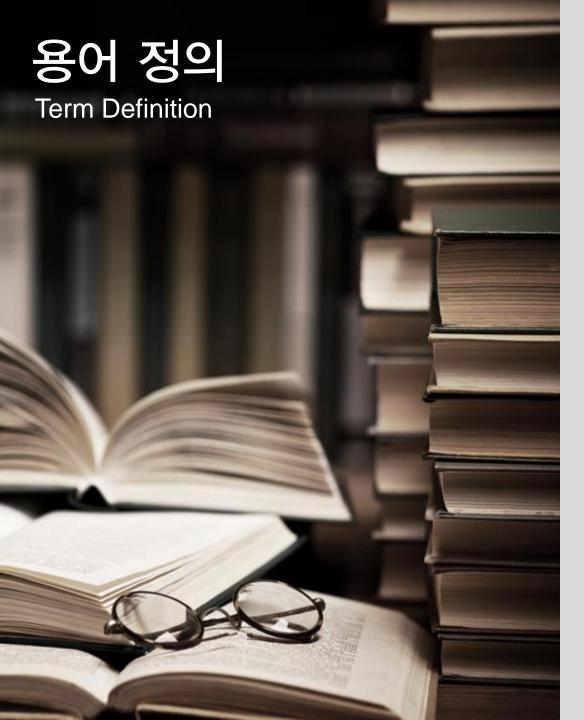


Content

- 1. 연구 배경
- 2. 선행 연구 조사
- 3. 연구목적
- 4. 분석방법
- 5. 분석결과
- 6. 결론 및 시사점



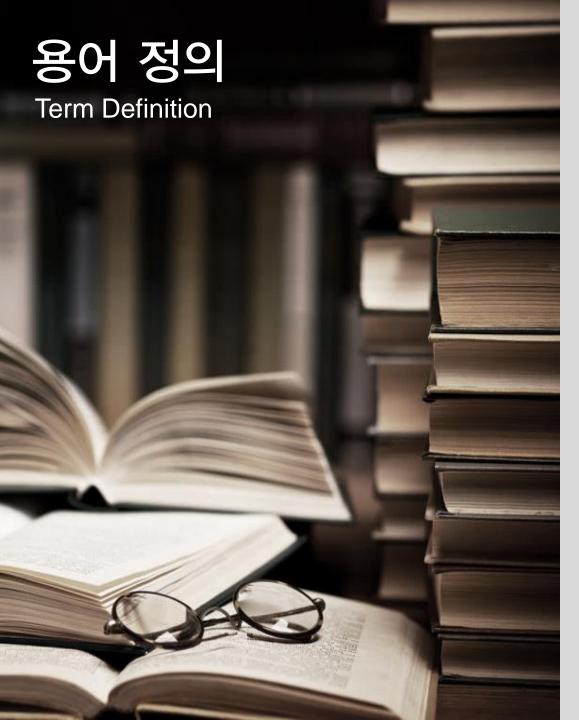


중앙은행

화폐를 발행하고 화폐의 총량을 관리하는 권한을 가진 기관

금융통화위원회

중앙은행(한국은행)의 정책결정기구 대한민국의 통화정책과 한국은행 운영에 관한 사항을 심의, 의결

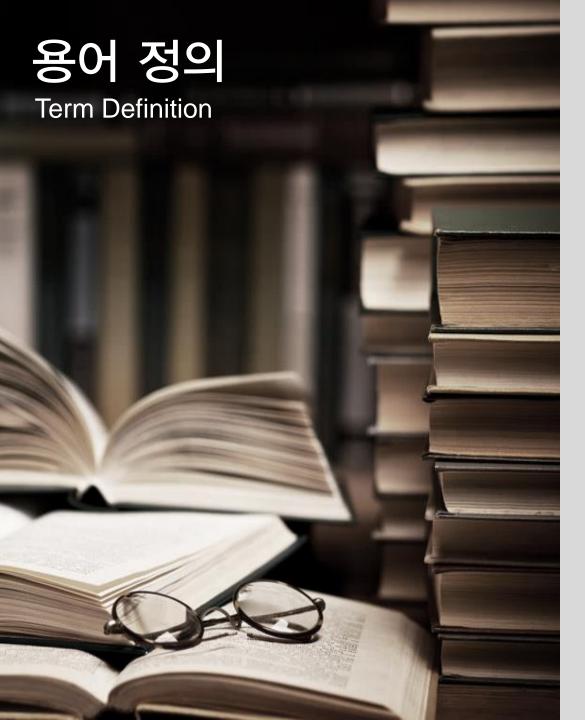


콜금리

일반 은행들 끼리 돈을 거래할 때의 금리 초기에는 기준금리를 대체

기준금리

나라의 기준이 되는 금리 일반 은행과 한국은행이 돈을 거래할 때의 금리



통화 정책

정부/중앙은행 등이 한 국가의 화폐 공급이나 금리 등을 경제 성장이나 안정성을 유지하기 위해 시행하는 일련의 조치

극성(polarity)

