



고령층을 위한 음성 기반 주문 서비스

신홍재 김원우 김성룡 한동훈

늘겜에

00 팀 소개



신 홍 재

팀장

RAG 아키텍처 설계
RAG LLM구현



김 원 우

팀원

NLU, FSM 설계
앱 설계



김 성 룡

팀원

추천 가게 선정
프로그램 설계
RAG LLM설계



한 동 훈

팀원

STT, NLU 학습
TTS 구현

목차

01 주제 선정 배경

02 프로젝트 개요

03 사용 모델 및 기술

04 시스템 구현

05 기대효과 및 활용방안

고령층을 위한 음성 기반 주문 서비스

주제 선정 배경

01

"로그인부터 막막"...노년층은 '비대면 스트레스'

더 벌어진 '디지털 격차'...더 커진 '노인 소외감'

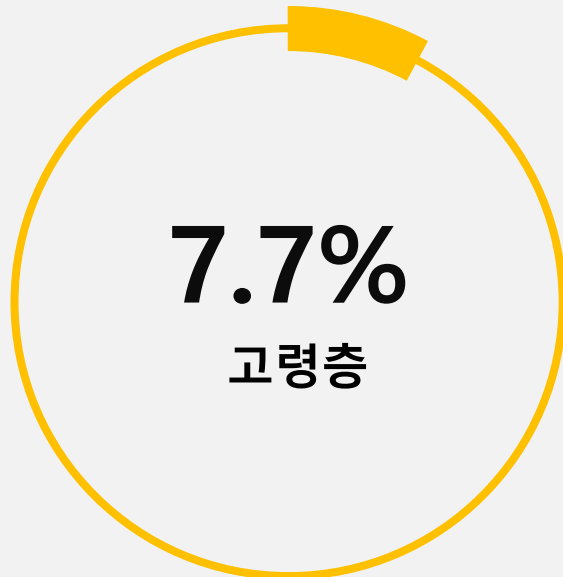
서울 고령층 10명 중 7명 "배달앱 안 써봤다"... 디지털장벽 '노인세' 여전

"고령층 '디지털 격차' 해소, 사회적 공감대 필요"

똑똑하지만 불친절한 디지털...노인들
"써보기도 전에 포기"

01 주제선정 배경

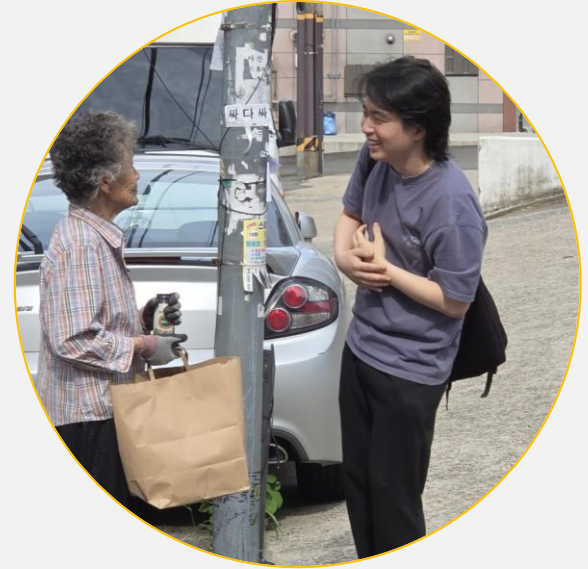
서울시 70세 이상 고령층의
음식 배달 앱 이용 경험률



낮은 이용률(7.7%)의 이유를 알아보기 위해,
가야지역 경로당 5곳에서 어르신들께 직접 여쭙었습니다.

“배달앱을 사용하지 않으시는 가장 큰 이유는 무엇인가요?”

