

키워드 기반 금융 정보 요약 서비스 구현

CUAI 4기 D-B팀

김민지(응용통계학과), 김민주(경영학부), 이건이(응용통계학과)

[요약] 본 프로젝트의 목적은 금융상품의 약관에 대해 금융소비자의 이해력을 신장하는데 있다. 이를 위해 금융상품 약관 별 키워드를 추출, 요약하여 금융소비자에게 제공해 소비자의 합리적인 의사결정에 도움을 준다. 키워드 추출에는 KeyBERT를 요약에는 TextRank를 사용했다.

1. 서 론

금융 소비자에 대한 설명 의무가 금융소비자법 6대 판매 원칙으로 제정되었다.

설명 의무

■ 중요사항 설명의무

- 금융상품판매업자 등은 일반금융소비자에게 계약 체결을 권유(금융상품판매업자가 자문에 응하는 것을 포함)하는 경우 및 일반금융소비자가 설명을 요청하는 경우에는 해당 금융상품에 관한 중요한 사항(일반금융소비자가 특정 사항에 대한 설명만을 원하는 경우 해당 사항으로 한정)을 일반금융소비자가 이해할 수 있도록 설명해야 합니다(「금융소비자 보호에 관한 법률」 제19조제1항).
- 금융상품판매업자 등은 설명을 할 때 일반금융소비자의 합리적인 판단 또는 금융상품의 가치에 중대한 영향을 미칠 수 있는 사항을 거짓으로 또는 왜곡(불확실한 사항에 대해 단정적 판단을 제공하거나 확실하다고 오인하게 할 소지가 있는 내용을 알리는 행위를 말함)하여 설명하거나 빠뜨려서는 안 됩니다(「금융소비자 보호에 관한 법률」 제19조제3항 및 「금융소비자 보호에 관한 법률 시행령」 제13조제9항).

※ 금융상품판매업자 등이 일반금융소비자에게 금융상품 유형별로 설명해야 하는 중요한 사항(상품의 내용, 위험, 이차율, 수익률, 금리 등)은 「금융소비자 보호에 관한 법률」 제19조제1항 및 「금융소비자 보호에 관한 법률 시행령」 제13조에서 확인할 수 있습니다.

■ 설명서 제공

- 금융상품판매업자 등은 위의 설명에 필요한 설명서를 일반금융소비자에게 제공해야 하며, 설명한 내용을 일반금융소비자가 이해하였음을 서명, 기명날인, 녹취 또는 그 밖의 방법으로 확인을 받아야 합니다(「금융소비자 보호에 관한 법률」 제19조제2항 본문).
- 다만, 금융소비자 보호 및 건전한 거래질서를 확립하기 위하여 필요할 경우에는 자문에 대한 답변 또는 세부정보 확인 방법을 제공하거나 기존 계약과 동일한 내용으로 계약을 갱신하는 경우 등에는 설명서를 제공하지 않을 수 있습니다(「금융소비자 보호에 관한 법률」 제19조제2항 단서 및 「금융소비자 보호에 관한 법률 시행령」 제14조제4항).

↑ 금융상품 6대 판매원칙 중 '설명 의무'에 대한 설명 (참고문헌 1번)

금융상품판매업자 등은 일반금융소비자에게 계약 체결을 권유하는 경우 및 일반금융소비자가 설명을 요청하는 경우에는 해당 금융상품에 관한 중요한 사항을 일반금융소비자가 이해할 수 있도록 설명해야 한다. 또한, 비대면 거래의 증가로 소비자의 충분한 약관 숙지가 어렵다는 문제가 존재한다. 즉, 쉽고 간단하게 중요한 정보를 전달하는 것이 중요하다고 할 수 있다.

따라서 본 논문에서는 금융 소비자 권익 제고를 위해, 약관 등 어려운 금융 정보를 한눈에 보기 쉽게 전달하는 키워드 기반 금융정보 요약 서비스를 제안한다.

2. 본 론

키워드 추출 및 요약에 사용된 데이터는 아래의 표와 같다.

