**타이레놀(1) & 후시딘(2)**

트위터, 코모란, 한나눔 비교

|  |  |
| --- | --- |
| URL(1) | |
| << https://www.tylenol.co.kr/products/tylenol-500mg>> | |
| 트위터 | ('<<', 'Punctuation'),  ('https://www.tylenol.co.kr/products/tylenol-500mg', 'URL'),  ('>>', 'Punctuation') |
| 비고 | 트위터는 URL을 정확하게 인식하여 분류함을 알 수 있다. |
| 코모란 | ('<', 'SS'), ('<', 'SS'),  ('https', 'SL'), (':', 'SP'), ('/', 'SP'), ('/', 'SP'),  ('www', 'SL'), ('.', 'SF'), ('tylenol', 'SL'), ('.co', 'NNP'),  ('.kr', 'NNP'), ('/', 'SP'), ('products', 'SL'), ('/', 'SP'),  ('tylenol', 'SL'), ('-', 'SS'), ('500', 'SN'), ('mg', 'SL'),  ('>', 'SS'), ('>', 'SS') |
| 비고 | ‘<<, >>’는 SS(따옴표, 괄호표, 줄표)로 잘 나눠졌지만,  나머지 URL은 SL(외국어), SP(쉼표, 콜론, 빗금), SF(마침표, 물음표) 등으로 세세하게 나눠졌음을 알 수 있다. |
| 한나눔 | ('<<', 'S'),  ('https', 'F'), ('://', 'S'), ('www', 'F'), ('.', 'S'),  ('tylenol', 'F'), ('.', 'S'), ('co', 'F'), ('.', 'S'),  ('kr', 'products'), ('/tylenol-500mg', 'N'),  ('>>', 'S') |
| 비고 | ‘<<, >>’는 S(기호)로 잘 표시되었고 URL부분은 S(기호)와 F(외국어)로 나눠졌다. |

|  |  |
| --- | --- |
| URL(2) | |
| << http://www.dong-wha.co.kr/product/content.asp?t\_idx=85&t\_page=1&d=&b=20&s=23 >> | |
| 트위터 | ('<<', 'Punctuation'),  ('http://www.dong-wha.co.kr/product/content.asp?t\_idx=85&t\_page=1&d=&b=20&s=23', 'URL'), ('>>', 'Punctuation') |
| 비고 | 트위터는 URL을 정확하게 인식하여 분류함을 알 수 있다. |
| 코모란 | ('<', 'SS'), ('<', 'SS'), ('http', 'SL'), (':', 'SP'), ('/', 'SP'), ('/', 'SP'),  ('www', 'SL'), ('.', 'SF'), ('dong', 'SL'), ('-', 'SS'), ('wha', 'SL'),  ('.co', 'NNP'), ('.kr', 'NNP'), ('/', 'SP'), ('product', 'SL'), ('/', 'SP'),  ('content', 'SL'), ('.', 'SF'), ('asp', 'SL'), ('?', 'SF'), ('t', 'SL'),  ('\_', 'SW'), ('idx', 'SL'), ('=', 'SW'), ('85', 'SN'), ('&', 'SW'), ('t', 'SL'),  ('\_', 'SW'), ('page', 'SL'), ('=', 'SW'), ('1', 'SN'), ('&', 'SW'), ('d', 'SL'),  ('=', 'SW'), ('&', 'SW'), ('b', 'SL'), ('=', 'SW'), ('20', 'SN'), ('&', 'SW'),  ('s', 'SL'), ('=', 'SW'), ('23', 'SN'), ('>', 'SS'), ('>', 'SS') |
| 비고 | ‘<<, >>’는 SS(따옴표, 괄호표, 줄표)로 잘 나눠졌지만,  나머지 URL은 SL(외국어), SP(쉼표, 콜론, 빗금), SF(마침표, 물음표),  SW(기타 기호), SN(숫자)등으로 세세하게 나눠졌음을 알 수 있다. |
| 한나눔 | ('<<', 'S'), ('http', 'F'), ('://', 'S'), ('www', 'F'), ('.', 'S'),  ('dong', 'F'), ('-', 'S'), ('wha', 'F'), ('.', 'S'), ('co', 'F'), ('.', 'S'),  ('kr', 'F'), ('/', 'S'), ('product', 'F'), ('/', 'S'), ('content', 'F'),  ('.', 'S'), ('asp', 'F'), ('?', 'S'), ('t\_idx=85&t\_page=1&d=&b=20&s=23', 'N'),  ('>>', 'S') |
| 비고 | ‘<<, >>’는 S(기호)로 잘 표시되었고 URL부분은 S(기호)와 F(외국어) 그리고  길게 한 묶음으로 묶여져 있는 부분은 N(체언)으로 나눠졌다. |