01 주제선정 배경



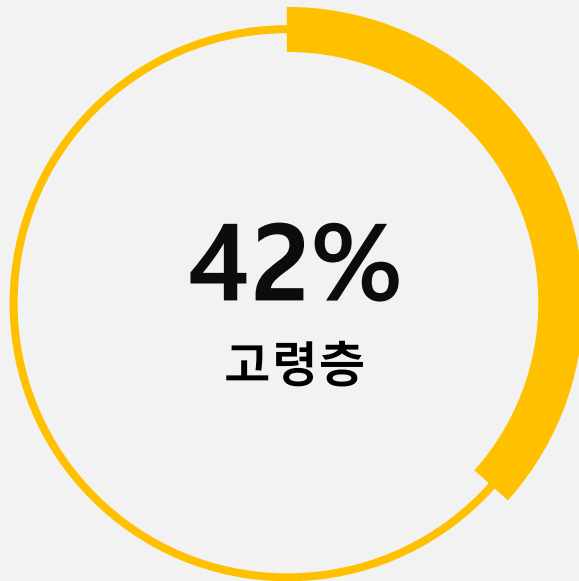
“ 전화주문이 익숙하기도하고,
배달 앱이란게 사용하기 어려워요...

”

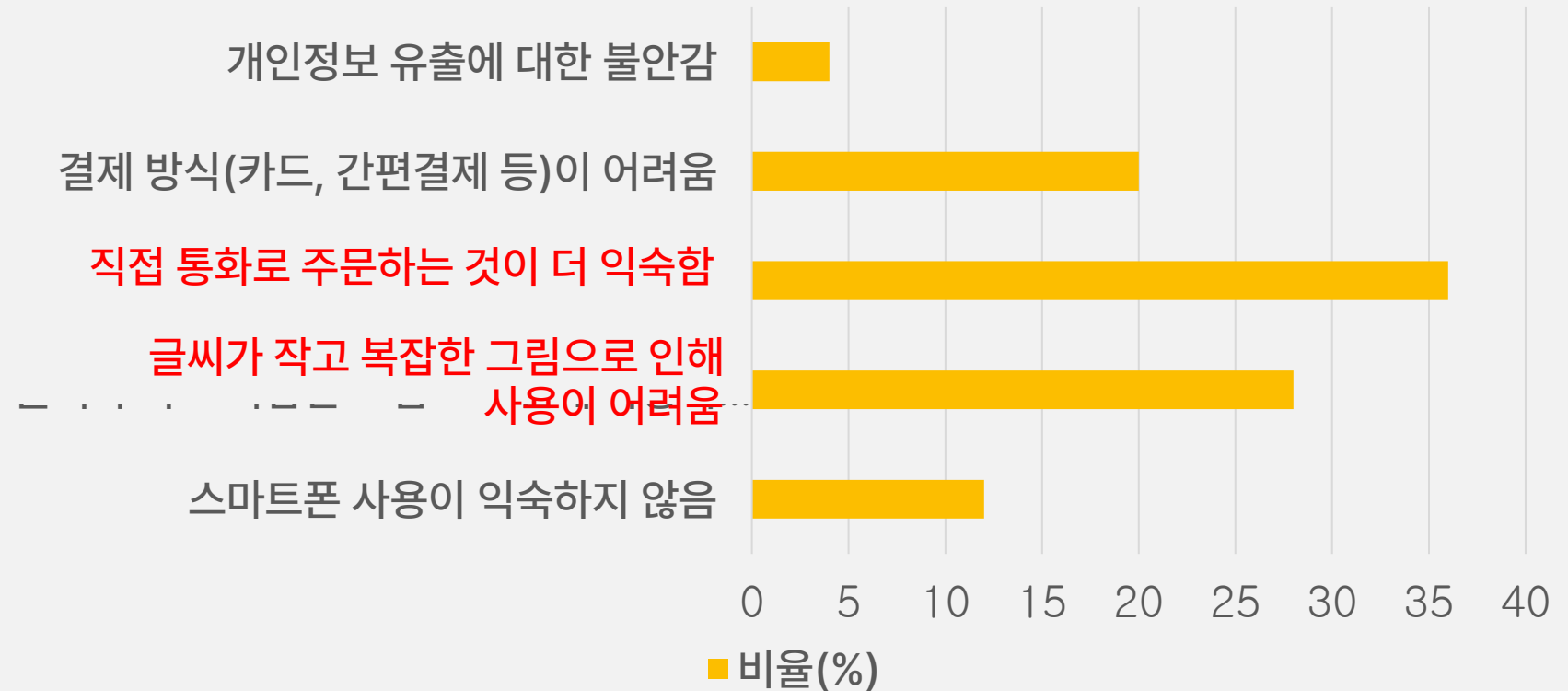
“ 자녀들에게 매번 물어보는 것도
눈치가 보이고 민망해 자꾸 시도조차 안 하게 돼요... ”

