

발제날짜:11월 6일 수요일

iLED 2학기 세미나 발제 배유진

텍스트 마이닝을
활용한 논증적
글쓰기 자동 채점
자질 추출 타당화
연구

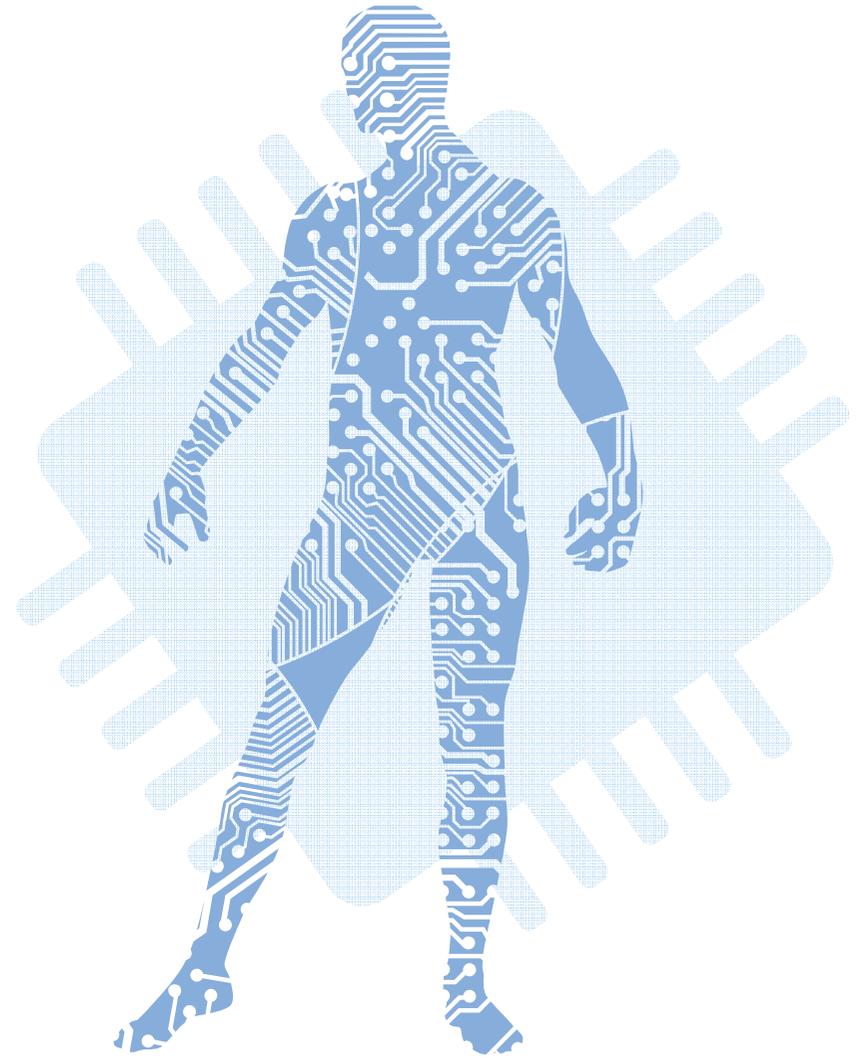
한국교육대학교 석사학위논문

국어교육전공 김승주

2019년 2월

목차

- 01** 서론
- 02** 이론적 배경
- 03** 연구 방법
- 04** 연구 결과
- 05** 결론



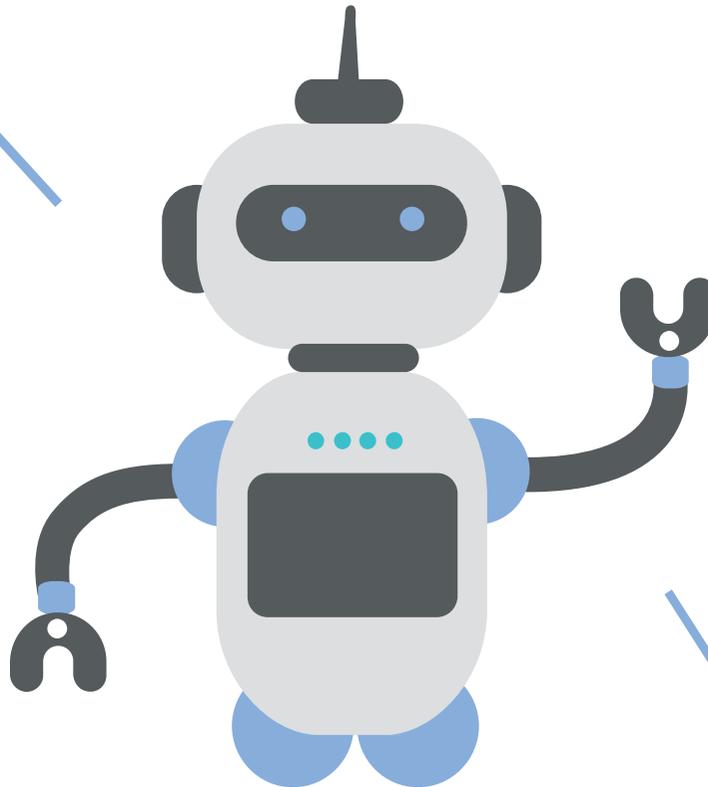


01. 서론

1. 연구의 필요성 및 목적

자유학기제
확대 적용, 실시

근본적인 평가
환경 개선



4차 산업혁명
지식 기반사회
도래

수학능력시험 논, 서술형
도입 관련
사회적 논의 진행

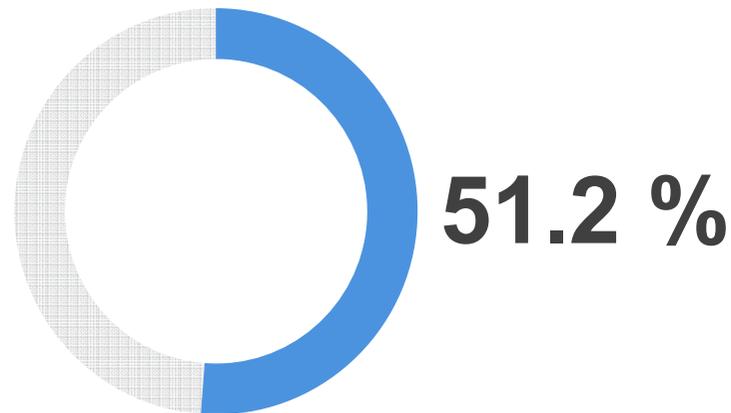


1. 연구의 필요성 및 목적

중등학교 교실에서 체계적인 글쓰기 활동 이루어지지 않음.

가은아(2016)

중등 교사 **51.2%** ‘학교에서 쓰기 수행을 잘 하지 않는 것’





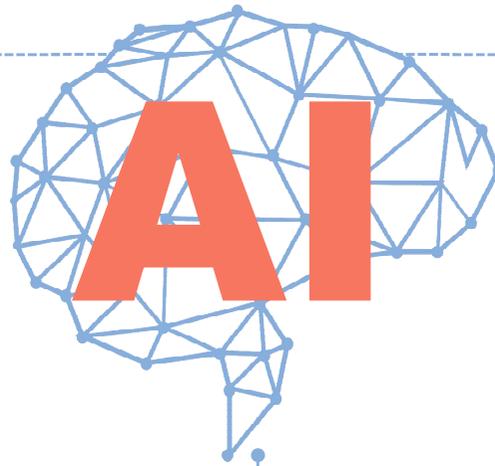
1. 연구의 필요성 및 목적

중등학교 교실에서 체계적인 글쓰기 활동 이루어지지 않음.

문제 원인

1. 쓰기 평가자의 인지적 부담의 문제
2. 현재 수능이 교육 현장에 갖는 강력한 환류 효과

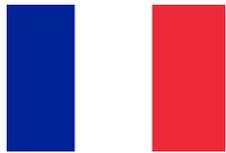
자동 쓰기 채점 시스템



GCE A-level
(General Certificated
