미래에셋 빅데이터 페스티벌 보고서

팀명: Team25(25)

학교: 숭실대학교

김성수(010-3881-8691)

윤지원(010-2472-3159)

정지현(010-2291-0670)

주제: 음성기반 금융 서비스 개발

Extension ID: com.mechat.finance

# 서론

금융투자협회 전자공시에 따르면, 현재 증권사의 국내 영업점은 2018년 1분기 기준 1001개로 2017년 말 대비 125개 감소했다. 이러한 증권사 영업점의 급격한 감소로 종합자산관리가 가능한 대형 점포를 강화 시키거나 비대면 거래를 강화 시키는 전략을 취하고 있다. 2016년 2월 도입한 증권사 비대면 계좌의 수는 1년만에 누적 75만 개를 돌파하였으며 올해 하반기 비대면 계좌 일임의 허용으로 비대면 채널을 사용한 고객 마케팅 및 서비스 제공의 중요성이 더욱 커졌다. 현재 증권사들은 기존의 웹 인터페이스를 사용자 친화적으로 개선, 어플리케이션(이하 “앱”) 개발, 카카오톡 플러스친구 계정을 운영 및 Clova 확장서비스를 통한 비대면 채널을 제공하고 있다.

현재 증권사들은 모바일 및 IoT 친화적인 신규 비대면 채널을 개설하고 있으며 그 종류는 크게 메신저(문자) 기반의 카카오톡 플러스친구와 음성 기반의 CEK가 있다. 카카오톡 플러스친구를 운영하고 있는 증권사의 예인 키움증권의 경우 고객이 질문을 하는 것이 불가하며, 매달 이벤트와 당월 혜택 소식을 링크형식으로 알려주는 방식으로 운영되고 있다. 다른 예시인 삼성증권 플러스 친구는 7개의 기능 버튼 인터페이스를 제공한다. 각 버튼을 클릭하면 해당 삼성증권 웹 페이지 또는 모바일 앱으로 이동 만들었다. CEK를 통한 음성인식 채널을 만드는 증권사로는 미래에셋 대우와 신한금융투자 알파 (이하”신한 알파”)가 유이하다. 미래에셋 대우의 경우 국내, 해외 시황을 알려주는 것에 한정된 서비스를 하고 있다. 신한 알파의 경우 종목검색, 증시 알림, 신한지주 시세 검색, 총 3가지 검색 기능을 제공하며 대화는 1번의 질문과 1번의 응답으로만 구성되어 있다.

음성인식 인터페이스는 “가장 자연스러운 인간과 기기간의 상호작용 방식”(Wired, 2016. 6.24)으로 평가되며 입력속도는 터치/텍스트 기반 인터페이스의 3배를 능가하고, Hands-free 기능으로 멀티태스킹이 가능하다. 텍스트 기반에서 양방향 음성 응답 인터페이스로의 이동은 금융투자 상품에 대한 장벽을 낮출 수 있다. 음성 인식 인터페이스가 가져올 즉각성, 효율성, 편리성을 증권투자 서비스에 더한다면, 사용자는 더욱 효율적인 방식으로 금융투자 인사이트를 얻고, 증권회사는 자동화된 서비스로 운영 비용을 효과적으로 줄일 수 있다.

현재 상용화된 비대면 채널들은 기능과 현재 증권사의 고객들과 잠재 고객들에게 유치할 만큼의 기능적, 편의적 매력도가 떨어진다. 이러한 이유로 본 팀은 CEK를 활용한 금융 서비스를 개발 하였다.

## 기능설명

GUI기반의 카카오톡 챗봇의 경우 버튼식 UI를 사용하게 되면 처음 사용하는 유저도 쉽게 사용할 수 있다. 반면, 음성 시스템의 경우 시각적인 인터페이스가 없기 때문에 처음 사용하는 사람들은 이 확장 서비스를 사용법의 무지로인한 당혹감을 느낄 수 있다. 이러한 문제를 극복하기 위해 Clova 앱의 경우 3개의 대표적인 사용법을 등록할 수 있으나 에이전트의 기능이 3가지보다 많을 경우, 사용자들이 그 외의 기능을 알 수 있는 길이 없다. 그래서 모이라는 사용자들에게 서비스 사용법을 제공하기 위해서 “기능 설명” 기능을 포함하였다. 이로써 시각적 가이드가 없는 음성인식 인터페이스의 단점을 즉각적인 음성 응답 서비스로 보완하였다.