01 주제선정 배경

고령층의
음식 배달 앱 사용 의향



배달앱 미사용 이유



01 주제선정 배경

기존 앱은 글자·버튼 중심이라
노인에게는 **진입장벽**이 높습니다.

그래서 생각했습니다.

"그냥 전화를 하듯 **말로 주문**할 수 있으면 얼마나 좋을까?"

저희는 노인분들이 부담없이 주문할 수 있도록

자연스러운 말하기만으로 주문 가능한 배달앱을 구현했습니다.

고령층을 위한 음성 기반 주문 서비스

프로젝트 개요

02

02 프로젝트 개요



최신 & 정확한 정보를 위해 RAG + LLM 적용

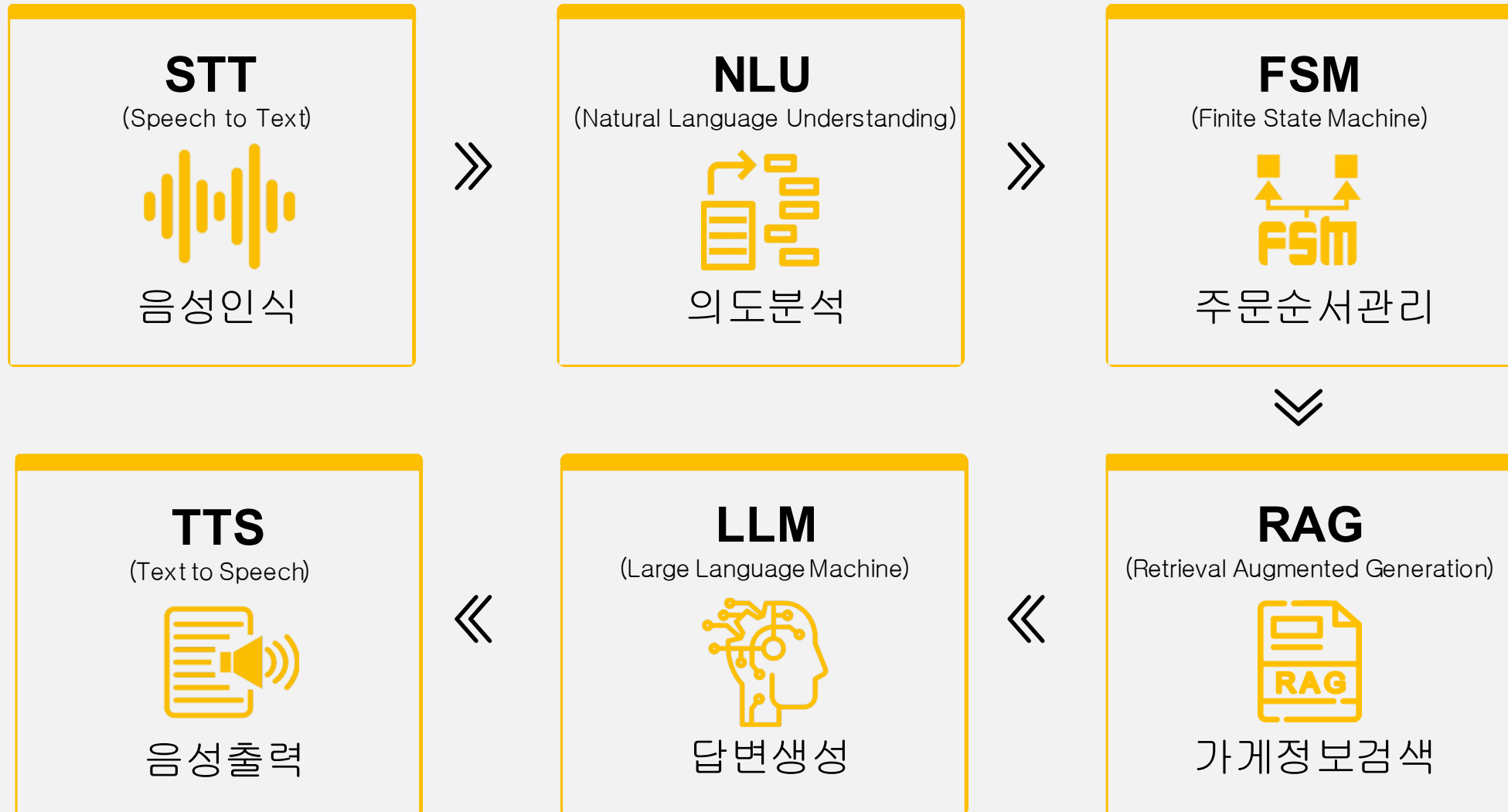


02 프로젝트 개요

주문 과정에 오류가 없도록
NLU → FSM → RAG / LLM 구조 적용



02 프로젝트 개요



고령층을 위한 음성 기반 주문 서비스

사용 모델 및 기술

03

03 사용 모델 및 기술

STT - 데이터



강원도
경상도

중, 노년층 한국어 방언 데이터
오디오, 나이, 성별,
지역, 문장 텍스트



충청도
전라도
제주도

중, 노년층 한국어 방언 데이터
오디오, 나이, 성별,
지역, 문장 텍스트

데이터 균등화 전

총 시간 : 약 2758시간