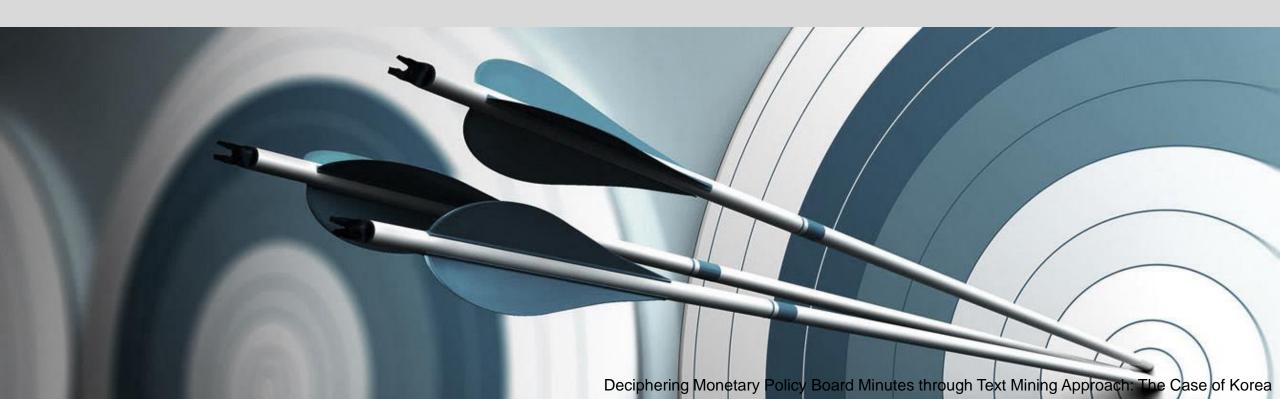
텍스트 안에 담겨있는 감정 기본적으로 긍정, 부정, 중립으로 나누어 진다. Beyond polarity의 경우에는 분노, 기쁨, 슬픔과 같은 감정의 상태도 포함된다. 중앙은행(한국은행)에서 대중에게 알리는 한국은행 금융통화위원회 의사록에는 중앙은행이 시행하는 통화정책의 방향, 현 경제상황에 대한 판단이 포함되어 있음 이러한 중앙은행의 의사록은 곧 대중들 사이의 소통, 즉 커뮤니케이션이라고 볼 수 있는데, 이는 시장에 직접적인 영향을 줄 수 있으므로 글로벌 금융 위기 이후 관심 증대 그러나, 중앙은행의 커뮤니케이션은 절제된 표현이 많아 일반적인 독해로는 명확한 의미를 파악하는 것이 힘듦

따라서, 텍스트 마이닝을 활용하여 이 의사록에 담긴 어조를 추출하여 지수화(수치화) 하고, 이를 통해 기준 금리의 변동에 얼마나 많은 영향을 미쳤는지에 대한 설명력과 예측력을 검증

ROTHIG

연구 목표

경제/경영 분야에 대한 새로운 감정 사전 제작 텍스트 마이닝을 활용하여 한국은행 금융통화위원회 의사록의 어조 분석 경제/경영 분야에서 한국어 텍스트 마이닝 활용방안 제시





Newspaper, Report, Deciphering Monetary

When? 2005~2017

How many? 230,000

How? Lemmatisation

한국어 특성상, 형용사, 동사 등 변형이 일어나는 형태소의 경우에 표제어를 통일하는 작업 수행

ex) [오른,오른다고,오른다는,올,올라서] 등의 경우 표제어 ['오르']로 처리



개별적인 단어 하나, 혹은 부정적인 단어와 부정적인 단어가 결합되면 의미 파악이 어려웠기 때문에 연속된 n개의 단어 조합 사용

ex) 낮은 + 실업률

그러나 무작정 n-gram 길이를 늘리는 것은 trade-off와 차원의 저주를 발생시킴

형태소조합 N-gram





5-gram 단어의 극성(polarity)을 구분하여 Positive/Negative를 나타내는 lexicon 생성





Polarity Lexicon

Market approach

Lexical approach

Supervised learning approach

Market approach

금통위 의사록이 공개된 시점의 시장변수의 움직임만으로 대중이나 은행의 주관적 판단을 배제하고단어의 극성을 파악 가능

시장변수(주식)을 독립변수, 단어를 종속변수로 두어 단어의 양수크기가 더 큰 것에 가중치를 준다.