|  |  |
| --- | --- |
| 명사 | |
| 아세트아미노펜, 단일성분, 무카페인, 무색소의 빠르고 마음편한 해열 진통제 | |
| 트위터 | ('아세트아미노펜', 'Noun'), (',', 'Punctuation'),  ('단일', 'Noun'), ('성분', 'Noun'), (',', 'Punctuation'),  ('무', 'Noun'), ('카페인', 'Noun'), (',', 'Punctuation'),  ('무', 'Noun'), ('색소', 'Noun'), ('의', 'Josa'),  ('빠르고', 'Adjective'), ('마음', 'Noun'), ('편한', 'Adjective'),  ('해열', 'Noun'), ('진통제', 'Noun') |
| 비고 | 주요 성분인 아세트아미노펜을 확실하게 나눠줬으며, 단일성분, 무카페인, 무색소 또한 확실하고 제대로 나눠졌음을 알 수 있다. |
| 코모란 | ('아세트아미노펜', 'NNP'), (',', 'SP'),  ('단일', 'NNG'), ('성분', 'NNG'), (',', 'SP'),  ('무', 'XPN'), ('카페인', 'NNP'), (',', 'SP'),  ('무', 'XPN'), ('색소', 'NNP'), ('의', 'JKG'),  ('빠르', 'VA'), ('고', 'EC'), ('마음', 'NNG'), ('편하', 'VA'),  ('ㄴ', 'ETM'), ('해열', 'NNG'), ('진통제', 'NNP') |
| 비고 | 명사에 관한 부분에 대해서는 트위터와 유사하게 분류되었다.  하지만, 트위터는 크게 명사로만 정의되어 있는 반면에, 코모란은  NNG(일반 명사), NNP(고유 명사), XPN(체언 접두사) 등으로 조금 더 자세하게 분류되어 있다. |
| 한나눔 | ('아세트아미노펜', 'N'), (',', 'S'), ('단일성분', 'N'), (',', 'S'),  ('무카페인,', 'N'), ('무', 'N'), ('색소', 'N'), ('의', 'J'), ('빠르', 'P'), ('고', 'E'), ('마음편한', 'N'), ('해열', 'N'), ('진통제', 'N') |
| 비고 | 나눠져야 할 부분이 확실하게 나눠지지 않았다. |

|  |  |
| --- | --- |
| 명사 | |
| 후시딘은 ‘퓨시드 나트륨’이라는 항생 물질이 주 성분으로 세균 증식을 막을 수는 있지만 | |
| 트위터 | ('후시딘', 'Noun'), ('은', 'Josa'), ('‘', 'Foreign'), ('퓨', 'Noun'),  ('시드', 'Noun'), ('나트륨', 'Noun'), ('’', 'Punctuation'), ('이라는', 'Josa'),  ('항생', 'Noun'), ('물질', 'Noun'), ('이', 'Josa'), ('주', 'Noun'),  ('성분', 'Noun'), ('으로', 'Josa'), ('세균', 'Noun'), ('증식', 'Noun'),  ('을', 'Josa'), ('막', 'Noun'), ('을', 'Josa'), ('수', 'Noun'), ('는', 'Josa'),  ('있지만', 'Adjective') |
| 비고 | 위의 모든 명사들이 제대로 Noun(명사)로 표현되어 분류되어 있음을 알 수 있다. |
| 코모란 | ('후시', 'NNP'), ('딘', 'NNP'), ('은', 'JX'), ('‘퓨시드', 'NA'),  ('나트륨', 'NNP'), ('’', 'SS'), ('이', 'VCP'), ('라는', 'ETM'), ('항생', 'NNG'),  ('물질', 'NNG'), ('이', 'JKS'), ('주', 'MM'), ('성분', 'NNG'), ('으로', 'JKB'),  ('세균', 'NNP'), ('증식', 'NNG'), ('을', 'JKO'), ('막', 'VV'), ('을', 'ETM'),  ('수', 'NNB'), ('는', 'JX'), ('있', 'VV'), ('지만', 'EC') |
| 비고 | ‘후시’,’딘’,’나트륨’,’세균’은 NNP(고유 명사)로 분류되었고,  ‘항생’,’물질’,’성분’,’증식’은 NNG(일반 명사)로 분류되었으며,  ‘퓨시드’는 앞의 작은따옴표랑 분리가 되지 못하고 NA(분석불능범주)로 분류되었다.  트위터가 코모란보다 더 잘 분류되었다고 볼 수 있다. |
| 한나눔 | ('후시딘', 'N'), ('은', 'J'), ('‘퓨시드', 'N'), ('나트륨’', 'N'), ('이', 'J'),  ('라는', 'E'), ('항생', 'N'), ('물질', 'N'), ('이', 'J'), ('주', 'N'),  ('성분', 'N'), ('으로', 'J'), ('세균', 'N'), ('증식', 'N'), ('을', 'J'),  ('막', 'N'), ('을', 'J'), ('수', 'N'), ('는', 'J'), ('있', 'P'), ('지만', 'E') |
| 비고 | 트위터랑 유사하게 모든 명사들이 N(체언)으로 분류되어 있음을 알 수 있다. |