남성 : 26.8%

여성 : 73.2%

경상도 : 29.9%

전라도 : 24.6%

강원도 : 20.0%

충청도 : 20.2%

제주도 : 5.2%



데이터 균등화 후

총 시간 : 약 1167시간

남성 : 52.4%

여성 : 47.6%

경상도 : 24.3%

전라도 : 23.8%

강원도 : 23.3%

충청도 : 23.1%

제주도 : 5.5%

03 사용 모델 및 기술

STT - 모델 테스트

표준어+방언 음성 테스트 결과

* WER: 단어 오류율, CER: 글자 오류율

STT 모델	평균 WER	평균 CER
Whisper-medium	50.06 %	25.17%
XLS-R-300M-Ko	83.18%	43.99 %
Google STT	68.29 %	39.94 %

03 사용 모델 및 기술

STT - 학습

LoRA 파인튜닝



ONNX 최적화



PTQ 정적 양자화



LoRA

기존 모델의 가중치는 고정된 채
특정 층에 저차원 행렬을 추가하여
소수의 파라미터만 학습하는 기법

풀 파인튜닝 vs LoRA

사용 메모리 : 12.32GB → 3.2GB (74%↓)
학습 시간 : 324시간 → 54시간 (83%↓)
예상 성능 : 100% → 99~98% (1~2%↓)