of Education
Advanced-level)



2020년부터 기존
객관식 중심 수업
폐지, 서술형 평가
도입



바칼로레아



SAT



아비투어

2. 선행 연구 검토

Ong et al.(2014) 대학생들이 작성한 논증적 에세이 52개를 대상으로 자동 채점을 시도함.

$$\text{학생 글 점수} = N_{\text{최근연구}} + N_{\text{가설}} + N_{\text{반박}} + \frac{N_{\text{주장}} + N_{\text{지지}}}{N_{\text{단락}}} + \frac{N_{\text{인용}}}{N_{\text{단락}}}$$

$N_x = (x\text{의 개수})$



2. 선행 연구 검토

이슬기(2017) 고등학교 3학년이 작성한 논설문 40편을 대상으로 텍스트 응집성 점수에 대한 자동 채점 가능성 탐색함.

Song et al(2014) 대학원생 300명이 제시된 텍스트에 대해 비판적인 의견을 작성하도록 한 논증 분석 글쓰기 과제를 대상으로 자동 채점을 실시함.



3. 연구 문제 및 제한점

구체적 연구 문제

1. 논증적 글쓰기 평가 요소는 재구조화될 수 있는가?
2. 형성적 피드백 제공이 가능한 평가 자질이 추출될 수 있는가?
3. 이들 자질은 실제 점수를 설명할 수 있는가?



3. 연구 문제 및 제한점

연구의 제한점

1. 논증 요소들이 휴리스틱 알고리즘을 기반으로 탐지되었다.
2. 자질과 점수와의 선형적 관계를 전제로 하였다.
3. 확률적 언어 모형을 기반으로 한 자질 등을 모델에 포함시키지 않았다.