[표1] 사용된 데이터

상품명	특징
KB내담대로적금(국민)	고객이 직접 상품설계
IBK_D day적금(기업)	고객이 정한 기한일에 맞춘 금액 설정
아름다운 용기 적금(신한)	친환경 서약, 고령자 우대 등 ESG상품
WON 적금(우리)	스마트폰 전용 가입 상품
급여하나 월복리 적금(하나)	월 복리 적용 상품
청년우대 주택청약(공통)	은행 공통 약관
26주 적금(카카오)	26주 계약기간, 꾸준한 납부 시 금리 상승
모임통장(카카오)	모임회비 관리 유용한 상품
세이프박스(카카오)	계좌안에 여유자금을 따로 보관
저금통(카카오)	소액 자동 저축해 최대 10만원까지 잔돈 저축

1) KeyBERT를 이용한 약관 키워드 추출(토픽 모델링)

약관의 키워드를 추출하기 전에, 형태소 분석기를 통해 명사만 추출한 문서를 만든다. KoNLPy의 오픈 소스 한국어 형태소 분석기 Okt를 이용하면 아래와 같은 결과를 얻을 수 있다. 카카오의 '26주 적금'을 예시로 들었다.

[('카카오', 'Noun'), ('뱅크', 'Noun'), ('26', 'Number'), ('주', 'Noun'), ('적금', 'Noun'), ('특약', 'Noun'), ('Wn', 'Foreign'), ('제', 'Noun'), ('1', 'Number'), ('...', 'Punctuation'), ('절차', 'Noun'), ('를', 'Josa'), ('거쳐', 'Verb'), ('제공', 'Noun'), ('됩니다', 'Verb'), ('.', 'Punctuation')]

위와 같이 품사가 태그되어 나오면, 이 중에서 명사(Noun)만 선택해준다.

['카카오', '뱅크', '주', '적금', '특약', '제', '조', '약관', '적용', '카카오', '뱅크', '적금', '이하', '...', '약관', '금융', '소비자', '보호', '관', '법률', '및', '카카오', '뱅크', '내부통제', '기준', '절차', '제공']

위에서 선택된 명사들을 공백과 함께 이어 붙여 명사만 들어있는 문서 형태로 만들어주고, KeyBERT의 extract_keywords() 모듈을 이용해 키워드를 추출한다. KeyBERT는 사용법이 매우 간

단한 키워드 추출 기술이다. 이 때 keyphrase_ngram_range 인자를 통해 몇 단어짜리 키워드를 추출할 지 설정해줄 수 있고, stop_words 인자를 통해 불용어를 설정해줄 수 있다.

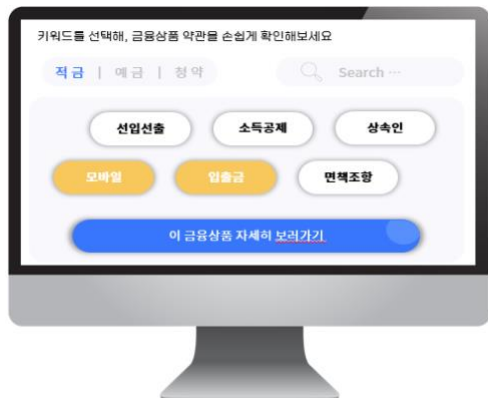
‘카카오’ 등 자주 등장하지만 크게 의미가 없는 단어를 stop_words로 설정해주고 keyphrase_ngram_range=(1, 1)로 설정하면 다음과 같은 결과가 나온다.

[('홈페이지', 0.7467), ('내부통제', 0.7443), ('입출금', 0.7104), ('모바일', 0.705), ('법적', 0.6987)]

keyphrase_ngram_range=(1, 2), 즉 두 글자 키워드를 추출하면 다음과 같은 결과를 얻을 수 있다.