사용자가 모이라의 기능 설명 Intent를 입력하면. 이 Intent는 모이라의 모든 기능을 나열하는 기능과 각 기능을 자세히 설명해 주는 기능을 구현한다. 불특정 기능을 탐색하는 발화는 아무 정보를 포함하지 않기에 Slot을 포함하지 않고, 특정 기능의 사용법을 문의하는 발화는 총 5가지Slot 값을 갖고 있는 menu Slot과 “문의” 와 유사한 동사로 이뤄져 있다. “기능 설명” Intent로 하여금 사용자들이 모이라를 사용하는 과정에서 효율적이고, 사용성 높고, 정확하게 의사소통할 수 있게 설계하였다.

## 펀드추천

펀드 추천은 사용자가 투자하고 싶은 펀드 상품을 펀드 종류와 수익률 기간에 따라 추천해주는 기능이다. 대화는 사용자가 펀드추천 요청하고 펀드 종류 선택한 후 수익률 기간 설정하여 모이라에게 펀드를 추천 받을 수 있도록 구성되어 있다. 펀드 추천 기능을 활용한 대화 가능 대상 시나리오는 다음과 같다.

1. 신규 유저 시나리오

본 서비스에 익숙하지 않은 사용자들에 발생 가능한 시나리오이다.

1. 사용자가 가장 기본적인 질문으로 대화를 시작한다

Ex. 요즘 가장 괜찮은 펀드는 뭐야?

1. 사용자가 원하는 펀드 종류를 알기 위해서 펀드 종류를 묻는 응답을 반환한다.

Ex. 어떤 종류의 펀드를 추천해 드릴까요?

1. 채권형, 인덱스, ETF, MMF, 혼합형, 주식형 총 6가지 펀드 종류에서 한 가지를 사용자가 골라 응답한다.

Ex. 채권형으로 부탁해.

1. 사용자가 원하는 상품의 수익률 기준을 얻기 위해 수익률 조회 기간을 묻는다.

Ex. 수익률 조회 기간은 어떻게 되나요?

1. 사용자는 6개월, 1년 중에서 조회기간을 선택한다.

Ex. 6개월로 부탁해.

1. 사용자가 선택한 펀드 종류와 수익률 조회 기간을 기준으로 상위 5가지 상품을 음성으로 응답해준다.
2. 숙련자 시나리오

모이라의 펀드추천 기능을 많이 이용해본 사용자들의 경우 질문의 시간을 단축 시키기 위해 해당 기능을 펀드추천 요청 – 펀드 타입 설정 – 수익률 조회 기간 설정의 3단계로 진행하는 것이 아닌 한 문장에 응답에 필요한 펀드 타입과, 수익률 조회 기간을 넣어 질문하는 경우가 있을 것이다. 또는 펀드를 고르는 일차적 기준이 펀드타입과 수익률인 개인 투자 고수들을 위한 시나리오이다.

1) 사용자가 원하는 펀드타입과 수익률 조회 기간을 말한다.

2) 모이라는 응답 받은 타입과 조회 기간을 기준으로 수익률 상위5개 펀드를 추천해준다.

펀드 추천 기능을 구현을 위해서 펀드추천 Intent, 펀드타입 Intent, 그리고 기간설정 Intent를 설계했다.

1. 펀드추천 Intent

이 Intent는 사용자들의 펀드 검색 의사를 확인하기 위해 설계되었다. 본 대화 Intent가 작동되기 위해 필요한 Slot이 없기 때문에 펀드를 수식하는 형용사, 펀드상품을 지칭하는 명사, 그리고 “추천” 또는 “검색” 의미의 동사를 조합하여32개의 발화 예시를 작성했다.

