[Labeling 예시] 분류가 잘된 case

|  |
| --- |
| **neg** |
| 🡪 영어 ‘so so’ 들어가도 분류 good!    🡪 neg\_prob가 0.5로 꽤 good! |
| **Neu** |
| -> ‘민들레 홀씨’라는 비유적인 표현 들어가도 labeling이 잘 되었음      🡪 부정단어와 긍정단어가 같이 있을 때, neu로 잘 labeling 되었음 |
| **Pos** |
| --> excellent!! |

[Labeling 관련 오류 정리]

|  |  |  |
| --- | --- | --- |
|  | **Opinion**  **(엑셀에서는 weird label 열)** | **H.F model** |
|  | ?? | neg |
|  | neu | neg |
| VS. | pos | neg |
|  | pos | neg |
|  | neu | neg |
|  | Neu? Pos? | neg |
| VS.    🡪 neg\_prob가 0.5로 꽤 good!  🡪 ‘터널의 끝이 보인다’를 neu로 예측 | Neu?  Pos?  혹은 neg의 확률이 조금 더 작아져야 하지 않을까? | neg |
| VS.    🡪 문자의 길이가 길어졌을 때, 잘 예측되는 것도 있지만, 잘 안되는 것(“설상가상(雪上加霜): 유럽 및 일본의 영향. 그러나 저평가 상태”를 positive로 예측)있음 | neg | neu |
|  | pos | neu |
|  | Pos | neu |
|  | Neg | pos |
|  | Neg | pos |

**[엑셀에서는 위의 Opinion을 weird\_label로 표시함]**

Opinion이 neg(부정)일 때 --> weird\_label = -1

Opinion이 neu(중립)일 때 --> weird\_label = 0

Opinion이 pos(긍정)일 때 --> weird\_label = 1