02. 이론적 배경



1. 논증과 논증적 글쓰기 평가

가. 논증적 글쓰기의 평가 요인과 평가 기준

표1. 선행 연구의 논술 능력 개념 층위

선행 연구	논술 능력
김태호(2016)	이해력 및 분석력/ 표현력/ 논증력
	<p>1. 논리적 내용 구성 능력</p> <p>2. 표현력</p>
원진숙(1998)	문법적 언어 능력/ 담화 구성 능력/ 논리적 사고력/ 문제 해결 능력/ 사회적 인지 능력

1. 논증과 논증적 글쓰기 평가

가. 논증적 글쓰기의 평가 요인과 평가 기준

	내용	조직	표현
원진숙 (1995)	<ul style="list-style-type: none"> • 내용의 통일성 • 주제의 선명성 • 논증의 타당성 • 결론의 적절성 • 사고력 	<ul style="list-style-type: none"> • 글의 구성 영역 • 문단 전개 방식 • 글의 논리 구조 	<ul style="list-style-type: none"> • 맞춤법 및 띄어쓰기 • 어휘의 정확성 • 문장의 정확성 • 글의 응집성 • 문체의 적절성
김태호 (2016)	<ul style="list-style-type: none"> • 논점 파악 • 제시문 이해 • 제시문 분석 • 논증의 합리성 • 논증의 일관성 • 논증의 적합성 • 자기 의견 생성 • 논증의 타당함 	<ul style="list-style-type: none"> • 구성의 적절성 	<ul style="list-style-type: none"> • 객관적 표현 • 정확한 표현 • 설득력

표2. 선행 연구의 논증적 글쓰기 세부 평가 기준

A 1. 논증과 논증적 글쓰기 평가

나. 논증적 글쓰기의 논증 평가

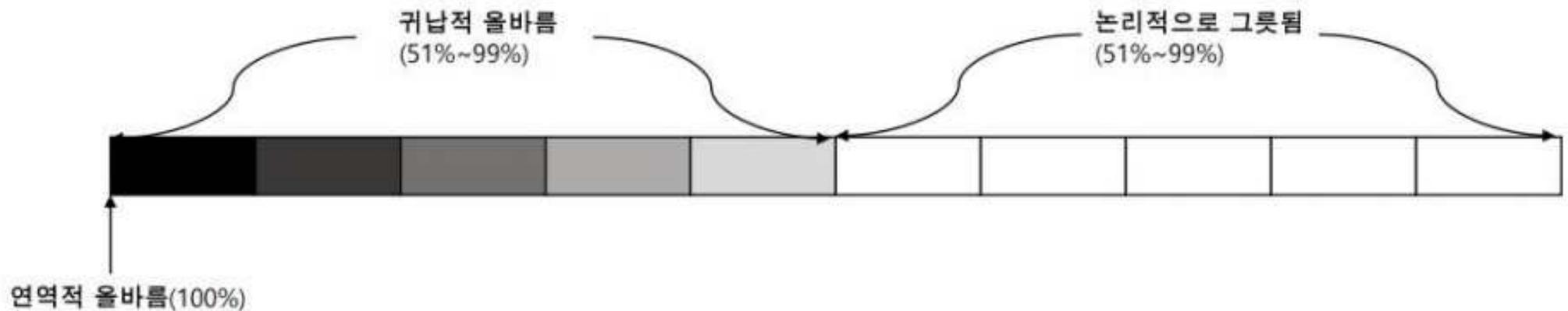


그림1. 논증의 올바름에 따른 분류(Skyrms & Brain, 1986; 박준호, 2006재인용)

A 1. 논증과 논증적 글쓰기 평가

나. 논증적 글쓰기의 논증 평가

표3. McCann(1989)의 논증 평가 기준

평가 요소	점수	평가 기준
주장(Claim)	6	진술된 입장이 명확하고 일반화가 완료됨
	4	필자가 준 정보로부터 의도를 추론해야만 하지만 주어진 정보가 충분하여 주제나 입장과 관련된 일반화가 가능함
	2	비록 주제에 대한 일반화는 이루어졌지만 필자의 주장이 명확하지 않고 전문성(특수성)이 부족
	0	주제에 관련된 주장 없음



2. 텍스트 마이닝을 활용한 논증적 글쓰기 자동 채점

가. 텍스트 마이닝과 아규먼트 마이닝 1) 논증 분석 체계 설정

표4. 쓰기 자동 채점 영역의 아규먼트 마이닝 관련 선행 연구의 논증 분석 체계

관련 선행 연구	논증 분석 체계
Ong et al.(2014)	(1) 인용(Citation) (2) 주장(Claim) (3) 가설(Hypothesis) (4) 현재연구(Current Study) (5) 반박(Oppose) (6) 지지(Supports)



