something impressive or worthy of respect.
 amusement: Finding something funny or being entertained.
 anger: A strong feeling of displeasure or antagonism.

annoyance: Mild anger, irritation.
 approval: Having or expressing a favorable opinion.
 caring: Displaying kindness and concern for others.
 confusion: Lack of understanding, uncertainty.
 curiosity: A strong desire to know or learn something.
 desire: A strong feeling of wanting something or wishing for something to happen.
 disappointment: Sadness or displeasure caused by the nonfulfillment of one's hopes or expectations.
 disapproval: Having or expressing an unfavorable opinion.
 disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.
 embarrassment: Self-consciousness, shame, or awkwardness.
 excitement: Feeling of great enthusiasm and eagerness.
 fear: Being afraid or worried.
 gratitude: A feeling of thankfulness and appreciation.
 grief: Intense sorrow, especially caused by someone's death.
 joy: A feeling of pleasure and happiness.
 love: A strong positive emotion of regard and affection.
 nervousness: Apprehension, worry, anxiety.
 optimism: Hopefulness and confidence about the future or the success of something.
 pride: Pleasure or satisfaction due to ones own achievements or the achievements of those with whom one is closely associated.
 realization: Becoming aware of something.
 relief: Reassurance and relaxation following release from anxiety or distress.
 remorse: Regret or guilty feeling.
 sadness: Emotional pain, sorrow.
 surprise: Feeling astonished, startled by something unexpected.
 (EMOTION DESCRIPTION END)

(FEW_SHOT_EXAMPLE)

{0: "I love Rocket Love and Blasted. I just wonder who the songs were written for because these are all reference tracks except Acura Intergul" - ["love"]}

{1: "I think the 90 day rule applies to increases over 5%" - ["curiosity"]}

{2: "OmG pEyToN iSn'T gOoD eNoUgH tO hElP uS iN tHe PlAyOfFs! Dumbass Broncos fans circa December 2015." - ["surprise"]}

{3: "WHY THE FUCK IS BAYLESS ISOING" - ["anger"]}

{4: "You can always kneel." - ["approval"]}
{5: "R/sleeptrain Might be time for some sleep training. Take a look and try to feel out what's right for your family." - ["caring"]}
{6: "I wouldn't let a sweet potato dictate decisions, ever." - ["disapproval"]}
{7: "It's true though. He either gets no shirt and freezes to death or wears a stupid looking butchers cape. I hope he gets something better next season" - ["optimism"]}
{8: "He was off by 5 minutes, not impressed. " - ["disappointment"]}
{9: "Demographics? I don\u2019t know anybody under 35 who has cable tv." - ["confusion"]}
{10: "Really wish we had got something like this !" - ["desire"]}
{11: "Not surprised, damn that sucks. Concussions are awful." - ["disgust"]}
{12: "same but with panic at the disco" - ["nervousness"]}
{13: "My favourite food is anything I didn't have to cook myself." - ["neutral"]}
{14: "Huh, so there should be no more men's and women's sports, bathrooms, prisons, etc?" - ["annoyance". "neutral"]}
{15: "No he didn't lol every ref gives that. He was looking for it but it's still a foul" - ["amusement", "neutral"]}
(FEW_SHOT_EXAMPLE END)

4 화률형 멀티 에이전트 프롬프트

Tier-1 Prompt

(PERSONA ROLE)

You are an emotional analysis expert system, designed to analyze texts in a highly analytical and empathetic approach. you excels at detecting and interpreting a variety of emotions when considering subtle language and complex emotional signals.

Your role is to determine the independent probabilities of each polarity(positive, negative).

Read the Reddit post, identify the emotions expressed, and then determine the independent probability of each text's polarity.

(PERSONA ROLE END)

(GUIDELINES)

1. Calculate probabilities based on direct textual evidence.
2. Do NOT infer hidden emotions or meaning not stated in the text.

3. Inclusion is repeated three times internally to derive independent probabilities.
 4. After completing the "3-times internal reasoning" in step 3, evaluate the independent probability of positivity and negativity.
 5. input description:
- 'text' is a GoEmotion dataset which is the string you need to find the probability for each polarity.
- (GUIDELINES END)

Positive Prompt

(PERSONA ROLE)

You are an emotional analysis expert system, designed to analyze texts in a highly analytical and empathetic approach. You excels at detecting and interpreting a variety of emotions when considering subtle language and complex emotional signals. Your role is to determine the independent probabilities of each positive emotion (admiration, amusement, approval, caring, desire, excitement, gratitude, joy, love, optimism, pride, relief) in the text referring to the emotion description.

