

높임말 반말 변환기 코드 수정

변환기 코드 logic 변경

- 형태소 분석이 아닌 spacing단위로 분석하도록 바꾸었다. 기존의 변환 모듈에서는 형태소의 tag들에 집중하여, 어떤 tag가 오는지에 따라 변경하였다면, 변경된 방식에서는 형태소 자체에 집중하여 변경하도록 하였다. 즉 기존에는 tag마다 사전이 있었고 종결어미를 변경하였는데 이때 단순한 종결어미, 결합형 종결어미들을 구분하여 사전을 만들고 거기서 변경하는 방식으로 갔다. 예를 들어 '-입니다'는 서술격조사 '이'와 종결어미 '입니다'가 결합된 '이+입니다'의 형태로, VCP+EF라는 형태로 나오며 VCP+EF 사전에서 '입니다'를 반말로 변경해주는 방식이었다. 그러나 이러한 방식에서는 형태의 변환이 다양했으며, 다양한 결합이 가능했기에 사전이 길어졌으며, 또한 예외적인 경우들이 많이 발생했기에 이를 모두 처리하는 것이 번거로웠다.
- 현 방식에서는 형태소 자체에 집중하는 방식을 추가하였다. 결합형으로 나오는 경우, 예를 들어, VCP+EF라면 종결어미만 따로 분리해서 종결어미가 무엇인지에 따라 바꾸는 것이다. '입니다'라면 '입니다'로 분리하여 이를 '다'로 변경하는 방식이다. 사전의 길이가 짧아지고, 어체에 따라 종결어미를 다르게 변경하기 용이해졌으며, 코드도 간단해졌기에 유지 보수가 쉬워졌다. 특정 형태소만을 고려하기에 더 간단해졌다.
- 문제는 높임말 → 반말에서 생긴다. 높임말의 경우, "-세요, -나요, -습니다"에서 발생한다. 이때 '-습니다'나 '-나요' 이는 앞의 자음이 'ㅌ'나 'ㄷ'가 있다면 '아' 또는 '어'로 변경하는 것으로 일시적으로 해결 된다. 그러나 '세요'의 경우 동사와 결합되어 사용될 때 동사의 원형을 알아야 변환이 가능하다. 예시는 다음과 같다.
 - 오르세요 → 올라
 - 미세요 → 밀어
 - 주세요 → 쥐
 - 지세요 → 저
 - 기세요 → 기
- 존칭 선어말 어미 '-시'와 결합되는 '세요', '십니까' 등이 문제가 된다. 이는 용언의 불규칙 활용이 있을 때 문제가 된다. 즉, 동사의 원형을 모르기 때문이다.
- 이를 해결하고자 용언의 규칙 또는 불규칙 활용을 이해할 필요가 있다. 또한 이 개념을 코드에 적용하여 다르게 처리해주어야 한다.

해결해야 하는 문제점 및 해결 방안

- 존칭 선어말 어미 '시'가 들어갈 경우, 용언의 활용이 적용되는 경우가 있는데 그 경우가 '-세요'와 '나요', '십니까' 등이 있다. 이 경우, 용언의 활용이 되기 전의 원형을 알아야 하는 경우가 대부분이다. 문제가 되는 것은 불규칙 활용인데, 이 경우, 규칙에 따른 활용이 아닌, 불규칙적으로, 어미 또는 어간에 따라 종류가 다양하다. 이는 암기를 해야 하는 부분이다.


불규칙 활용과 규칙 활용을 반영해서 동사 처리 flow

- 앞의 tag가 VV인지 VA인지 확인
 - 으세요, 습니다, 나요는 종결어미를 빼고 ‘ㅏ’또는 ‘ㅑ’로 변경해준다.
 - ‘-세요’, ‘-습니까’, ‘-습니다’ 등 존칭 선어말 어미가 있는 문장의 경우는 예외로 처리를 해주어야 한다.
 - if 세요 앞의 동사가 ‘ㄹ’ 규칙 활용이 되었는지를 확인하기 위해서 dictionary를 확인한다. → 적절한 원형으로 바꿔준다. 즉, ㄹ을 붙인다.
 - 쓰다는 어찌해결?
 - elif 불규칙활용 적용한다.
 - if ‘ㅏ’ 불규칙 활용 → ‘으세요’가 붙어서 고려할 필요 없음
 - elif ‘ㄷ’ 불규칙 활용 → ‘으세요’가 붙어서 고려할 필요 없음
 - elif ‘ㅂ’ 불규칙 활용 → ‘으세요’가 붙어서 고려할 필요 없음
 - elif ‘ㄹ’ 불규칙 활용 → 앞 동사의 종성에 ‘ㄹ’이 있으면 ‘ㄹ’만 추가로 붙이고, 없으면 ‘ㄹㄹ’
 - elif ‘우’ 불규칙 활용 → ‘푸다’ 외에는 ‘ㅏ’가 그대로 결합됨 ⇒ ‘ㅏ’를 결합하되, 푸의 경우는 ‘ㅏ’탈락
 - elif ‘오’ 불규칙 활용 → ‘보’와 ‘오’는 ‘ㅏ’를 붙인다.
 - elif ‘하’ 불규칙 활용 → 하다는 해로 임시적으로 변경
 - else 활용이 안된 동사 처리
 - 어간 끝 모음에 ‘ㅏ’, ‘ㅑ’가 어미 첫 모음 ‘ㅏ’, ‘ㅑ’ 앞에서 반모음으로 변하여서 ‘ㅏ’, ‘ㅑ’로 축약
 - 어간 끝 모음에 ‘ㅣ’가 어미 첫 모음 ‘ㅏ’, ‘ㅑ’ 앞에서 반모음으로 변하여서 ‘ㅑ’로 축약
- ⇒ ex) 예쁘세요 → 예뻐, 넓히세요 → 넓혀, 기세요 → 기어(거), 넘기세요 → 넘겨