실행결과 67% 의 정확도

Naive Bayes Classification

 $\frac{P(B|A)P(A)}{P(A|B)} = \frac{P(B|A)P(A)}{P(B)}$

조건부 확률을 사용하는 확률 모델 모형의 단순함에도 불구하고, 실제 데이터에 잘 적합

실행결과, 86% 의 정확도

Lexical approach

Market Approach와는 달리, 단어들 간의 관계에 초점을 맞추는 접근방식. 어휘적 접근법 직관적이지만, 반어법과 같은 단어를 구분하지 못함

실행결과 67% 의 정확도

센티멘트

Sentiment

감정점수라고 부르며, 극성(polarity)를 측정한 점수



분석방법 Analysis Method

1. Descriptive Statistics

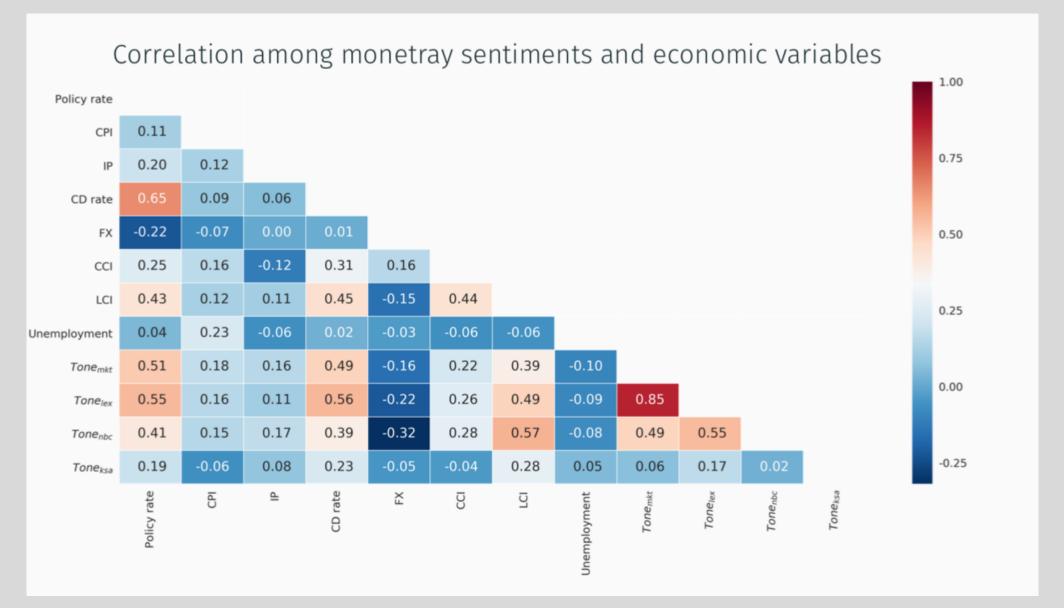
- a. Correlation Heatmap
- b. Descriptive Statistics
- c. Correlation matrix

2. Linear Regression (t0, t1, t2)

- a. Ordered Probit Models
- b. OLS



Correlation Heatmap



Descriptive Statistics

Table 3: Descriptive statistics of economic variables

Statistic	Mean	Median	St. Dev.	Min	Max
BOK_t	2.892	2.750	1.153	1.250	5.250
ΔBOK_t	-0.014	0.000	0.163	-1.000	0.250
π_t	0.306	0.156	1.202	-1.632	3.736
$\Delta\pi_t$	-0.011	-0.013	0.345	-0.889	0.852
$(y_t - y^*)$	0.0001	0.003	0.039	-0.228	0.065
$\Delta(y_t - y^*)$	-0.00001	0.001	0.023	-0.116	0.066
$\Delta\pi_t^e$	-0.005	0.000	0.134	-0.300	0.400
Δy_t^e	0.012	0.000	0.276	-0.600	1.000
Tone _{mkt}	-0.067	-0.111	0.465	-0.867	0.842
Tone _{lex}	-0.072	-0.096	0.295	-0.800	0.679
Tone _{nbc}	-0.388	-0.393	0.155	-0.833	0.053
Tone _{ksa}	0.550	0.554	0.101	0.178	0.818