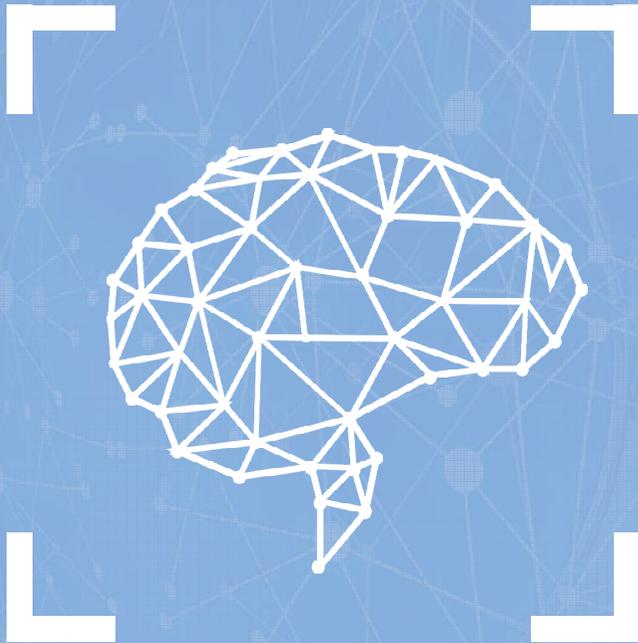
2. 텍스트 마이닝을 활용한 논증적 글쓰기 자동 채점

가. 텍스트 마이닝과 아규먼트 마이닝 2) 논증 분석 알고리즘 설정

표5. Ong et al.(2014)의 논증 요소 및 탐지 규칙

논증 요소	탐지 규칙
주장(Claim)	문장에 {suggest, evidence, shows, Essential, indicating} (대소 문자 구별)로 시작하는 문자열 포함
지지(Supports)	문장이 우연적 연결사(Contingency connective)로 시작하고 4자리 숫자 불포함(so, because...)
인용(Citation)	문장에 4자리 숫자가 포함
반박(Oppose)	비교 담화 연결사(Comparison discourse connective)으로 시작(but, still...) {conflict, opposose}로 시작하는 문자열 및 4 자리 숫자 (인용문의 연도로 의도된) 문자열 포함

Artificial Intelligence



머신러닝 (Machine Learning)

머신러닝은 데이터 패턴을 찾아내는 문제를 학습의 관점에서 접근한다.

먼저, 모델은 트레이닝 셋(training set)을 바탕으로 데이터의 패턴을 학습하고

이후 학습의 결과가 일반화(generalization) 될 수 있는지를 확인하기 위해 테스트 셋(test set)에서 성능 검사를 거친다.

Artificial Intelligence



머신러닝 (Machine Learning)

- 지도학습(Supervised Learning)은 데이터에 대한 정답 라벨(label)을 명시적으로 제시하여 학습이 이루어지도록 하는 방법
- 비지도학습(Unsupervised Learning)은 정답 라벨이 주어지지 않은 상태에서 학습이 이루어지도록 하는 방법
- 강화학습(Reinforcement Learning)은 주어진 환경 속에서 시행착오를 거치는 과정을 통해 최적의 결과를 찾도록 하여 학습이 이루어지게끔 하는 방법

2. 텍스트 마이닝을 활용한 논증적 글쓰기 자동 채점

가. 텍스트 마이닝과 아규먼트 마이닝 2) 논증 분석 알고리즘 설정

표6. Song et al.(2014)의 논증 요소 탐지를 위한 자질

번호	자질명	설명
1	word n-gram	n(1~3)개 단어로 이루어진 연속적인 문자열에 대한 이진 변수로 특정한 n-gram이 존재하면 1 그렇지 않으면 0 부여
2	이전 문장과 다음 문장의 word n-gram	word n-gram과 유사한 자질
3	문장 길이(이진 변수)	2t(1~10)보다 긴 단어 토큰이 있는지의 여부를 나타내는 이진 변수
4	문장 위치	문장 번호를 텍스트 내 문장 수로 나눈 값
5	품사 태그	다양한 품사의 존재 여부에 대한 이진 변수
6	제시문과의 일치도	다음의 특질은 에세이와 제시문 사이의 어휘의 중복 정도를 기초로 자카드 계수로 산출