1. 펀드타입 Intent

사용자가 원하는 펀드 종류를 입력 받기 위해 만든 Intent이다. 본 Intent는 사용자가 FundType Slot를 필수적으로 응답하도록 구성돼 있다. 사용자가 응답하여야 하는 FundType Slot은 혼합형, 채권형, MMF형, ETF형, 인덱스형, 주식형 총 6가지가 있다. FundType Slot 1 개와 “추천”이라는 동사의 조합으로 이루어진 발화 예시 67개를 구성하였다.

1. 기간설정 Intent

사용자가 펀드의 수익률의 기준을 설정하기 위한 Intent이다. 수익률 조회 기간은 펀드를 추천 할 때 필요한 정보이기에, terms Slot을 통해 사용자가 말한 수익률 기간을 저장한다. 이 Slot은 펀드 다모아의 기준인 6개월 수익률과 1년 수익률 2가지로 이뤄져 있다. 모이라 신규 유저는 순차적 진행에 따라 펀드타입 Intent에서 이미 “수익률 기간 설정”이라는 다음 동작을 제안 받았기 때문에, terms Slot 과 “추천하다”라는 의미의 동사의 변형으로 이루어진 발화를 입력할 것이며, 모이라 서비스에 익숙한 유저들은 해당 동작에 대한 암시를 받지 않기 때문에, terms Slot, FundType Slot과 “추천” 또는 “검색”의 의미로 쓰이는 동사로 이루어진 발화를 입력할 것이라 예상된다. 그렇기 때문에 총 103개의 발화 예시를 작성하였다.

사용자가 입력한 발화를 서버에서는 어떻게 처리를 하는가? 앞서 언급하였듯이 펀드추천 프로세스는 펀드추천 – 펀드타입 설정- 수익률 조회기간 설정 3단계의 프로세스를 표준으로 동작되므로 사용자가 마지막 수익률 조회기간을 설정 하였을 때 앞서 선택한 펀드타입이 수익률 기간과 함께 함수로 넘겨져야 한다. 서버에서는 이러한 질문에 대한 응답을 보내는 동시에 이전 대화의 기록을 “SessionAttribute”에 저장하여 넘겨지도록 구축했다. 이러한 Mulit-Term 대화 시스템은 사용자의 잘못된 입력 혹은 모이라가 사용자의 질문 파악을 제대로 못한 경우 응답 실패를 반환하게 되고 이 때 이전의 대화 기록이 사라지는 문제가 발생할 수 있다. 서버에서는 이러한 문제를 제거하기 위해 응답 실패의 메시지에 이전 대화의 “SessionAttribute”를 복사해 서버에 전달하였다. 이러한 방법을 통해 사용자는 대화 응답 실패에도 다음 단계로 이어질 수 있다.

펀드추천 프로세스에 익숙한 사용자들은 3단계 프로세스 대신 최종 응답에 필요한 요소들을 모두 충족한 1개의 문장으로 펀드 추천을 받을 수도 있다. 이 경우 앞에 표준 프로세스와 다르게 SessionAttribute가 없기 때문에 서버에서 같은 규칙으로 응답을 반환할 수 없다. 따라서 받은 요청 json에 “SessionAttribute”가 존재 여부를 확인한 후 딕셔너리가 비어 있다면 펀드 타입, 수익률 기간을 slots에서 추출하여 함수로 전달하는 방식으로 구축하였다.

마지막으로 서버는 응답의 확실성을 위해 응답에 필요한 요소를 생략하고 최종 단계로 진입하는 것을 제한할 필요가 있다. 가령 본 펀드추천 대화 프로세스에서 펀드 타입을 설정하지 않은 상태로 “6개월 수익률” 과 같이 수익률 기간을 설정하는 경우 서버는 펀드타입을 먼저 설정하라는 메시지를 보내도록 설정 하였다.

사용자가 올바른 형태로 질문을 하였다면 서버는 json에서 추출한 펀드 유형과 조회수를 전달 인자로 자체 SQL 서버로 접근해 저장된 DB에서 해당 기간에 수익률이 높은 5개 펀드를 가져온다.