03 사용 모델 및 기술

STT - 최적화

LoRA 파인튜닝



ONNX 최적화



PTQ 정적 양자화



ONNX 변환

연산을 그래프로 표준화하여
다양한 프레임워크에서 공통으로
사용할 수 있는 포맷으로 변환

ONNX Runtime

중복 연산 및 노드 제거, 연산 병합 등
그래프 최적화 기법을 적용하여
모델의 연산 효율 및 추론 속도 향상

03 사용 모델 및 기술

STT - 양자화

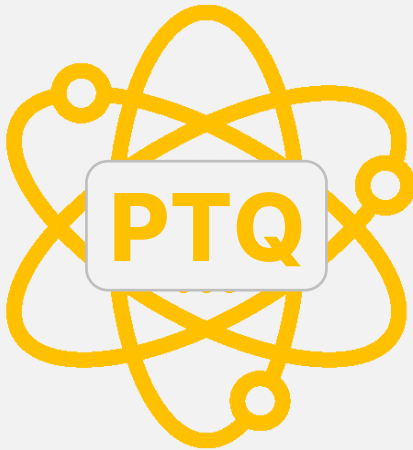
LoRA 파인튜닝



ONNX 최적화



PTQ 정적 양자화



양자화

모델의 가중치와 활성화 값을
실수형에서 정수형으로 변환하여
모델의 크기와 속도를 개선하는 방법

PTQ 정적 양자화

실제 서비스 환경과 유사하게 구성된
데이터를 적용하여 양자화 과정 중
성능 저하를 최소화하는 양자화 기법

03 사용 모델 및 기술

STT - 학습 결과

표준어+방언 음성 평가 결과

* WER: 단어 오류율, CER: 글자 오류율

STT 모델	평균 추론 속도	평균 WER	평균 CER
Whisper-medium 파인튜닝	0.79 초	18.74%	6.92%
Whisper-medium	1.21 초	50.06 %	25.17%
XLS-R-300M-Ko	0.08 초	83.18%	43.99 %
Google STT	1.83 초	68.29 %	39.94 %

03 사용 모델 및 기술

NLU (Natural Language Understanding)



NLU란?

사람이 사용하는 언어의 의도와 의미를
기계가 분석하여 이해하는 기술

NLU의 활용

사용자의 질의에서 의도를 정확히 파악
주문 과정에서의 오류를 최소화

03 사용 모델 및 기술

NLU - 데이터

문장 생성

음식 주문 상황

짜장면 1개랑 짬뽕 2개 주세요
짬뽕이 맛있는 중국집 추천해줘
수라반점에서 주문하고 싶어
배달 시간 얼마나 걸리나요
...
237개

문장 증강

맞춤법 오류, 사투리

짜증면 한게랑 짬뽕 두게 주새오
짜 장 면 한개랑 짬 뽕두개주세요
짱개 둘 짬뽕 하나 주이소
짜장 두개 짬뽕 하나 주드래요
...
21875개

의도 매핑

문장	의도
"짬뽕이맛있는중국집"	추천
"수라반점에주문할게"	가게선택
"짜장면1개랑짬뽕2개"	메뉴선택
"이렇게주문할게"	주문확정
"응", "그렇게해줘"	긍정응답
"아니", "다시주문할게"	부정응답
"가게리뷰요약해줘"	가게문의
"오늘날씨좋네"	잡담

03 사용 모델 및 기술

NLU - 학습 결과

NLU 모델 학습 결과

NLU 모델	정확도	초당 처리 개수
KLUE -RoBERTa-large	96.3%	183개
KoBigBird -RoBERTa-large	96.0%	32개

03 사용 모델 및 기술

NLU - 적용

의도 종류

문장	의도
"짬뽕이맛있는중국집"	추천
"수라반점에주문할게"	가게선택
"짜장면1개랑짬뽕2개"	메뉴선택
"이렇게주문할게"	주문확정
"응", "그렇게해줘"	긍정응답
"아니", "다시주문할게"	부정응답
"가게리뷰요약해줘"	가게문의
"오늘날씨좋네"	잡담

추천

가게 선택 중00:39

가게: -

주문: -

어떤 음식을 드시고 싶으세요?

맛있는 짬뽕집 추천해 주이소

제가 추천해 드리는
"짬뽕관-서면점"로 하실까요?
아니면 후보 5곳 중 골라 보세요

주문 확정

이대로 주문할게요

'명품짬뽕 2그릇' 20000원입니다
주문이 맞으실까요?