Correlation matrix

Table 4: Correlation matrix of economic variables

	BOK _t	Δ BOK $_t$	π_t	$\Delta\pi_{t}$	$(y_{t} - y^{*})$	$\Delta(y_t - y^*)$	$\Delta\pi_t^e$	Δy_t^e	Tone _{mkt}	Tone _{lex}	Tone _{nbc}
BOK _t											
Δ BOK $_t$	0.12										
π_t	0.57***	-0.06									
$\Delta\pi_t$	0.06	0.10	0.14								
$(y_t - y^*)$	0.26**	0.54***	0.04	0.22**							
$\Delta(y_t - y^*)$	-0.08	0.25**	-0.09	0.06	0.30***						
$\Delta\pi_t^e$	0.19*	0.14	0.29***	0.21*	0.23**	0.03					
$\Delta y_t^{\triangleright}$	-0.23**	0.16	-0.28***	-0.09	-0.02	0.31***	-0.25**				
Tone _{mkt}	0.53***	0.51***	0.40***	0.12	0.40***	0.08	0.13	0.05			
Tone _{lex}	0.33***	0.56***	0.16	0.19*	0.48***	0.06	0.11	0.09	0.85***		
Tone _{nbc}	-0.02	0.43***	-0.04	0.27**	0.42***	0.14	-0.06	0.23**	0.49***	0.55***	
Tone _{ksa}	-0.16	0.23**	-0.39***	0.06	0.18*	0.09	0.13	0.23**	0.06	0.17*	0.01

Linear Regression (t0, t1, t2) Ordered Probit Models

Table 5: Results from ordered probit models of monetary sentiments score at t_0

	Dependent variable: MPD _t								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
MPD _{t-1}	0.616**	0.578**	0.539**	-0.718*	- 0.557	0.277	0.525*		
	(2.326)	(2.146)	(1.969)	(-1.778)	(-1.462)	(0.941)	(1.898)		
$\Delta(\pi_{t}-\pi^{*})$		0.172	0.042	-0.390	-0.534	-0.356	0.044		
		(0.520)	(0.125)	(-0.749)	(-1.226)	(-0.952)	(0.130)		
$\Delta(y_t - y^*)$		8.374*	6.785	7.757	11.688*	6.871	6.839		
		(1.797)	(1.370)	(1.285)	(1.797)	(1.301)	(1.374)		
$\Delta\pi_t^e$			1.900 * *	1.952	1.897*	2.295**	1.852**		
			(2.071)	(1.524)	(1.680)	(2.400)	(1.996)		
Δy_t^e			0.400	0.049	-0.004	0.036	0.355		
			(0.898)	(0.076)	(-0.007)	(0.076)	(0.763)		
Tone _{mkt,t}				5.466***					
шкс, с				(4.884)					
Tone _{lex,t}				(,	4.451***				
lex, t					(5.725)				
Tone _{nbc,t}					(0.720)	2.921***			
noc,t						(3.300)			
Tone _{ksa,t}						(5.555)	0.399		
RSa ,t							(0.341)		
. 2									
Pseudo R ²	0.037	0.062	0.092	0.45	0.371	0.162	0.092		
Observations	145	145	145	145	145	145	145		

Linear Regression (t0, t1, t2) OLS

Table 8: Results from OLS of monetary sentiments score at t_0

	Dependent variable: ΔΒΟΚ _t								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
Δ BOK $_{t-1}$	0.334***	0.301***	0.282***	0.097	0.022	0.161**	0.261***		
	(4.237)	(3.846)	(3.592)	(1.236)	(0.281)	(2.030)	(3.293)		
$\Delta(\pi_t - \pi^*)$		0.027	0.021	0.007	-0.014	-0.024	0.020		
		(0.737)	(0.572)	(0.200)	(-0.435)	(-0.640)	(0.534)		
$\Delta(y_t - y^*)$		1.449***	1.196**	1.163 * *	1.333***	1.141**	1.211**		
		(2.673)	(2.104)	(2.262)	(2.689)	(2.125)	(2.139)		
$\Delta\pi_t^e$			0.158	0.106	0.109	0.200**	0.134		
			(1.571)	(1.155)	(1.240)	(2.095)	(1.320)		
Δy_t^e			0.067	0.063	0.048	0.029	0.045		
			(1.327)	(1.381)	(1.095)	(0.594)	(0.866)		
Tone _{mkt,t}				0.156***					
тис, с				(5.650)					
Tone _{lex,t}					0.298***				
ιελ, τ					(6.727)				
Tone _{nbc,t}						0.374***			
HDC, C						(4.248)			
Tone _{ksa,t}						(0.207		
rsa,t							(1.551)		
Observations	145	145	145	145	145	145	145		
Adjusted R ²	0.105	0.141	0.149	0.304	0.355	0.242	0.158		

