* **감성분석(Sentiment Analysis)**

**지도학습기반의 분석 :** 지도학습은 학습데이터와 타깃 레이블 값을 기반으로 감정 분석 학습을 수행한 뒤 이를 기반으로 다른 데이터의 감성 분석을 예측하는 방법으로 일반적인 텍스트 기반의 분류와 거의 동일

**감성 어휘사전을 이용한 분석 :** 감성 어휘 사전은 감성 분석을 위한 용어와 문맥에 대한 다양한 정보를 가지고 있음. 이를 이용해 긍정, 부정 감성 여부를 판단

문서 내 텍스트가 나타내는 여러가지 주관적인 단어와 문맥을 기반으로 감성(Sentiment)수치를 계산하는 방법

* **감성 어휘사전기반의 감성분석**

**SentiWordNet :** NLTK 패키지의 WordNet과 유사하게 감성 단어 전용의 WordNet을 구현한 것

SentiWordNet은 Synset별로 3가지 감성 점수(sentiment score)을 할당함

긍정 감성 지수, 부정 감성 지수, 객관성 지수

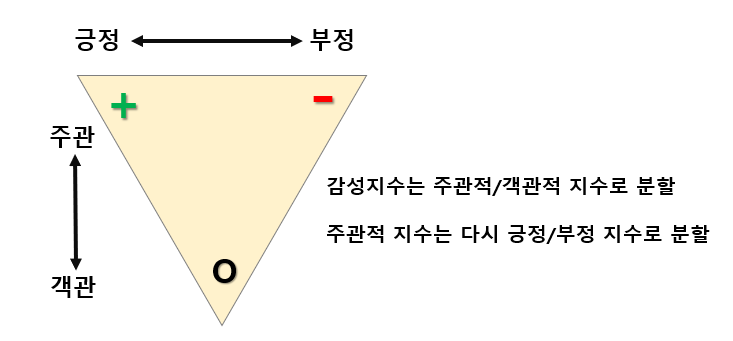
문장별로 단어들의 **긍정/부정 감성지수를 합산하여 최종 감성지수를 계산** 후 **긍정, 부정 결정**

**VADER :** 주로 소셜미디어 텍스트에 대한 감성 분석을 제공하기 위한 패키지

뛰어난 감성 분석 결과를 제공, 비교적 빠른 수행시간, 대용량 텍스트 데이터에 유용한 패키지

**Pattern :** 예측 성능 측면에서 가장 주목받는 패키지

* **SentiWordNet의 감성지수**



* **SentiWordNet을 이용한 감성 분석**

1. 문서(Document)를 문장(Sentence) 단위로 분해
2. 다시 문장을 단어(Word) 단위로 토큰화하고 품사 태깅
3. 품사 태깅된 단어 기반으로 synset 객체와 senti\_synset 객체를 생성
4. Senti\_synset에서 긍정 감성/ 부정 감성 지수를 구하고 이를 모두 합산해 특정 임계치 값 이상일 때 긍정 감성, 그렇지 않을 때 부정 감성으로 결정

* **VADER**

1. VADER는 소셜 미디어의 감성 분석 용도로 만들어진 룰 기반의 Lexicon
2. VADER는 SentimentIntensityAnalyzer 클래스를 이용해 쉽게 감성분석을 제공
3. SentimentIntensityAnalyzer 객체를 생성한 뒤 문서별로 polarity\_scores()메서드를 호출해 감성점수를 구한 뒤, 해당 문서의 감성 점수가 특정 임계값 이상이면 긍정, 아니면 부정
4. ‘neg’는 부정 ‘neu’는 중립 ‘pos’는 긍정

compound는 neg,neu,pos score를 적절히 조합해 -1~1 감성지수를 표현한 값

1. compound score를 기반으로 부정, 긍정 여부 결정

0.1 이상이면 긍정 그 이하이면 부정, 임계 값 조절 가능