펀드 추천 응답에 필요한 데이터는 펀드다모아에서 크롤링 하였다. 펀드다모아 사이트의 경우 기본으로 조회한 펀드 종류들을 1년 수익률 기준으로 제공하고 있다. 하지만 펀드 추천 기능의 경우 사용자가 1년 수익률 외에도 6개월 수익률을 요구 할 수 있기 때문에 6개월 수익률에 대한 정보도 필요하다. 펀드다모아 페이지는 6개월 수익률 상위 5개를 가져오려면 이용자가 직접 ‘6개월 수익률’ 버튼을 눌러줘야 하기 때문에 selenium을 이용하여 브라우저를 제어해 6개월 수익률 및 1년 수익률 기준 각각 상위 5개 펀드종목들을 가져오도록 하였다.

데이터를 사용자에게 전달하는 방식에 있은 매 사용자의 요청마다 크롤링을 실행하는 것과 크롤링한 데이터를 자체 DB에 저장하여 불러오는 방법이 있다. 전자의 경우 저장을 하지 않고 사용자의 요청과 동시에 크롤링을 하기 때문에 SQL 최신의 데이터를 보장하는 장점이 있다. 하지만 셀레니움을 사용할 경우 다수의 사용자가 동시 다발적으로 크롤링을 요청하는 경우 응답 반환이 실패할 수 있기 때문에 본 팀은 펀드다모아에서 크롤링한 데이터를 자체 SQL에 저장하고 출력하는 방식을 사용하였다.

## 주식추천

주식 추천 기능은 사용자가 원하는 조건에 따른 주식 종목 5개를 추천하는 기능이다. 주식추천 Intent가 해당 기능을 구현한다.

-주식추천 Intent

사용자의 주식 추천 요청을 처리하는 Intent이다. 해당 발화는 사용자가 원하는 조건 없이 주식 추천 기능을 사용한 후 원하는 조건을 입력하는 방법과 최초의 질문에 사용자가 원하는 조건을 포함한 요청 두가지 종류가 있다. 우선, 모이라 이용이 익숙하지 않은 사용자가 물어보는 최초의 주식 추천 요청은 입력 받아야하는 정보가 없기 때문에, Slot 이 없는 상태로 발화가 만들어 진다. 이러한 초기 요청 발화들은 추천 받고 싶은 대상을 수식하는 형용사와 그 대상, 그리고 “추천”하다의 유의어로 구성된다. 초기 요청 발화를 입력 받은 후, 모이라는 설정 값을 요청한다. 여기서 지칭하는 설정 값은 총 4가지로 산업군, 설정 기준 변수, 제한 범위, 제한 수치이다. 사용자가 응답한 위 4가지 변수를 입력 받기 위해 Industry, Dic (개념), condition, number라는 4가지 Slot을 설계했다. Dic Slot은 기본적으로 금융 투자 개념 및 증권 관련 용어들이 입력되어 있다. 사용자는 3가지 혹은 4가지 Slot으로 구성된 발화를 입력한다. 만약 3 개의 Slot을 넣어 발화를 구성한다면 Dic, condition, 그리고 number로 이뤄진다. 4개의 Slot으로 구성된 Industry, Dic, condition, 그리고 number Slot들을 적절히 조합하여 변형한 문장들이다. 해당 Intent는 총 79개의 발화문이 현재 입력되어 있다.

서버에서는 사용자가 어떤 추천을 받고 싶은 기준을 제시하면 서버는 질문에 대한 slot value들을 변수화 하여 쿼리로 넘겨 준다. 쿼리문은 변수가 2개인 질문과 1개인 질문으로 나뉘는데 이는 산업군을 지정에 따른 변수 개수의 차이이다. 서버에는 산업군 변수가 지정되지 않을 경우 디폴트 값인 “”로 설정하여 하나의 함수에서 산업군이 지정된 경우와 미지정된 경우 모두를 처리할 수 있도록 추출하도록 설계하였다.