센티멘트

Sentiment

감정점수라고 부르며, 극성(polarity)를 측정한 점수



통화 감정 분석

Monetary Sentiment Analysis

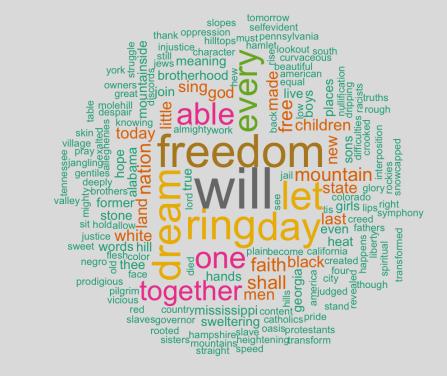
모든 문장의 톤을 4개의 방법으로 측정하고, 센티멘트 점수를 합산하여 산출 Tone.mkt, Tone.lex, Tone.nbc, Tone.ksa



영어 감정 분석

Comparison with English Sentiment Analysis

구글 번역기로 금통위 의사록의 모든 문장을 영어로 번역하고, 각 문장의 톤을 3가지의 방법으로 측정(Tone,LM, HIV4, Google) 그 이후 모든 문장의 센티멘트 점수를 합산하여 산출



경제 불확실성 분석

Economic Uncertainty Analysis

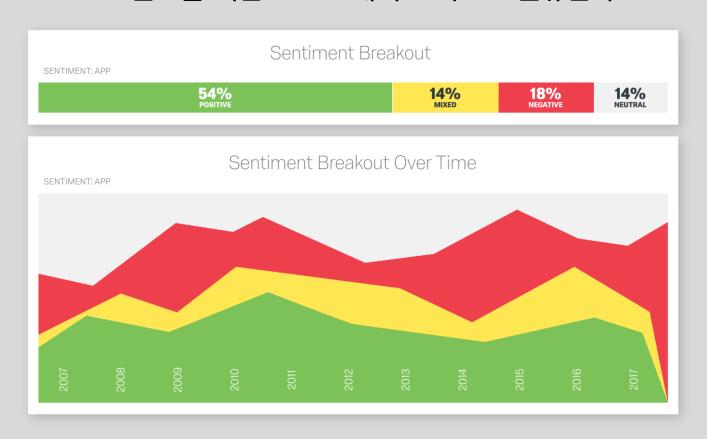
Polarity lexicon 생성 과정처럼 Market approach와 Lexical approch를 활용하여 Uncertainty Lexicon을 생성하여 분석



주제 표현 분석

Topic Sentiment Analysis

문장별로 센티멘트점수 산출하여 그 분포를 알아보고, 그 분포를 기반으로 36개의 토픽으로 분류한다



연구자들은 자체 제작한 경제/금융 형태소 사전과 n-gram을 사용하여 통화정책에서 나타나는 정서(감정)을 정량화 하는 텍스트 기반 지표를 개발

개발한 지표는 현재 및 미래의 통화 정책을 다른 지표들보다 훨씬 더 잘 설명할 수 있음 특히 한국어로 수행한 연구는 다른 언어의 연구보다 높은 효율을 보여줌 금통위 의사록에서 추출해낸 지수는 여타 변수에 비해 기준금리에 대한 설명력과 예측력이 높은 것으로 나타남

또한, 기존 테일러 준칙의 GDP갭률과 인플레이션 등과 함께 추출한 지수를 설명 변수로 추가할 경우, 현재, 그리고 미래의 금리도 상당부분 설명 가능 만약, 금통위 의사록이 현재 시행되고 있는 기준금리나 통화정책 심리 지표에 잠재적으로 영향을 미치는 변수라고 해석한다면, 이러한 통화정책의 효과를 분석하는 VAR시스템이나 DSGE 모델에 우리의 연구를 적용하면 좋을 것 또한 이 연구결과는 중앙은행의 의사소통이 얼마나 대중들과 효율적으로 소통하는지를 평가하는 자체적인 지표로 사용될 수 있다. 중앙은행이 기존에 의도한 바는 상관없이, 연구자들의 지표는 통화 정책의 미래 결정을 설명하는 데 도움을 줄 수 있다.

마지막으로, 연구자들이 수행한 방법은 수치적으로 측정하기 어려운 거시 경제적 불확실성, 미래 통화 정책 태도에 대한 대중들의 기대, 주식 시장의 감정을 측정하기 위해 다른 지표를 구성하는 데에도 보다 쉽게 적용할 수 있다. (Text 기반이기 때문)



Thank You

12864