|  |  |
| --- | --- |
| 형용사(1-1) | |
| 이러한 중대한 피부 반응은 치명적일 수 있다. | |
| 트위터 | ('이러한', 'Adjective'), ('중대한', 'Adjective'), ('피부', 'Noun'),  ('반응', 'Noun'), ('은', 'Josa'), ('치명', 'Noun'), ('적', 'Suffix'),  ('일', 'Noun'), ('수', 'Noun'), ('있다', 'Adjective'),  ('.', 'Punctuation') |
| 비고 | ‘이러한’, ‘중대한’같은 형용사들이 제대로 ‘Adjective’로 분류되어 있다. |
| 코모란 | ('이러', 'XR'), ('하', 'XSA'), ('ㄴ', 'ETM'), ('중대', 'XR'), ('하', 'XSA'),  ('ㄴ', 'ETM'), ('피부', 'NNP'), ('반응', 'NNG'), ('은', 'JX'),  ('치명', 'NNG'), ('적', 'XSN'), ('이', 'VCP'), ('ㄹ', 'ETM'),  ('수', 'NNB'), ('있', 'VX'), ('다', 'EF'), ('.', 'SF') |
| 비고 | 코모란에서는 XR(어근), XSA(형용사 파생 접미사), ETM(관형형 전성 어미)로 세세하게 나눠져 있으며, VA(형용사)로 언급이 되어 있지는 않습니다. |
| 한나눔 | ('이러하', 'P'), ('ㄴ', 'E'), ('중대한', 'N'), ('피부', 'N'), ('반응', 'N'),  ('은', 'J'), ('치명적', 'N'), ('이', 'J'), ('ㄹ', 'E'), ('수', 'N'),  ('있', 'P'), ('다', 'E'), ('.', 'S') |
| 비고 | ‘이러한’은 그게 동사, 형용사, 보조 용언을 모두 포함하는 P(용언)과 E(어미)로 이루어져있으며, ‘중대한’은 형용사가 아닌 N(명사)로 분류되어 있기 때문에, 제대로 분리가 되지 않았음을 알 수 있습니다. |

|  |  |
| --- | --- |
| 형용사(1-2) | |
| 속방정으로 4시간동안 빠르고 강한 약효과를 발휘 | |
| 트위터 | ('속', 'Modifier'), ('방정', 'Noun'), ('으로', 'Josa'),  ('4시간', 'Number'), ('동안', 'Foreign'), ('빠르고', 'Adjective'),  ('강한', 'Adjective'), ('약효', 'Noun'), ('과를', 'Josa'), ('발휘', 'Noun') |
| 비고 | 형용사(1)과 동일하게 ‘빠르고’, ‘강한’을 크게 Adjective(형용사)로 분류하였다. |
| 코모란 | ('속방', 'NNG'), ('정', 'NNP'), ('으로', 'JKB'), ('4', 'SN'),  ('시간', 'NNP'), ('동안', 'NNP'), ('빠르', 'VA'), ('고', 'EC'),  ('강하', 'VA'), ('ㄴ', 'ETM'), ('약', 'MM'), ('효과', 'NNG'), ('를', 'JKO'),  ('발휘', 'NNG') |
| 비고 | ‘빠르’오아 ‘강하’을 VA(형용사)로 표현하였으며, 두 단어를 연결해주는 ‘고’를 EC(연결어미)로 잡아주었고, ‘강한’은 ETM(관형형 전성 어미)와 같이 분류해주었다. 트위터랑은 다르게 매우 세세하게 분류한다. |
| 한나눔 | ('속방정으로', 'N'), ('4시간', 'N'), ('동안', 'X'), ('빠르', 'P'),  ('고', 'E'), ('강하', 'P'), ('ㄴ', 'E'), ('약효과', 'N'), ('를', 'J'),  ('발휘', 'N') |
| 비고 | ‘빠르고’, ‘강한’ 모두 형용사(1)과 동일하게 P(용언)과 E(어미)로 표현되었다. |

|  |  |
| --- | --- |
| 형용사(2) | |
| 피부의 깊은 곳에 응어리가 생긴 상태 | |
| 트위터 | ('피부', 'Noun'), ('의', 'Josa'), ('깊은', 'Adjective'), ('곳', 'Noun'),  ('에', 'Josa'), ('응어리', 'Noun'), ('가', 'Josa'), ('생긴', 'Verb'),  ('상태', 'Noun') |
| 비고 | ‘깊은’ 부분이 Adjective(형용사)로 제대로 분류가 되었다고 볼 수 있다. |
| 코모란 | ('피부', 'NNP'), ('의', 'JKG'), ('깊', 'VA'), ('은', 'ETM'), ('곳', 'NNG'),  ('에', 'JKB'), ('응어리', 'NNG'), ('가', 'JKS'), ('생기', 'VV'), ('ㄴ', 'ETM'),  ('상태', 'NNG') |
| 비고 | ‘깊은’이 VA(형용사)인 ‘깊’과 ETM(관형형 전성 어미)인 ‘은’으로 제대로 잘 분류가 되었다고 볼 수 있다. |
| 한나눔 | ('피부', 'N'), ('의', 'J'), ('깊', 'P'), ('은', 'E'), ('곳', 'N'), ('에', 'J'), ('응어리', 'N'), ('가', 'J'), ('생기', 'P'), ('ㄴ', 'E'), ('상태', 'N') |
| 비고 | 코모란과 유사하게 ‘깊은’이 P(용언)인 ‘깊’과 E(어미)인 ‘은’으로 분류가 되었다고 볼 수 있다. |