[('법률 내부통제', 0.8379), ('만기일로부터', 0.8283), ('중액 입출금', 0.8278), ('법적 지급', 0.8266), ('금융 소비자', 0.8226)]

각 약관의 키워드에서 모델이 중요하다고 판단했지만 ‘법정공휴일’, ‘영업일’, ‘뱅킹’ 등 큰 의미가 없는 단어들을 제외해준 뒤 서비스에 적용할 수 있다. 서비스에 적용했을 때의 예시는 다음과 같다



아래와 같이 각 키워드 별로 약관 내용을 확인할 수 있다. (편의상 각 상품 약관 별로 구별하여 서술했다.)

1. 국민은행 - KB내맘대로적금
보이스피싱: 적금 가입자는 보험가입서비스를 받을 수 있고, 4가지 플랜 중 플랜B[피싱] 선택 시 보이스피싱 피해 최대 1000만원을 보상 받을 수 있다. 명의 변경: 명의변경으로 인한 예금주 변경 시 보험가입서비스가 제공되지 않는다.

2. 기업은행 - IBK-D day 적금
적금 종류 정기적금: 이 적금의 종류는 정기적금(자유적립식)으로 한다.
전용 상품 영업: 이 통장은 비대면채널 전용상품으로 영업점 창구를 통해서 가입할 수 없고, 실물통장이 발급되지 않는다.

3. 신한은행 - 아름다운 용기 적금
선입선출: 이 예금은 만기일 전영업일까지 2회에

한하여 일부해지가 가능하며, 일부해지금에 대한 이자는 선입선출법에 따라 제12조에 의해 지급한다. (제12조 중도해지 이자율 · 예금주가 만기일 전에 예금을 찾고자 하는 경우에는 입금일로부터 예금을 찾는 날 전날까지의 기간에 대하여 가입일에 영업점 및 인터넷 홈페이지에 고시된 중도해지이자율에 의해 지급합니다.)

4. 공통 - 청년우대 주택 청약
포함 금융기관: 이 저축의 가입은 청약저축, 청약예금, 청약부금 및 주택청약종합저축을 포함하여 전 금융기관 1인 1계좌에 한한다.

소득공제 저축: 제9조 소득공제/이 저축은 소득공제 가능한 상품으로 「조세특례제한법」 제87조에서 정한 요건에 해당하는 경우 소득공제를 받을 수 있다. 전환 임대주택: 제3조 납입기간/이 저축의 납입기간은 가입한 날부터 국민주택(분양전환되지 않는 임대주택은 제외)과 민영주택의 입주자로 선정된 날까지로 한다.

5. 카카오뱅크 - 26주 적금
입출금 통장: 제8 조 26주적금서비스/26주적금서비스는 매주마다 자동이체 금액이 자동으로 증액되어 카카오뱅크 입출금통장에서 이 적금으로 자동이체하는 서비스이다.
명의 변경: 이 적금은 양도에 의한 변경은 불가하나, 상속에 의한 명의변경은 가능하다.

2) Textrank를 통한 약관 요약

본 연구에서는 금융상품 약관을 요약하기 위해 다음과 같이 진행하였다. 금융상품 약관은 다른 Text데이터와 달리 그 수가 적어서 적어 학습을 충분히 하기 어렵다는 판단 하에 Textrank 알고리즘을 사용해 문서요약을 진행하였다. Textrank 알고리즘은 Google의 Pagerank 알고리즘에서 착안한 비지도기반의 알고리즘이다. Pagerank는 중요한 웹 문서일수록 다른 페이지로부터 링크를 많이 받는다는 것을 전제로 하이퍼링크를 통해 상대적 중요도에 따라 웹문서에 가중치를 부여하는 알고리즘으로, 이를 텍스트에 적용한 것이 Textrank알고리즘이다. Textrank에서는 각 단어 혹은 문장 등을 노드로 간주하고 이를 토대로 각 단어 혹은 문장의 가중치를 계산하여 중심을 찾아 이를 바탕으로 요약을 진행한다. 중요도는 다음과 같이 계산된다.