2. 텍스트 마이닝을 활용한 논증적 글쓰기 자동 채점

나. 논증적 글쓰기 자동 채점과 자질 추출

지도학습 기반의 쓰기 자동 채점을 위한 연구 절차

1단계. 트레이닝 셋으로 사용될 코퍼스 데이터를 구축하는 것

2단계는 평가 자질을 추출하는 것

3단계는 모델 생성 및 검증 과정

2. 텍스트 마이닝을 활용한 논증적 글쓰기 자동 채점

나. 논증적 글쓰기 자동 채점과 자질 추출

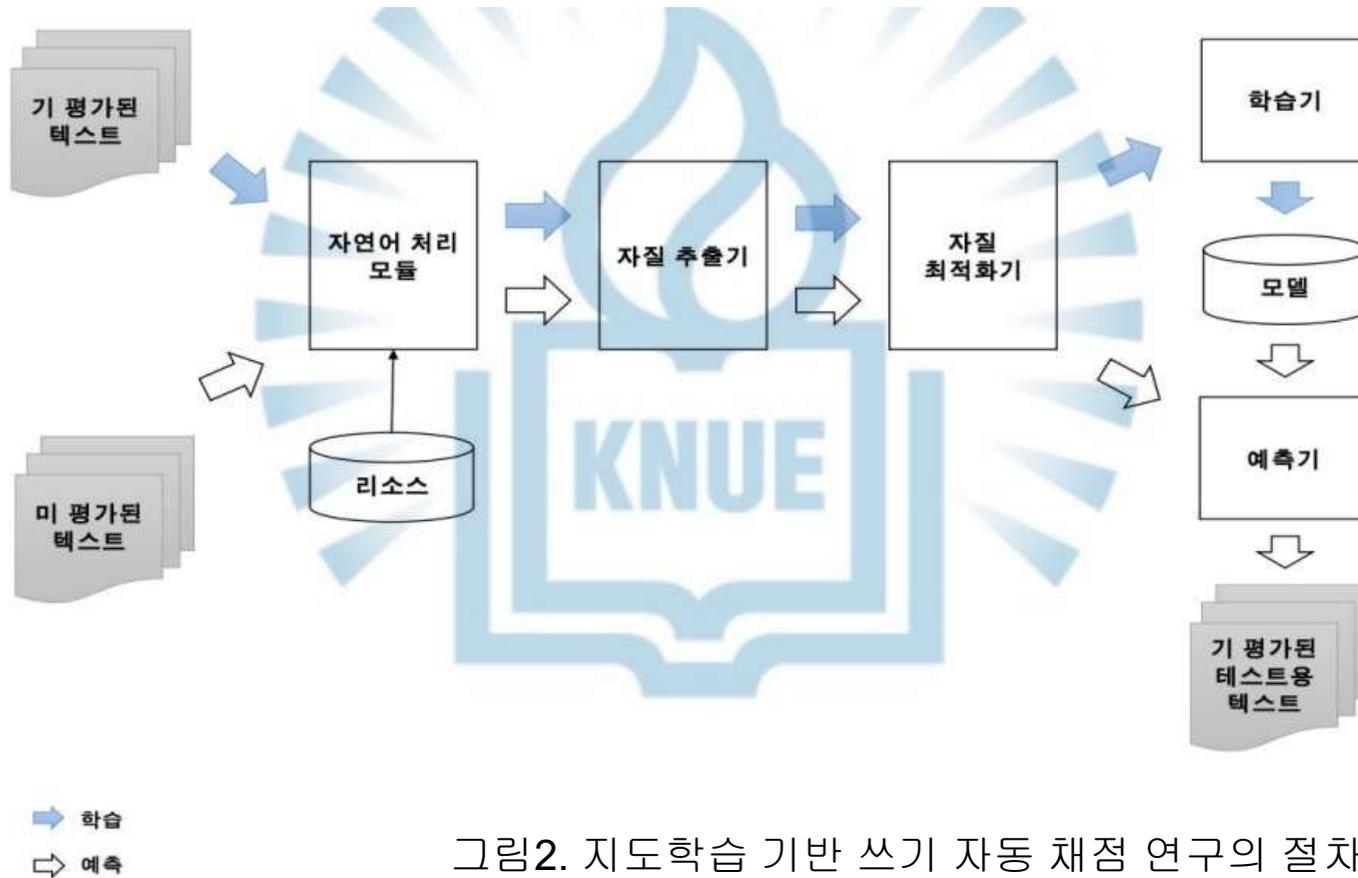


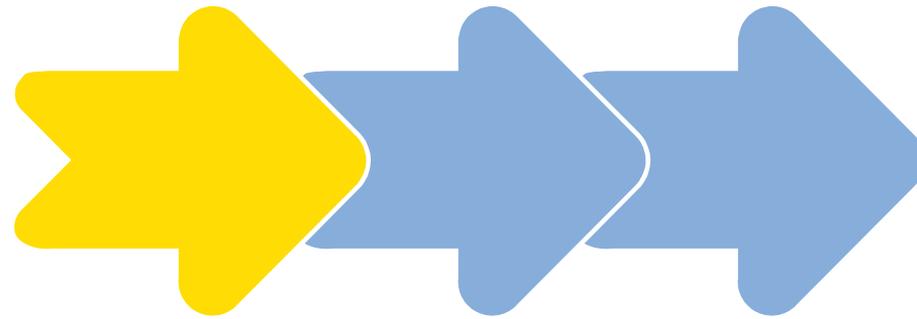
그림2. 지도학습 기반 쓰기 자동 채점 연구의 절차(Yi et al., 2015)