|  |  |
| --- | --- |
| 부사(1) | |
| 매일 세잔 이상 정기적으로 술을 마시는 사람 | |
| 트위터 | ('매일', 'Noun'), ('세잔', 'Noun'), ('이상', 'Noun'), ('정기', 'Noun'),  ('적', 'Suffix'), ('으로', 'Josa'), ('술', 'Noun'), ('을', 'Josa'),  ('마시는', 'Verb'), ('사람', 'Noun') |
| 비고 | 정기적으로 또한 부사이지만, Noun(명사), Suffix(접사), Josa(조사)로 나누어서 분류되었으며, ‘매일’은 부사로 사용되었지만 Noun(명사)로 표시되었기 때문에, 분류가 제대로 이루어졌다고 보기는 어렵다. |
| 코모란 | ('매일', 'MAG'), ('세', 'MM'), ('잔', 'NNP'), ('이상', 'NNG'),  ('정기', 'NNG'), ('적', 'XSN'), ('으로', 'JKB'), ('술', 'NNG'),  ('을', 'JKO'), ('마시', 'VV'), ('는', 'ETM'), ('사람', 'NNG') |
| 비고 | 트위터와 유사하게 ‘정기적으로’는 NNG(일반 명사)와 XSN(명사파생 접미사), JKB(부사격 조사)로 나눠서 분류되었지만, ‘매일’은 MAG(일반 부사)로 정확하게 분류가 되어 제대로 분류가 이루어졌다고 볼 수 있습니다. |
| 한나눔 | ('매일', 'M'), ('세잔', 'N'), ('이상', 'N'), ('정기적', 'N'), ('으로', 'J'),  ('술', 'N'), ('을', 'J'), ('마시', 'P'), ('는', 'E'), ('사람', 'N') |
| 비고 | ‘매일’을 관형사와 부사로 이루어진 M(수식언)으로 분류하였고, ‘정기적으로’는 앞의 두 방법과는 다르게 ‘정기적’을 하나의 명사로 묶어서 분류하였다. |

|  |  |
| --- | --- |
| 부사(2) | |
| 환부(질환 부위)를 깨끗이 하고 1일 1~2회 적당량을 넓게 펴발라 붙인다 | |
| 트위터 | ('환부', 'Noun'), ('(', 'Punctuation'), ('질환', 'Noun'), ('부위', 'Noun'), (')', 'Punctuation'), ('를', 'Noun'), ('깨끗이', 'Adverb'), ('하고', 'Verb'),  ('1일', 'Number'), ('1~2회', 'Number'), ('적당량', 'Noun'), ('을', 'Josa'),  ('넓게', 'Adjective'), ('펴', 'Verb'), ('발라', 'Noun'), ('붙인다', 'Verb') |
| 비고 | ‘깨끗이’가 Adverb(부사)로 제대로 분류되었음을 볼 수 있다. |
| 코모란 | ('환부', 'NNG'), ('(', 'SS'), ('질환', 'NNG'), ('부위', 'NNG'), (')', 'SS'),  ('를', 'JKO'), ('깨끗이', 'MAG'), ('하', 'VV'), ('고', 'EC'), ('1', 'SN'),  ('일', 'NNB'), ('1', 'SN'), ('~', 'SO'), ('2', 'SN'), ('회', 'NNB'),  ('적', 'NNG'), ('당량', 'NNG'), ('을', 'JKO'), ('넓', 'VA'), ('게', 'EC'),  ('펴', 'VV'), ('어', 'EC'), ('바르', 'VV'), ('아', 'EC'), ('붙이', 'VV'),  ('ㄴ다', 'EF') |
| 비고 | ‘깨끗이’가 MAG(일반 부사)로 제대로 분류되었음을 볼 수 있다. |
| 한나눔 | ('환부(질환', 'N'), ('부위', 'N'), (')', 'S'), ('를', 'J'), ('깨끗', 'N'),  ('이', 'J'), ('하', 'P'), ('고', 'E'), ('1일', 'N'), ('1', 'N'), ('~', 'S'),  ('2회', 'N'), ('적당량', 'N'), ('을', 'J'), ('넓', 'P'), ('게', 'E'),  ('펴발라', 'N'), ('붙이', 'P'), ('ㄴ다', 'E') |
| 비고 | ‘깨끗이’가 부사가 아닌 N(체언)인 ‘깨끗’과 J(조사)인 ‘이’로 분류되었다.  즉, 위의 두 방법과 비교를 해보면, 분류가 제대로 되었다고 볼 수 없다. |

|  |  |
| --- | --- |
| 동사(1) | |
| 의사 또는 약사와 상의해야 한다. | |
| 트위터 | ('의사', 'Noun'), ('또는', 'Adverb'), ('약사', 'Noun'), ('와', 'Josa'), ('상의', 'Noun'), ('해야', 'Verb'), ('한다', 'Verb'), ('.', 'Punctuation') |
| 비고 | ‘상의해야 한다’가 Noun(명사)인 ‘상의’와 Verb(동사)인 ‘해야’와 ‘한다’로 분류되었다. |
| 코모란 | ('의사', 'NNG'), ('또는', 'MAJ'), ('약사', 'NNP'), ('와', 'JC'), ('상의', 'NNG'),  ('하', 'XSV'), ('아야', 'EC'), ('하', 'VX'), ('ㄴ다', 'EF'), ('.', 'SF') |
| 비고 | ‘상의해야’ 부분이 ‘상의’라는 NNG(일반명사)와 XSV(동사 파생 접미사)인 ‘하’와 EC(연결 어미)인 ‘아야’로 분류되었으며, ‘한다’는 VX(보조 용언)인 ‘하’와  EF(종결어미)인 ‘ㄴ다’로 분류되어서 동사를 자세하게 나눠 표현했음을 알 수 있다. |
| 한나눔 | ('의사', 'N'), ('또는', 'M'), ('약사', 'N'), ('와', 'J'), ('상의해', 'N'),  ('이', 'J'), ('야', 'E'), ('하', 'P'), ('ㄴ다', 'E'), ('.', 'S') |
| 비고 | ‘한다’ 부분은 코모란 방법과 유사하게 P(용언)과 E(어미)로 분류되었지만,  ‘상의해야’ 부분은 ‘상의해’가 새로운 N(체언)으로 분류되었으며 ‘야’ 부분은  J(관계언)과 E(어미)로 분류되어 있으므로, 분석이 잘 되었다고 볼 수 없다. |