$$TR(V_i) = \frac{w_{ji}}{\sum_{v_k \in \text{In}(V_i)} \sum_{v_k \in \text{Out}(V_j)} w_{jk}} TR(V_j) \cdot (1 - d) + d \cdot \sum_{v_k \in \text{In}(V_i)} \sum_{v_k \in \text{Out}(V_j)} w_{jk} TR(V_j)$$

위의 식에서 V_i 는 임의의 단어를 V_j 는 단어 i와 인접한 단어, V_k 는 단어 j와 인접한 단어를 말한다. $TR(V_i)$ 는 현재 단계에서 단어 i의 중요도, $TR(V_j)$ 는 이전단계에서 단어 j의 중요도를 말한다. w_{jk} 와 w_{ji} 는 가중치를 나타낸다. 위의 계산식을 바탕으로 문서별로 중요한 문장을 추출할 수 있다.

Textrank 알고리즘을 구현하기 위해 python의

gensim패키지의 summarize를 사용했고, 요약문의 길이를 본문의 길이에 따라 정하는 summarize의 옵션으로 ratio를 0.05와 디폴트 값인 0.2를 사용했다. 각 약관에 대해 요약을 실시한 결과, 몇 가지 예시를 보자면 다음과 같다.

1. 급여하나 월복리 적금 약관 요약(ratio=0.05사용)

이 예금은 적립금마다 입금일부터 만기일 전날까지의 기간에 대하여 예금 가입일(또는 재예치일) 당시 영업점 및 은행 홈페이지에 고시한 기본금리에 제 9조 및 제 10 조에서 정한 금리를 더하여 월복리로 계산한 이자를 만기일에 일시 지급합니다.

②이 예금의 만기일 전에 지급을 청구한 때에는 적립금마다 입금일부터 지급일 전날까지 기간에 대하여 예금가입일(또는 재예치일) 당시 영업점 및 은행 홈페이지에 고시한 중도해지금리를 단리로 썬한 이자를 지급합니다.

2. 26주 적금 약관 요약(ratio=0.2사용)

이 적금은 가입일 당시 카카오뱅크 모바일앱 이하 모바일앱 이하 한다) 및 인터넷홈페이지에 게시된 이 적금의 계약기간별 금리를 적용합니다 .

3. 자동이체 종료일은 이 서비스 가입일로부터 26주에 해당하는 날이며 변경은 불가합니다

5. 이 서비스는 모바일앱 자동이체 조회 화면에서 조회 및 해지가 가능합니다

①제 8 조의 26 주적금서비스를 통해 1 주차부터 7 주차까지 자동이체 납입에 연속 성공한 경우 연 0.20%p 제공

②제 8 조의 26 주적금서비스를 통해 1 주차부터 26 주차까지 자동이체 납입에 연속 성공하는 경우 연 0.30%p 추가 제공

① 이 적금의 가입일로부터 만기일 이전까지 고객이 만기 시 자동해지에 동의한 경우에는 만기일에 본인의 실명확인된 입출금이자유로운예금 이하 근거계좌 라 한다 으로 이 적금을 자동해지하여 입금 처리 합니다

이 적금의 긴급출금은 가입일로부터 만기일 전일까지 최대 2 회까지 가능하며, 긴급출금 시에는 먼저 납입한 금액부터 출금됩니다.

만기일 전에 해지를 요청하는 경우 가입일로부터 해지일 전일까지의 기간에 대하여 가입일 당시 모바일앱 및 인터넷홈페이지에 게시한 기간별 중도해지금리를 적용합니다.

만기일 이후 해지를 요청하는 경우 만기일로부터 해지일 전일까지의 기간에 대하여 가입일 당시 모바일앱 및 인터넷홈페이지에 게시한 기간별 만기후금리를 적용합니다.

위의 2개의 결과는 각 상품의 특성을 보여주는 조항을 중심으로 요약이 진행되었다.

3. 아름다운 용기 적금 약관 요약(ratio=0.05)

예금주가 만기일 전에 예금을 찾고자 하는 경우에는 입금일로부터 예금을 찾는 날 전날까지의 기간에 대하여 가입일에 영업점 및 인터넷 홈페이지에 고시된 중도해지이자율에 의해 지급합니다.