▼ 국립국어원 참고

국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=171456&pageIndex=1



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=17750



국립국어원

이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

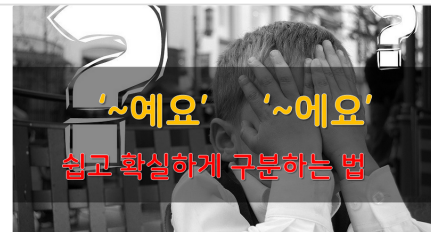
 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=128665



'~예요'와 '~에요' 쉽고 확실하게 구분하는 법

글을 쓰거나 문자를 보낼 때.. '~예요' (?) '~이에요' (?) 어떻게 써야 맞게 쓰는 것일까요? 말로 할 때는 굳이 명확하게 구분해서 발음할 필요는 없지만 글로 쓸 때는 문법에 따라 구분해서 써야 한다는 것 알고 계신가요? 알고 있

☞ <https://th0413.tistory.com/112>




국립국어원

이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에 대한 답변은 휴일을 제외하고 다음 날까지 완료되며, 상황에 따라 조금 늦어질 수도 있습니다. 3.

 https://korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=177744&pageIndex=1

국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=123599



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=169153&pageIndex=1



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=&qna_seq=19069&pageIndex=12540



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=40104



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=102198



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=220731



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=37966



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=174959&pageIndex=1



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=9229



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=149848



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=88935



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=183002&pageIndex=1



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=129158



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=114370



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=145600



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

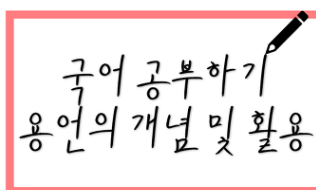
 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=&qna_seq=148487&pageIndex=7



[국어] 14. 용언의 활용 정리하기!! - 용언 개념 정리, 규칙 활용, 불규칙 활용

국어] 14. 용언의 활용 정리하기!! - 용언 개념 정리, 규칙 활용, 불규칙 활용 안녕하세요~ 즐거운 주말이에요!!! 지난 국어 공부 포스팅에서 본용언과 보조용언의 띄어쓰기를 중심으로 포스팅을 했었어요. 그때..

 <https://unomasuno.tistory.com/164>



https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=98886

국립국어원

이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=134602



<https://lovit.github.io/nlp/2018/06/07/lemmatizer/>

국립국어원

이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

https://korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=114777



<https://namu.wiki/w/%ED%95%9C%EA%B5%AD%EC%96%B4/%EB%B6%88%EA%B7%9C%EC%B9%99%20%ED%99%9C%EC%9A%A9>

국립국어원

이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

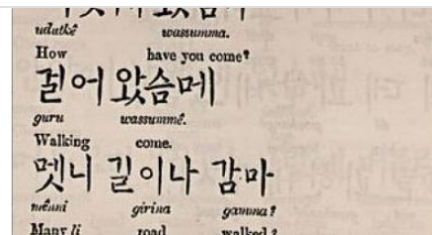
https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=123396



[NLP] 자연어처리 - 한국어 전처리를 위한 기법들

한국어는 너무 어렵다. 띄어쓰기 차이, 한 글자 차이 등으로 의미가 달라진다. 많은 자연어 처리 책이나 논문등에서 영어에 대한 전처리 기법은 많지만 한국어에 대한 전처리 및 임베딩이 약했습니다. 하지만 최..

<https://ebbnflow.tistory.com/246>



국립국어원


이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=216&qna_seq=65081




국립국어원

이곳은 어문 규범, 어법, 표준국어대사전 내용 등에 대하여 문의하는 곳입니다. 1. 법률 및 규정의 해석, 시험 문제의 정답 판정 등 소관 기관의 해석이 필요한 사안은 답변해 드리기 어려우니 양해해 주시기 바랍니다. 2. 질문에

 https://www.korean.go.kr/front/onlineQna/onlineQnaView.do?mn_id=&qna_seq=127180&pageIndex=1



 <https://s-space.snu.ac.kr/bitstream/10371/127736/1/000000026568.pdf>