|  |  |
| --- | --- |
| 동사(2) | |
| 1일 3～4회 적당량을 환부(질환 부위)에 부드럽게 바른다. | |
| 트위터 | ('1일', 'Number'), ('3', 'Number'), ('∼', 'Foreign'), ('4회', 'Number'),  ('적당량', 'Noun'), ('을', 'Josa'), ('환부', 'Noun'), ('(', 'Punctuation'),  ('질환', 'Noun'), ('부위', 'Noun'), (')', 'Punctuation'), ('에', 'Josa'),  ('부드럽게', 'Adjective'), ('바른다', 'Verb'), ('.', 'Punctuation') |
| 비고 | ‘바른다’가 밑의 2 방법과는 다르게 하나의 묶음으로 Verb(동사)로 분류된다. |
| 코모란 | ('1', 'SN'), ('일', 'NNB'), ('3', 'SN'), ('～', 'SO'), ('4', 'SN'),  ('회', 'NNB'), ('적', 'NNG'), ('당량', 'NNG'), ('을', 'JKO'), ('환부', 'NNG'),  ('(', 'SS'), ('질환', 'NNG'), ('부위', 'NNG'), (')', 'SS'), ('에', 'JKB'),  ('부드럽', 'VA'), ('게', 'EC'), ('바르', 'VA'), ('ㄴ다', 'EF'), ('.', 'SF') |
| 비고 | ‘바른다’가 VA(형용사)인 ‘바르’와 EF(종결 어미)인 ‘ㄴ다’로 분류되어 있다.  즉, 동사가 아닌 형용사로 분류되었기 때문에 제대로 분석이 잘 수행되지 않았다. |
| 한나눔 | ('1일', 'N'), ('3～4회', 'N'), ('적당량', 'N'), ('을', 'J'), ('환부(질환', 'N'),  ('부위', 'N'), (')', 'S'), ('에', 'J'), ('부드럽', 'P'), ('게', 'E'),  ('바르', 'P'), ('ㄴ다', 'E'), ('.', 'S') |
| 비고 | 코모란과 비슷하게 P(용언)인 ‘바르’와 E(어미)인 ‘ㄴ다’으로 분류되어 있다.  하지만 한나눔의 단점은 P(용언)이 동사, 형용사, 보조 용언을 다 품고 있기 때문에 정확하게 무엇을 뜻하는지 알 수 없다. |

|  |  |
| --- | --- |
| 외국어(1) | |
| 스티븐스 - 존슨 증후군(SJS), 독성 표피 괴사용해(TEN)와 같은 중대한 피부 반응이 보고되었고, | |
| 트위터 | ('스티븐스', 'Noun'), ('-', 'Punctuation'), ('존슨', 'Noun'),  ('증후군', 'Noun'), ('(', 'Punctuation'), ('SJS', 'Alpha'),  ('),', 'Punctuation'), ('독성', 'Noun'), ('표피', 'Noun'), ('괴사', 'Noun'),  ('용해', 'Noun'), ('(', 'Punctuation'), ('TEN', 'Alpha'),  (')', 'Punctuation'), ('와', 'Verb'), ('같은', 'Adjective'),  ('중대한', 'Adjective'), ('피부', 'Noun'), ('반응', 'Noun'), ('이', 'Josa'),  ('보고', 'Noun'), ('되었고', 'Verb'), (',', 'Punctuation') |
| 비고 | ‘SJS’ 및 ‘TEN’ 모두 Alpha(알파벳)으로 잘 표현되어 있음을 알 수 있다. |
| 코모란 | ('스티븐스', 'NNP'), ('-', 'SS'), ('존슨', 'NNP'), ('증후군', 'NNP'),  ('(', 'SS'), ('SJS', 'SL'), (')', 'SS'), (',', 'SP'),  ('독성', 'NNG'), ('표피', 'NNP'), ('괴사', 'NNP'), ('용해', 'NNP'),  ('(', 'SS'), ('TEN', 'SL'), (')', 'SS'),  ('와', 'JC'), ('같', 'VA'), ('은', 'ETM'), ('중대', 'XR'), ('하', 'XSA'),  ('ㄴ', 'ETM'), ('피부', 'NNP'), ('반응', 'NNG'), ('이', 'JKS'), ('보고', 'NNG'),  ('되', 'XSV'), ('었', 'EP'), ('고', 'EC'), (',', 'SP') |
| 비고 | ‘SJS’와 ‘TEN’모두 SL(외국어)로 제대로 잘 표현되어 있음을 알 수 있다. |
| 한나눔 | ('스티븐스', 'N'), ('-', 'S'), ('존슨', 'N'), ('증후군(SJS),', 'N'),  ('독성', 'N'), ('표피', 'N'), ('괴사용해(TEN)', 'N'), ('와', 'J'), ('같', 'P'),  ('은', 'E'), ('중대한', 'N'), ('피부', 'N'), ('반응', 'N'), ('이', 'J'),  ('보고', 'N'), ('되', 'X'), ('었고', 'E'), (',', 'S') |
| 비고 | ‘SJS’와 ‘TEN’모두 따로 분류되지 않고, 앞의 단어와 같이 묶여서 N(체언)으로 분류되었다. |