이 예금은 만기일 전영업일까지 2 회에 한하여 일부해지가 가능하며, 일부해지금에 대한 이자는 선입선출법에 따라 제 12 조에 의해 지급합니다.

4. IBK_D day 적금 약관 요약(ratio=0.05)

이 적금은 만기일 이전에 중도해지를 신청하는 경우 입금액마다 입금일부터 해지일 전일까지의 기간에 대하여 가입일 당시 은행이 고시한 중도해지금리를 적용합니다.

이 적금은 만기일로부터 지급일 전일까지의 기간에 대하여 만기일 당시 은행이 고시한 만기 후 금리를 적용합니다.

반면 아래의 2개의 결과는 상품의 특성을 보여주는 요약이 진행되지 않았다.

3. 결 론

본 프로젝트에서는 총 10개의 금융상품의 약관에 대해, 약관 별로 키워드를 추출하고, 문서 요약을 시도하였다. 코로나 이후, 언택트 시대가 더욱 가속화되면서 금융 또한 마찬가지로의 변화를 겪고 있는 시점에 금융상품에 대한 금융소비자들의 이해력은 더욱 요구된다. 따라서, 어렵다고 느껴지는 금융상품에 대한 진입장벽을 낮추기 위해서는 각 약관별로 핵심을 요약해주는 것이 금융소비자에게는 편리한 의사결정 및 소비자의 권리보호라는 측면에서, 금융판매자에게는 증명책임전환 및 비용감소라는 측면에서 두 주체 모두 유익한 의사결정이 가능해진다. 또한 추출된 키워드와 요약문과 사용자의 관점에서 접근성을 높이는 UX Writing과의 연계를 통해 AI UX Writer를 구현, 각 개인별 눈높이에 맞는 서비스를 제공해 건강한 금융활동에 도움이 될 것이다.

이를 위해 보완해야할 점은 다음과 같다. 첫째, 키워드 추출 및 요약의 성능을 더 높인다. 프로젝트 결과 하나은행의 '급여하나 월복리 적금'과 같이 요약 결과가 좋은 것도 있었으나, 기업은행의 'IBK_D day'와 같이 결과가 좋지 않은 것도 있었다. 따라서, 키워드 추출 및 요약의 성능을 더욱 향상시키기 위해, 더 많은 금융약관 데이터를 가지고 kobertsum, ke-t5등의 다양한 문서요약 알고리즘을 실행하는 것이 추후 과제로 남아있다. 둘째, 어려운 금융용어를 알기 쉽게 변환해야 한다. 금융용어 대부분은 일상에서 자주 쓰이지 않는 말로, 소비자자들이 이를 한번에 이해하는 것은 힘들다. 따라서 '금융감독원'의 '알기쉬운 금융용어' 등을 이용해 용어를 변환하는 작업이 필요하다.

참고 문헌

1)

<https://easylaw.go.kr/CSP/CnpClsMain.laf?popMenu=ov&csmSeq=1530&ccfNo=1&cciNo=2&cnpClsNo=1> : 금융상품 6대 판매원칙.

2)

배원식, 차정원 “Textrank 알고리즘을 이용한 문서 범주화”, 컴퓨팅의 실제 및 레터 16, 2010.01, 110-114

이기욱, 금소법 시행령 모호성에 금융사 ‘가우뚱’...“설명의무 조항, 보완 필요”, 시사저널e

<http://www.sisajournal-e.com/news/articleView.html?idxno=225062>

(2022.01.22 방문)

금융위원회, 「금융상품 설명의무의 합리적 이행을 위한 가이드라인」

박은경 (2021). 비대면 보험판매와 보험자의 약관 설명의무. 법학연구, 21(1), 21-45

한국은행, 금융안정보고서, 2020.10

Federico Barrios(2016), Variations of the Similarity Function of TextRank for Automated Summarization