if industry == '’:

if feature == 'pbr' or feature == 'per' and condition =='<=':

cursor.execute("SELECT `name`,`returns` FROM `ppp`.`stockInfo` WHERE `%s` %s %s and `%s` > 0 ORDER BY `returns` DESC limit 5"%(feature,condition,number,feature))

else:

cursor.execute("SELECT `name`,`returns`,FROM `ppp`.`stockInfo` WHERE `%s` %s %s ORDER BY `returns` DESC limit 5"%(feature,condition,number))

else:

if feature == 'pbr' or feature == 'per' and condition =='<=':

cursor.execute("SELECT `name`,`returns` FROM `ppp`.`stockInfo` WHERE `%s` %s %s and `%s`>0 and `sector`= '%s' ORDER BY `returns` DESC limit 5"%(feature,condition,number,feature,industry))

else:

cursor.execute("SELECT `name`,`returns` FROM `ppp`.`stockInfo` WHERE `%s` %s %s and `sector`= '%s' ORDER BY `returns` DESC limit 5"%(feature,condition,number,industry))

또한 PBR과 PER은 일반적으로 값이 낮을수록 좋기 때문에 사용자가 조건의 제한선을 “이하”로 물어 볼 가능성이 높다. 이 경우 변수의 범위에 대한 하한선을 두지 않을 경우 음수의 값이 추출되는 오류가 발생할 수 있다. 따라서 사용자가 PBR 혹은 PER이 특정 값 이하인 종목을 추출하려 할 때 미리 하한선 0을 지정한 쿼리를 SQL 서버에 전송하여 오류를 방지한다.

cursor.execute("SELECT `name`,`returns` FROM `ppp`.`stockInfo` WHERE `%s` %s %s and `%s`>0 and `sector`= '%s' ORDER BY `returns` DESC limit 5"%(feature,condition,number,feature,industry))

주식 추천에 필요한 데이터들은 네이버 금융의 시가총액 페이지와 FinanceDataReader API를 통해 가져왔다. 네이버 금융 시가총액 페이지의 기본 선택변수는 거래량, 시가총액, PER, ROE, 외국인비율, 상장주식수 이렇게 6개이다. 하지만 모이라 주식 추천 기능에서 필요한 변수는 위 6개에서 거래량을 제외하고 PBR을 포함시켜야 했기 때문에 Selenium 을 통해 브라우저를 동적으로 제어하여 화면에 표시되는 데이터를 수정한 후 크롤링 하여 자체 DB에 저장 하였다.

크롤링한 데이터들 중 Float 타입으로 저장되는 변수들의 경우 python의 결측치인 NaN을 mysql로 전송하는 것이 불가하기 때문에 크롤링한 데이터 프레임에 대한 전처리가 필수적이다. 결측치를 채워야 하는 FLOAT 변수 PBR, PER, ROE, Returns는 변수의 특성에 따라 다르게 넣어 주었다. 먼저 낮을수록 좋은 값인 PBR,PER은 sys.maxsize 함수를 통하여 python 정수 변수의 최대 값을 넣었으며 높을수록 좋은 값인 ROE, Returns의 경우 -sys.maxsize를 넣었다.

## 미래에셋 추천펀드

미래에셋 상품 추천은 사용자에게 미래에셋대우가 선정하는 월간 추천상품을 제공하는 기능이다. 미래에셋은 매달 해외주식, 국내 주식, 그리고 대체투자 3가지 영역에서 정량과 정성평가를 종합하여 추천 상품을 선정하고 있다. 본 미래에셋 추천 상품 검색 기능은 미래추천 Intent와 자산군 Intent를 통해 작동한다.

-미래추천 Intent

사용자가 미래에셋 추천 상품을 검색하고 싶은 의도를 확인하기 위한 Intent이다. 본 의도는 확인해야 하는 정보가 없으므로 무 (無) Slot 형식의 발화로 구성하였다. 기본적으로 해당 Intent의 발화는 “미래에셋”, “추천 상품”, 그리고 “추천”의 의미를 가진 동사로 구성되고 총 30개의 발화 예시가 있다. 미래에셋의 추천 상품은 매달 약 10개이기에 모든 결과를 음성 응답으로 반환하기에는 지나치게 길다. 양방향 음성 인식 인터페이스의 장점인 즉각성과 효율성을 살리기 위해 최초 요청 발화에 자산군 3 가지 선택 사항을 주며 다음 행동을 유도한다.