|  |  |
| --- | --- |
| 외국어(2) | |
| 1g 중 : 퓨시드산나트륨(Sodium fusidate) | |
| 트위터 | ('1', 'Number'), ('g', 'Alpha'), ('중', 'Noun'), (':', 'Punctuation'),  ('퓨', 'Noun'), ('시드', 'Noun'), ('산', 'Noun'), ('나트륨', 'Noun'),  ('(', 'Punctuation'), ('Sodium', 'Alpha'), ('fusidate', 'Alpha'),  (')', 'Punctuation') |
| 비고 | ‘g’과 ‘Sodium’, ‘fusidate’ 모두 Alpha(알파벳)으로 제대로 잘 분류되었음을 알 수 있다. |
| 코모란 | ('1', 'SN'), ('g', 'SW'), ('중', 'NNB'), (':', 'SP'),  ('퓨시드산나트륨(Sodium', 'NA'), ('fusidate', 'SL'), (')', 'SS') |
| 비고 | ‘g’은 SW(기타 기호{논리/수학 기호, 화폐 기호})로 분류되었으며,  ‘Sodium’은 앞의 ‘퓨시드산나트륨’과 묶여서 제대로 분류가 되지 않았으며  ‘fusidate’는 SL(외국어)로 제대로 분류가 됬음을 알 수 있다. |
| 한나눔 | ('1g', 'N'), ('중', 'N'), (':', 'S'), ('퓨시드산나트륨(Sodium', 'N'),  ('fusidate', 'F'), (')', 'S') |
| 비고 | ‘g’은 ‘1’과 따로 분류가 되지 않고 묶여서 N(체언)으로 분류가 됬으며,  코모란과 유사하게 ‘Sodium’은 앞의 ‘퓨시드산나트륨’과 묶여서 분류가 되지 않았고  ‘fusidate’는 F(외국어)로 제대로 분류가 됬음을 알 수 있다. 즉, 외국어는 트위터에서 잘 분류가 되었다. |

|  |  |
| --- | --- |
| 숫자(1) | |
| 포장단위는 일반의약품 : 10정, 안전상비의약품 : 8정 | |
| 트위터 | ('포장', 'Noun'), ('단위', 'Noun'), ('는', 'Josa'), ('일반의약품', 'Noun'), (':', 'Punctuation'), ('10', 'Number'), ('정', 'Noun'), (',', 'Punctuation'), ('안전', 'Noun'), ('상비', 'Noun'), ('의약품', 'Noun'), (':', 'Punctuation'), ('8', 'Number'), ('정', 'Noun') |
| 비고 | 10, 8 등 숫자들은 따로 Number(숫자)로 분류가 제대로 되어있음을 알 수 있다. |
| 코모란 | ('포장', 'NNP'), ('단위', 'NNP'), ('는', 'JX'), ('일반의약품', 'NNP'),  (':', 'SP'), ('10', 'SN'), ('정', 'NNP'), (',', 'SP'),  ('안전', 'NNG'), ('상비', 'NNG'), ('의약품', 'NNP'), (':', 'SP'),  ('8', 'SN'), ('정', 'NNP') |
| 비고 | 10, 8 등 숫자들은 따로 SN(숫자)로 분류가 제대로 되어있음을 알 수 있다. |
| 한나눔 | ('포장단위', 'N'), ('는', 'J'), ('일반의약품', 'N'), (':', 'S'), ('10정', 'N'), (',', 'S'), ('안전상비의약품', 'N'), (':', 'S'), ('8정', 'N'), ('#', 'N') |
| 비고 | 10, 8 등 숫자들은 따로 분류되지 않고, 뒤에 한글과 같이 N(체언)으로 분류된다. |

|  |  |
| --- | --- |
| 숫자(2) | |
| 후시딘연고(5g/4.5g/10g) | |
| 트위터 | ('후시', 'Noun'), ('딘', 'Noun'), ('연고', 'Noun'), ('(', 'Punctuation'), ('5', 'Number'), ('g', 'Alpha'), ('/', 'Punctuation'), ('4.5', 'Number'),  ('g', 'Alpha'), ('/', 'Punctuation'), ('10', 'Number'), ('g', 'Alpha'),  (')', 'Punctuation') |
| 비고 | 정확하게 숫자들(‘5’, ‘4.5’, ‘10’)만 제대로 Number(숫자)로 분류가 되었음을 알 수 있다. ‘4.5’는 하나의 실수 형태로 분류되었다. |
| 코모란 | ('후시', 'NNP'), ('딘', 'NNP'), ('연고', 'NNG'), ('(', 'SS'),  ('5', 'SN'), ('g', 'SL'), ('/', 'SP'), ('4', 'SN'), ('.', 'SF'),  ('5', 'SN'), ('g', 'SL'), ('/', 'SP'), ('10', 'SN'), ('g', 'SL'),  (')', 'SS') |
| 비고 | 트위터랑 유사하게 숫자들만 따로 잘 SN(숫자)으로 분류가 되었지만,  ‘4.5’는 ‘4’와 ‘.’, ‘5’로 총 3가지로 따로 분류가 되었음을 알 수 있다. |
| 한나눔 | ('후시딘연고(5g', '4'), ('.', 'N'), ('5g', '10g') |
| 비고 | 어떻게 나뉘어 졌는지 모를 정도로 분류가 잘 되지 않았다. |