03. 연구 방법



1. 연구 절차



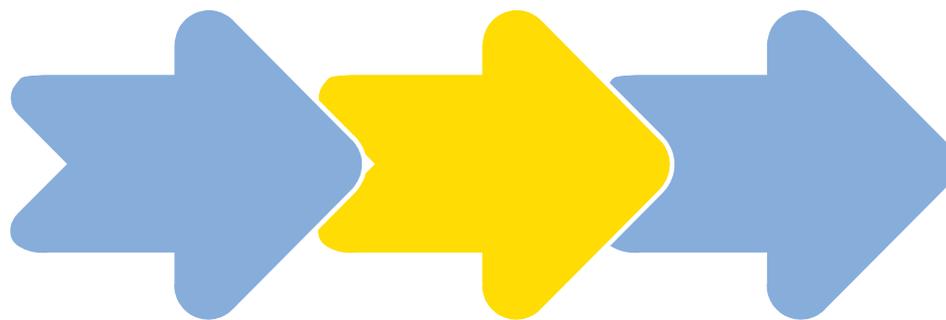
01 코퍼스 구축
쓰기 과제 및 평가 기준 구안 → 과제 투입 → 평가 실시 → 주석 코퍼스 구축

02 텍스트 전처리

03 자질 추출 및 타당화



1. 연구 절차



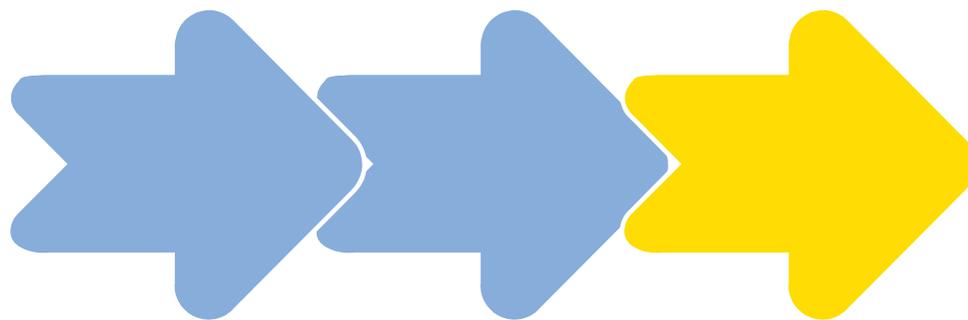
02 텍스트 전처리
풍사 태깅 → 아규먼트 마이닝

01 코퍼스 구축

03 자질 추출 및 타당화



1. 연구 절차



03 자질 추출 및 타당화
자질 설계 → 단계적 회귀분석 → 파악 요소 재구조화

01 코퍼스 구축

02 텍스트 전처리



2. 연구 대상

가. 쓰기 과제

‘학생부 종합전형은 폐지되어야 한다.’는 논제에 찬성 또는 반대 중 하나의 입장을 정해 글을 작성해야 함.

두 편의 자료 글과 논설문의 제목과 해당 신문사, 일시 등의 출처 표시와 함께 각 논설의 논지를 뒷받침하는 통계 자료가 제시됨.

자료를 자유롭게 활용하되, 내용을 그대로 베끼지 말 것과 1000자 내외로 글을 작성할 것이 조건으로 제시됨.

A

2. 연구 대상

가. 쓰기 과제

표7. 코퍼스 집단 내 분포

	남	여	계
지역 A	25	26	51
지역 B	21	25	46
계	46	51	97

2. 연구 대상

나. 평가 기준 구안 및 평가

표8. 논증적 글쓰기 평가 기준

평가 범주	평가 요소	평가 기준
내용	논증의 명료성	하나의 입장을 선택하여 자신의 주장을 명료하게 제시하였는가?
	논증의 적절성	주장과 근거의 연결관계가 적절한가?
	논증의 타당성	주장을 뒷받침하는 근거가 타당하고 충분한가?
조직	조직의 통일성	하나의 입장에 따라 통일성있게 글이 조직되었는가?
	구조의 유기성	문단과 문장내용의 연결이 유기적이고 자연스러운가?
	조직의 효과성	예상독자를 설득하기 위해 효과적으로 글을 조직하였는가?
표현	어법의 정확성	맞춤법, 띄어쓰기, 문장호응등이 정확한가?
	어조의 적합성	설득력있고 논리적인 어조로 표현하였는가?
	단어사용의 명료성	구체적이고 명확한 단어를 사용하여 글의 내용을 전달하였는가?

A

3. 텍스트 전처리

가. 논증 분석 체계

표9. 논증 분석 체계

번호	논증 요소	설명
(1)	핵심주장	논제에 대한 필자의 입장을 담고 있는 요소
(2)	반론 제시 및 반박 주장	예상 가능한 반론을 제시하고 이에 대한 반박 주장을 담고 있는 요소
(3)	주장	핵심주장에 대한 하위주장을 나타내는 요소
(4)	근거 및 보충설명	핵심주장, 반론 제시 및 반박 주장, 주장에 대한 근거를 제시하거나 문제의 배경(background)을 나타내는 등의 논증의 보충 요소