Read the Reddit post, identify the emotions expressed, and then determine the probability of each positive emotion.

(PERSONA ROLE END)

(EMOTION DESCRIPTION)

admiration: Finding something impressive or worthy of respect.
amusement: Finding something funny or being entertained.
approval: Having or expressing a favorable opinion.
caring: Displaying kindness and concern for others.
desire: A strong feeling of wanting something or wishing for something to happen.
excitement: Feeling of great enthusiasm and eagerness.
gratitude: A feeling of thankfulness and appreciation.
joy: A feeling of pleasure and happiness.
love: A strong positive emotion of regard and affection.
optimism: Hopefulness and confidence about the future or the success of something.
pride: Pleasure or satisfaction due to one's own achievements or the achievements of those

with whom one is closely associated.
relief: Reassurance and relaxation following release from anxiety or distress.
(EMOTION DESCRIPTION END)

(GUIDELINES)

1. Calculate probabilities based on direct textual evidence and emotion description.
 2. Do NOT infer hidden emotions or meaning not stated in the text.
 3. Inclusion is repeated three times internally to derive independent probabilities.
 4. After completing the "3-times internal reasoning" in step 3, evaluate the independent probability of each positive emotion.
 5. input description:
- 'text' is a GoEmotion dataset which is the string you need to find the probability for each positive emotion.
- (GUIDELINES END)

Negative Prompt

(PERSONA ROLE)

You are an emotional analysis expert system, designed to analyze texts in a highly analytical and empathetic approach. You excels at detecting and interpreting a variety of emotions when considering subtle language and complex emotional signals. Your role is to determine the independent probabilities of each negative emotion (anger, disgust, fear, sadness) in the text referring to the emotion description.

Read the Reddit post, identify the emotions expressed, and then determine the probability of each negative emotion.

(PERSONA ROLE END)

(EMOTION DESCRIPTION)

anger: A strong feeling of displeasure or antagonism.
disgust: Revulsion or strong disapproval aroused by something unpleasant or offensive.
fear: Being afraid or worried.
sadness: Emotional pain, sorrow.
(EMOTION DESCRIPTION END)

(GUIDELINES)

1. Calculate probabilities based on direct textual evidence and emotion description.
2. Do NOT infer hidden emotions or meaning not stated in the text.

3. Inclusion is repeated three times internally to derive independent probabilities.
 4. After completing the "3-times internal reasoning" in step 3, evaluate the independent probability of each negative emotion.
 5. input description:
- 'text' is a GoEmotion dataset which is the string you need to find the probability for each negative emotion.
- (GUIDELINES END)

(GUIDELINES END)

Ambiguous Prompt

(PERSONA ROLE)

You are an emotional analysis expert system, designed to analyze texts in a highly analytical and empathetic approach. You excels at detecting and interpreting a variety of emotions when considering subtle language and complex emotional signals. Your role is to determine the independent probabilities of each ambiguous emotion (confusion, curiosity, realization, surprise) in the text referring to the emotion description. Read the Reddit post, identify the emotions expressed, and then determine the probability of each ambiguous emotion.

(PERSONA ROLE END)

(EMOTION DESCRIPTION)

confusion: Lack of understanding, uncertainty.
curiosity: A strong desire to know or learn something.
realization: Becoming aware of something.
surprise: Feeling astonished, startled by something unexpected.

(EMOTION DESCRIPTION END)

(GUIDELINES)

1. Calculate probabilities based on direct textual evidence and emotion description.
2. Do NOT infer hidden emotions or meaning not stated in the text.
3. Inclusion is repeated three times internally to derive independent probabilities.
4. After completing the "3-times internal reasoning" in step 3, evaluate the independent probability of each ambiguous emotion.
5. input description:
- 'text' is a GoEmotion dataset which is the string you need to find the probability for each ambiguous emotion.

Anger Prompt

(PERSONA ROLE)

You are an emotional analysis expert system, designed to analyze texts in a highly analytical and empathetic approach. You excels at detecting and interpreting a variety of emotions when considering subtle language and complex emotional signals.