|  |  |
| --- | --- |
| 조사(1) | |
| 다음 질환에도 사용할 수 있습니다. | |
| 트위터 | ('다음', 'Noun'), ('질환', 'Noun'), ('에도', 'Josa'), ('사용', 'Noun'),  ('할', 'Verb'), ('수', 'Noun'), ('있습니다', 'Adjective'), ('.', 'Punctuation') |
| 비고 | ‘에도’가 Josa(조사)로 분류된 것을 알 수 있습니다. 트위터에서는 여러 가지의 조사를 그냥 크게 Josa 하나로 통일하여 사용합니다. |
| 코모란 | ('다음', 'NNG'), ('질환', 'NNG'), ('에', 'JKB'), ('도', 'JX'),  ('사용', 'NNG'), ('하', 'XSV'), ('ㄹ', 'ETM'), ('수', 'NNB'),  ('있', 'VX'), ('습니다', 'EF'), ('.', 'SF') |
| 비고 | ‘에도’가 JKB(부사격 조사)인 ‘에’와 JX(보조사)인 ‘도’로 나뉘어서 트위터와 한나눔보다 더 자세하게 분류되었음을 알 수 있다. |
| 한나눔 | ('다음', 'N'), ('질환', 'N'), ('에도', 'J'), ('사용', 'N'), ('하', 'X'),  ('ㄹ', 'E'), ('수', 'N'), ('있', 'P'), ('습니다', 'E'), ('.', 'S') |
| 비고 | 트위터와 동일하게 ‘에도’는 크게 J(관계언)으로 분류되었습니다. |

|  |  |
| --- | --- |
| 조사(2) | |
| 크림 제형으로 보습효과가 있어 거친 피부에 효과적 | |
| 트위터 | ('크림', 'Noun'), ('제', 'Modifier'), ('형', 'Noun'), ('으로', 'Josa'),  ('보습', 'Noun'), ('효과', 'Noun'), ('가', 'Josa'), ('있어', 'Adjective'),  ('거친', 'Verb'), ('피부', 'Noun'), ('에', 'Josa'), ('효과', 'Noun'),  ('적', 'Suffix') |
| 비고 | ‘으로’, ‘가’, ‘에’ 등 조사의 의미로 사용된 단어들이 정확하게 Josa(조사)로 분류되었음을 알 수 있다. |
| 코모란 | ('크림', 'NNP'), ('제', 'XPN'), ('형', 'NNG'), ('으로', 'JKB'), ('보습', 'NNG'), ('효과', 'NNG'), ('가', 'JKS'), ('있', 'VV'), ('어', 'EC'), ('거치', 'VV'),  ('ㄴ', 'ETM'), ('피부', 'NNG'), ('에', 'JKB'), ('효과', 'NNG'), ('적', 'XSN') |
| 비고 | ‘으로’, ‘에’는 JKB(부사격 조사)로 분류되었으며, ‘가’는 JKS(주격 조사)로 분류되었다. 코모란이 트위터와 한나눔보다 조금 더 자세하게 분류함을 추측할 수 있다. |
| 한나눔 | ('크림', 'N'), ('제', 'X'), ('형', 'N'), ('으로', 'J'), ('보습효과', 'N'),  ('가', 'J'), ('있', 'P'), ('어', 'E'), ('거칠', 'P'), ('ㄴ', 'E'),  ('피부', 'N'), ('에', 'J'), ('효과적', 'N') |
| 비고 | ‘으로’, ‘가’, ‘에’ 모두 J(관계언)으로 잘 분류되어 있음을 알 수 있다. |

|  |  |
| --- | --- |
| 관형사(1) : 체언 앞에 놓여서, 그 체언의 내용을 자세히 꾸며 주는 품사 | |
| 다음과 같은 사람은 이약을 복용하기 전에 의사와 상의할 것 | |
| 트위터 | ('다음', 'Noun'), ('과', 'Josa'), ('같은', 'Adjective'), ('사람', 'Noun'),  ('은', 'Josa'), ('이', 'Determiner'), ('약', 'Noun'), ('을', 'Josa'),  ('복용', 'Noun'), ('하기', 'Verb'), ('전', 'Noun'), ('에', 'Josa'),  ('의사', 'Noun'), ('와', 'Josa'), ('상의', 'Noun'), ('할', 'Verb'),  ('것', 'Noun') |
| 비고 | ‘이’가 Determiner(관형사)로 사용되어 뒤의 ‘약’을 꾸며준다. |
| 코모란 | ('다음', 'NNG'), ('과', 'JC'), ('같', 'VA'), ('은', 'ETM'), ('사람', 'NNG'),  ('은', 'JX'), ('이', 'MM'), ('약', 'NNG'), ('을', 'JKO'), ('복용', 'NNG'),  ('하', 'XSV'), ('기', 'ETN'), ('전', 'NNG'), ('에', 'JKB'), ('의사', 'NNG'),  ('와', 'JC'), ('상의', 'NNG'), ('하', 'XSV'), ('ㄹ', 'ETM'), ('것', 'NNB') |
| 비고 | ‘이’가 MM(관형사)로 사용되어 트위터와 같이 뒤의 ‘약’을 꾸며준다. |
| 한나눔 | ('다음', 'N'), ('과', 'J'), ('같', 'P'), ('은', 'E'), ('사람', 'N'),  ('은', 'J'), ('이약', 'N'), ('을', 'J'), ('복용', 'N'), ('하', 'X'),  ('기', 'E'), ('전', 'N'), ('에', 'J'), ('의사', 'N'), ('와', 'J'),  ('상의', 'N'), ('하', 'X'), ('ㄹ', 'E'), ('것', 'N') |
| 비고 | ‘이’가 관형사로 사용되지 않고, 뒤의 약과 붙어서 새로운 N(체언)으로 분류되었다. |