-자산군 Intent

사용자가 희망하는 자산군을 입력받아 해당 자산군의 추천 상품을 반환하는 Intent이다. 이 Intent는 자산군이라는 정보를 획득해야하기 때문에, AssetGroup이라는 Slot을 생성하였다. 본 Slot은 국내주식, 해외주식, 그리고 대체투자등 3가지 값이 있다. 앞서 미래추천 Intent에서 다음 동작 제안을 받거나, 혹은 특정 자산군의 추천 상품 검색을 위한 최초 발화 모두 AssetGroup Slot만을 포함하기 때문에 발화는 AssetGroup Slot과 “검색” 동사의 조합으로 총 30개가 있다.

서버에서는 사용자가 입력한 자산군에 알맞은 펀드 이름을 반환해주도록 구축되어 있다. 미래에셋 추천 상품 데이터의 경우 추천 상품이 1달 주기로 업데이트 되며 데이터의 개수가 적기 때문에 SQL 서버가 아닌 excel 파일로 따로 구축해 놓았다.

## 용어사전

사용자가 궁금하거나 알고 싶은 금융용어를 설명해주는 기능이다. 사용자가 특정 금융용어를 검색하면, 음성으로 해당 정의를 쉬운 말로 풀어서 응답해주는 서비스를 제공한다. 용어 검색이 주 목표이기 때문에 Single-Turn 대화 시나리오만 존재한다. 용어 설명 Intent가 본 기능을 구현한다.

-용어설명 Intent

사용자가 금융용어를 검색을 요청을 처리하는 Intent이다. 사용자가 검색하고 싶은 금융용어를 입력 받으면, 그에 상응하는 정의를 알기 쉽게 반환해준다. 사용자가 모이라에게 물어보고 싶은 용어 Intent가 동작하기 때문에 Dictionary 라는 Slot을 설계하여 금융 기초 용어 102개의 저장했다. 용어와 정의 출처로는 전자 공시 서비스 용어사전, 금융소비자 정보 포탈 금융용어사전, 매경과 한경 용어사전을 참고하였다. Dictionary Slot과 “검색” 과 유의한 동사를 조합하여 총 46개의 발화목록이 있다.

서버는 사용자가 질문한 금융용어에 매칭되는 설명을 “Terms.xlsx” 에서 찾아 사용자에게 알려준다.

## 주가확인

사용자가 특정 주식의 가격을 문의하였을 때, 해당 주가를 검색하여 알려주는 기능이다. 사용자가 손쉽게 주식 종목만 언급하면, 주가를 검색할 수 있기에, 조회하고 싶은 종목 이름은 모든 발화에 꼭 필요한 요소이다. 따라서 Single-Turn 대화 시나리오만 존재한다.

-주가확인 Intent

사용자가 주식 가격 검색을 요청하는 Intent이다. 주식 종목 이름이라는 정보를 입력받아야하기 때문에 stockname이라는 Slot을 만들었다. 본 Slot에는 2018년 7월 기준으로 국내 시장에 상장된 2203개의 종목코드가 있다. stockname Slot, “조회”와 유의한 동사의 조합으로 98개의 발화 예시를 등록되어있다.

사용자가 특정 종목의 현재 가격에 대한 요청이 들어오면 서버는 네이버 금융에서 slot에 있는 종목 코드에 해당하는 사이트로 접근해 현재 가격을 가져온다. 또한 clova 앱 등을 사용하여 CEK에 접근할 시 1개의 계정에 1개의 clova userID가 저장 되는 것을 이용하여 사용자가 주가 확인을 할 때마다 서버는 SQL에 사용자의 UserID 테이블에 검색한 종목의 코드를 저장한다. 이는 추후에 추천 기능을 추가하기 위한 데이터를 구축할 수 있을 것으로 예상한다.