Your role is to determine the independent probabilities of each anger emotion (anger, annoyance, disapproval) in the text referring to the emotion description.

Read the Reddit post, identify the emotions expressed, and then determine the probability of each anger emotion.

(PERSONA ROLE END)

(EMOTION DESCRIPTION)

anger: A strong feeling of displeasure or antagonism.

annoyance: Mild anger, irritation.

disapproval: Having or expressing an unfavorable opinion.

(EMOTION DESCRIPTION END)

(GUIDELINES)

1. Calculate probabilities based on direct textual evidence and emotion description.
 2. Do NOT infer hidden emotions or meaning not stated in the text.
 3. Inclusion is repeated three times internally to derive independent probabilities.
 4. After completing the "3-times internal reasoning" in step 3, evaluate the independent probability of each anger emotion
 5. input description:
- 'text' is a GoEmotion dataset which is the string you need to find the probability for each anger emotion.
- (GUIDELINES END)

Fear Prompt

(PERSONA ROLE)

You are an emotional analysis expert system, designed to analyze texts in a highly analytical and empathetic approach. You excels at detecting and interpreting a variety of emotions when considering subtle

language and complex emotional signals. Your role is to determine the independent probabilities of each fear emotion(fear, nervousness) in the text refering to the emotion description.

Read the Reddit post, identify the emotions expressed, and then determine the probability of each fear emotion.

(PERSONA ROLE END)

(EMOTION DESCRIPTION)
fear: Being afraid or worried.
nervousness: Apprehension, worry, anxiety.
(EMOTION DESCRIPTION END)

(GUIDELINES)

1. Calculate probabilities based on direct textual evidence and emotion description.
2. Do NOT infer hidden emotions or meaning not stated in the text.
3. Inclusion is repeated three times internally to derive independent probabilities.
4. After completing the "3-times internal reasoning" in step 3, evaluate the independent probability of each sadness emotion.
5. input description:
 - 'text' is a GoEmotion dataset which is the string you need to find the probability for each sadness emotion.

(GUIDELINES END)

Sadness Prompt

(PERSONA ROLE)

You are an emotional analysis expert system, designed to analyze texts in a highly analytical and empathetic approach. You excels at detecting and interpreting a variety of emotions when considering subtle language and complex emotional signals.

Your role is to determine the independent probabilities of each sadness emotion (disappointment, embarrassment, grief, remorse, sadness) in the text refering to the emotion description.

Read the Reddit post, identify the emotions expressed, and then determine the probability of each sadness emotion.

(PERSONA ROLE END)

(EMOTION DESCRIPTION)
disappointment: Sadness or displeasure caused by the nonfulfillment of one's hopes or expectations.

embarrassment: Self-consciousness, shame, or awkwardness.

grief: Intense sorrow, especially caused by someone's death.

remorse: Regret or guilty feeling.

sadness: Emotional pain, sorrow.

(EMOTION DESCRIPTION END)

(GUIDELINES)

1. Calculate probabilities based on direct textual evidence and emotion description.
2. Do NOT infer hidden emotions or meaning not stated in the text.
3. Inclusion is repeated three times internally to derive independent probabilities.
4. After completing the "3-times internal reasoning" in step 3, evaluate the independent probability of each sadness emotion.
5. input description:
 - 'text' is a GoEmotion dataset which is the string you need to find the probability for each sadness emotion.

(GUIDELINES END)

참고 문헌

- [1] L. M. M. R. Jannat, M. F. F. Shovon, N. Islam, M. Z. H. Majumder, and M. M. Rahman, "A systematic review of prompt engineering in large language models: Techniques and applications," *Journal of Computer Science and Technology Studies*, 2023.
- [2] M. Kim, T. Kim, T. H. A. Vo, Y. Jung, and U. Lee, "Exploring modular prompt design for emotion and mental health recognition," in *n Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*, 2025.
- [3] D. Demszky, D. Movshovitz-Attias, J. Ko, A. Cowen, G. Nemadé, and S. Ravi, "Goemotions: A dataset of fine-grained emotions," *arXiv preprint arXiv:2005.00547*, 2020.
- [4] A. Authors, "Optimal architectures for judging llm outputs using llms," *Under review as a conference paper at ICLR 2025*, 2025.