|  |  |
| --- | --- |
| 관형사(2) | |
| 그 위에 적당한 거즈 또는 붕대로 고정한다. | |
| 트위터 | ('그', 'Noun'), ('위', 'Noun'), ('에', 'Josa'), ('적당한', 'Adjective'),  ('거즈', 'Noun'), ('또는', 'Adverb'), ('붕대', 'Noun'), ('로', 'Josa'),  ('고정', 'Noun'), ('한다', 'Verb'), ('.', 'Punctuation') |
| 비고 | ‘그’가 관형사로 사용되었지만 트위터에서는 ‘그’를 관형사가 아닌 Noun(명사)로 분류하였기 때문에 분류가 제대로 잘 되지 않았다고 추측 할 수 있다. |
| 코모란 | ('그', 'MM'), ('위', 'NNG'), ('에', 'JKB'), ('적당', 'XR'), ('하', 'XSA'),  ('ㄴ', 'ETM'), ('거즈', 'NNP'), ('또는', 'MAJ'), ('붕대', 'NNP'), ('로', 'JKB'),  ('고정', 'NNG'), ('하', 'XSV'), ('ㄴ다', 'EF'), ('.', 'SF') |
| 비고 | ‘그’가 의도에 맞게 MM(관형사)로 제대로 잘 분류가 되었다고 볼 수 있다. |
| 한나눔 | ('그', 'M'), ('위', 'N'), ('에', 'J'), ('적당', 'N'), ('하', 'X'), ('ㄴ', 'E'),  ('거즈', 'N'), ('또는', 'M'), ('붕대', 'N'), ('로', 'J'), ('고정', 'N'),  ('하', 'X'), ('ㄴ다', 'E'), ('.', 'S') |
| 비고 | ‘그’가 관형사와 부사를 일컫는 M(수식언)으로 잘 분류가 되었다고 볼 수 있다. |

|  |  |
| --- | --- |
| 미등록어 & 이모티콘(1) | |
| 쇽: 쇽, 아나필락시양 증상 | |
| 트위터 | ('쇽', 'Noun'), (':', 'Punctuation'), ('쇽', 'Noun'), (',', 'Punctuation'),  ('아나필락시양', 'Noun'), ('증상', 'Noun') |
| 비고 | ‘쇽’이 Unknown(미등록어)가 아닌 Noun(명사)로 분류되었다.  즉, 미등록어에 관해서는 분석이 제대로 이루어졌다고 말 할 수 없다. |
| 코모란 | ('쇽:', 'NA'), ('쇽,', 'NA'), ('아나', 'NNP'), ('필', 'NNP'), ('락', 'NNG'),  ('시', 'NNB'), ('양', 'NNP'), ('증상', 'NNP') |
| 비고 | ‘쇽’이라는 의미 없는 오타들을 정확하게 NA(분석불능범주)로 분류해준다. |
| 한나눔 | ('쇽:', 'N'), ('쇽,', 'N'), ('아나필락시양', 'N'), ('증상(과민성유사증상', 'N') |
| 비고 | 트위터와 유사하게 그냥 ‘쇽’은 하나의 N(체언)으로 분류되었으며, 뒤의 명사 또한 제대로 분류가 되지 않았다고 볼 수 있다. |

|  |  |
| --- | --- |
| 미등록어 & 이모티콘(2) | |
| 여드름에는 효과를 보기가 힘들다고 하네요..ㅠ | |
| 트위터 | ('여드름', 'Noun'), ('에는', 'Josa'), ('효과', 'Noun'), ('를', 'Josa'),  ('보기', 'Noun'), ('가', 'Josa'), ('힘들다고', 'Adjective'), ('하네요', 'Verb'),  ('..', 'Punctuation'), ('ㅠ', 'KoreanParticle') |
| 비고 | ‘ㅠ’를 눈물을 흘리는 하나의 이모티콘 이라고 생각하여 KoreanParticle(이모티콘)으로 잘 분류되었음을 알 수 있다. |
| 코모란 | ('여드름', 'NNP'), ('에', 'JKB'), ('는', 'JX'), ('효과', 'NNG'), ('를', 'JKO'),  ('보', 'VV'), ('기', 'ETN'), ('가', 'JKS'), ('힘들', 'VA'), ('다고', 'EC'),  ('하네요..ㅠ', 'NA') |
| 비고 | 코모란에서는 이모티콘을 따로 분류하는 기능이 없기 때문에 ‘하네요..ㅠ’자체를 NA(분석불능범주)로 생각하여 분류했다고 추측 할 수 있다. |
| 한나눔 | ('여드름', 'N'), ('에는', 'J'), ('효과', 'N'), ('를', 'J'), ('보', 'P'),  ('기', 'E'), ('가', 'J'), ('힘들', 'P'), ('다', 'E'), ('고', 'J'), ('하', 'P'),  ('네', 'E'), ('요', 'J'), ('..', 'S'), ('ㅠ', 'N') |
| 비고 | 한나눔에서는 미등록어 및 이모티콘 모두 분류하는 기능이 없기 때문에 (1)에서와 유사하게 ‘ㅠ’또한 하나의 N(체언)으로 분류해놓았다. |