3. 텍스트 전처리

가. 논증 분석 체계

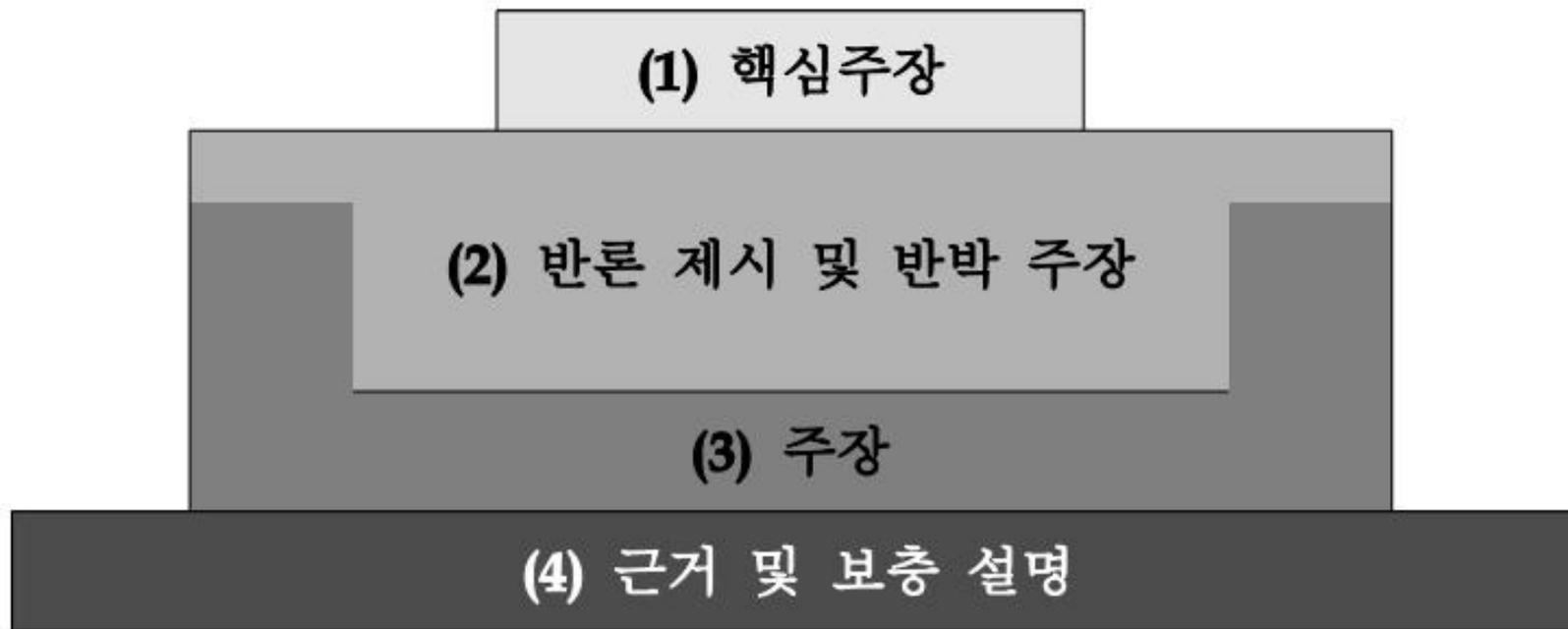


그림3. 논증 모형



3. 텍스트 전처리

나. 논증 분석 알고리즘

표10. 전제 지시어와 결론 지시어(김희정 & 박은진, 2004:70)

전제 지시어	결론 지시어
왜냐하면	따라서
...라는 사실을 고려한다면	그래서
...이니	그러므로
...이므로	그러니까
...에서 알 수 있듯이	...이 도출된다
...라는 사실 때문에 우리는	...라고 결론 내릴 수 있다.
...인 까닭에	그것은 ...를 함축한다.
...인 이유로	...임이 틀림없다.
그 까닭은	그 결과 ...이다.
그 이유는	...을 의미한다.



3. 텍스트 전처리

나. 논증 분석 알고리즘

표11. 핵심주장 탐지 규칙

논증 요소	번호	규칙 내용	탐지 예시
핵심 주장	규칙 1	문장 내 문자열에 순서대로 아래 A, B, C 집합 당 하나의 원소를 포함하고 A= {‘학생부종합전형’, ‘이 전형’, ‘이 제도’} B= {‘유지’, ‘폐지’, ‘폐기’, ‘없애’} C= {‘되’, ‘하’, ‘야’, ‘찬성’, ‘반대’, ‘좋’, ‘옳’, ‘올바르’} 이어서 D 집합의 원소를 포함하지 않는 문장 D= {‘면’, ‘ㄴ다면’, ‘ㄹ경우’, ‘하는+N’, ‘한다는+N’}	○ 나는 (A)학생부종합전형 (B)폐지에 대해 (C)반대한다. ○ (A)학생부종합전형으로만 학생들을 뽑는 것이 아니기 때문에 굳이 학생부종합전형을 (B)없애(C)야 한다는 생각은 하지 않는다. ○ 특목고, 자사고를 못 간 게 죄가 아닌데 이런 글을 볼 때면 죄가 되는 것 같기도 해서 (A)이 전형은 (B)폐지(C)되어야 한다.
	규칙 2	문장 내 문자열에 순서대로 A, B 의 문자열 그리고 ‘면’ 또는 ‘ㄴ다면’ 그리고 ‘안’을 포함하는 문장	○ 학생부종합전형은 폐지되면 안 됩니다.



04. 연구 결과



1. 논증적 글쓰기 평가 요소 재구조화

가. 평가 요소 간 상관관계 분석

표12. 평가 영역 간 상관관계 분석

	내용	조직
조직	.880*	
표현	.740*	.766*

*. 상관관계가 0.01 수준에서 유의합니다(양측).

1. 논증적 글쓰기 평가 요소 재구조화

가. 평가 요소 간 상관관계 분석

표13. 논증 관련평가 요소

평가 범주	평가 기준
내용	1) 하나의 입장을 선택하여 자신의 주장을 명료하게 제시하였는가?
	2) 주장과 근거의 연결관계가 적절한가?
	3) 주장을 뒷받침하는 근거가 타당하고 충분한가?
조직	4) 하나의 입장에 따라 통일성있게 글이 조직되었는가?
	5) 문단과 문장내용의 연결이 유기적이고 자연스러운가?
	6) 예상독자를 설득하기 위해 효과적으로 글을 조직하였는가?
표현	7) 맞춤법, 띄어쓰기, 문장호응등이 정확한가?
	8) 설득력있고 논리적인 어조로 표현하였는가?
	9) 구체적이고 명확한 단어를 사용하여 글의 내용을 전달하였는가?

A

1. 논증적 글쓰기 평가 요소 재구조화

나. 평가 요소 및 총점 간 상관관계 분석

논증적 글쓰기 평가에 있어 논증 텍스트의 ‘조직’과 ‘내용’이 텍스트의 전반의 수준을 결정하는 데에 매우 중요한 역할을 하고 있다는 것을 나타냄.

특별히 논증의 적절성에 대한 평가 점수가 총점에 미치는 영향이 컸으나, 다른 영역들 역시 이와 거의 유사한 수준에서 총점과의 상관관계를 보였음.

A 1. 논증적 글쓰기 평가 요소 재구조화

다. 전문가 집단 자문 및 평가 요소 재구조화

표14. 논증적 글쓰기 평가 요소 재구조화 결과

번호	평가 요소	평가 기준
(1)	논증의 명료성	하나의 입장을 선택하여 자신의 주장을 명료하게 제시하였는가?
(2)	논증의 타당성	주장을 뒷받침하는 근거가 충분한가?

2. 평가 자질 추출

가. 논증의 명료성 관련 자질

표15. 논증의 명료성 관련 자질

번호	자질명	설명
1	핵심주장 제시 여부	글의 처음 또는 마지막 25%(글자 수 기준)내에 핵심주장으로 태깅된 문장이 서론 및 결론에서 모두 나타나지 않을 경우를 1로 한 번 이상 나타날 경우를 0으로 코딩되도록 한 더미 변수
2	핵심주장 재진술 여부	글의 처음 또는 마지막 25%(글자 수 기준) 내에 핵심주장으로 태깅된 문장이 서론 및 결론에서 모두 나타날 경우 1로 그렇지 않은 경우가 0으로 코딩되도록 한 더미 변수
3	논증 주요 조직 점수	핵심주장 또는 반론 제시 및 반박 주장 또는 주장으로 태깅된 문장 수

2. 평가 자질 추출

나. 논증의 타당성 관련 자질

표16. 논증의 타당성 관련 자질

번호	자질명	설명
1	하위 주장 점수	주장 또는 반박 주장으로 태깅된 문장 수
2	근거 및 보충설명 과소 여부	전체 문장 수의 20%이하로 이유 및 근거를 사용한 경우 1, 그렇지 않은 경우 0으로 코딩되도록 한 더미 변수
3	1인칭 주장 표현 점수	주장 또는 반론 고려 및 반박 주장으로 탐지된 문장 중 ‘나’ 또는 ‘내’를 포함하는 문장의 수



3. 평가 자질 타당화

가. 논증의 명료성

논증의 명료성은 하나의 입장을 선택하여 자신의 주장을 명료하게 제시하였는지를 기준으로 산출된 점수임.

나. 논증의 타당성

논증의 타당성은 주장을 뒷받침하는 근거가 타당하고 충분한지를 기준으로 산출된 점수임.



05. 결론

시사점



설계한 평가 자질들은 논증과 관련된 국어교육학적 접근을 바탕으로 학생에게 형성적 피드백을 제공할 수 있다.



쓰기 자동 채점을 위해서는 자동 채점을 위한 특수한 쓰기 과제가 설계될 필요가 있다.



평가 기준의 설계에서부터 교육 측정 및 평가 이론들을 도입하여 쓰기 과제 별로 요구되는 학생의 쓰기 능력과 이에 따른 평가 기준을 타당화하는 작업이 선행되어야 할 필요가 있다.



평가 자질의 설계에 있어서도 보다 심도 깊은 논의가 필요하다.



Thank You