03 사용 모델 및 기술

FSM (Finite State Machine)



FSM이란?

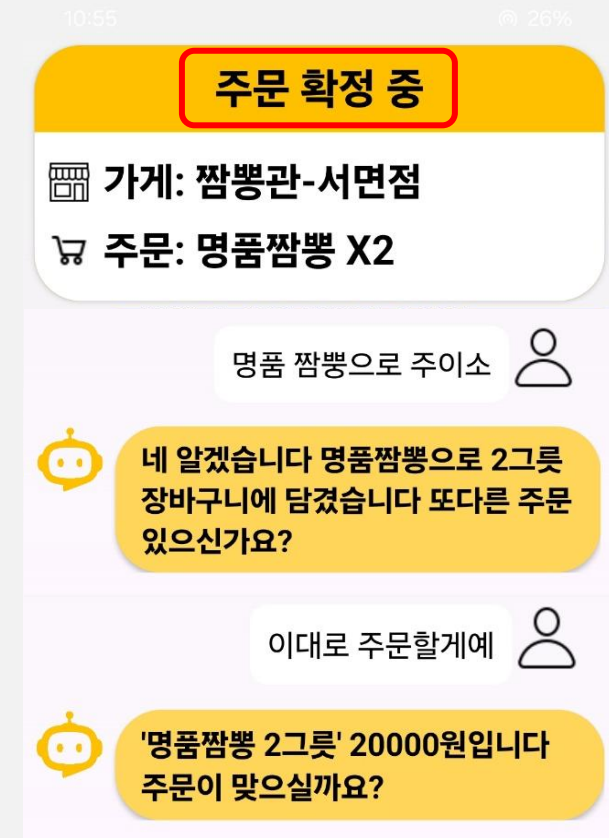
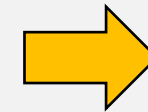
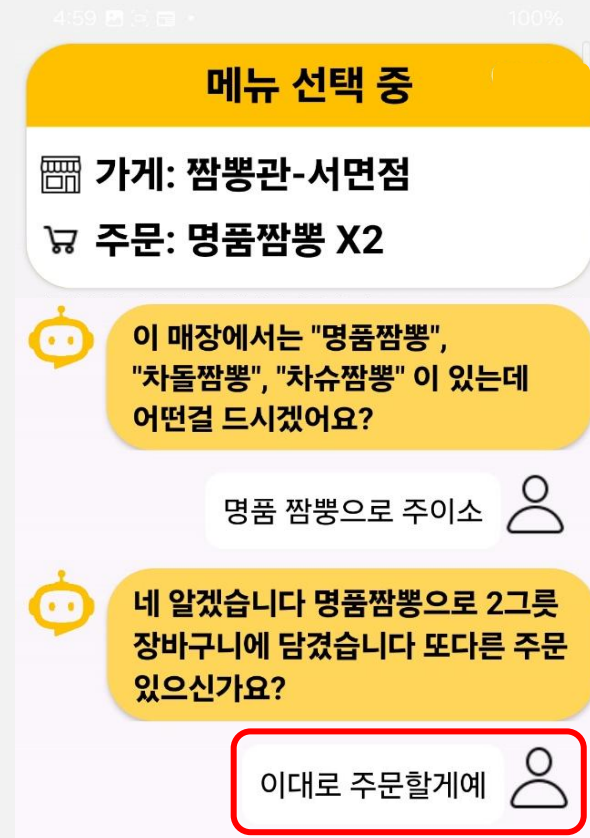
한 번에 하나의 상태만 가질 수 있고,
특정한 조건이나 이벤트가 발생하면
한 상태에서 다른 상태로 변화하는 시스템

FSM의 활용

사용자 입력의 의도에 따라
주문 상태를 다음 단계로
자연스럽게 전환되도록 제어

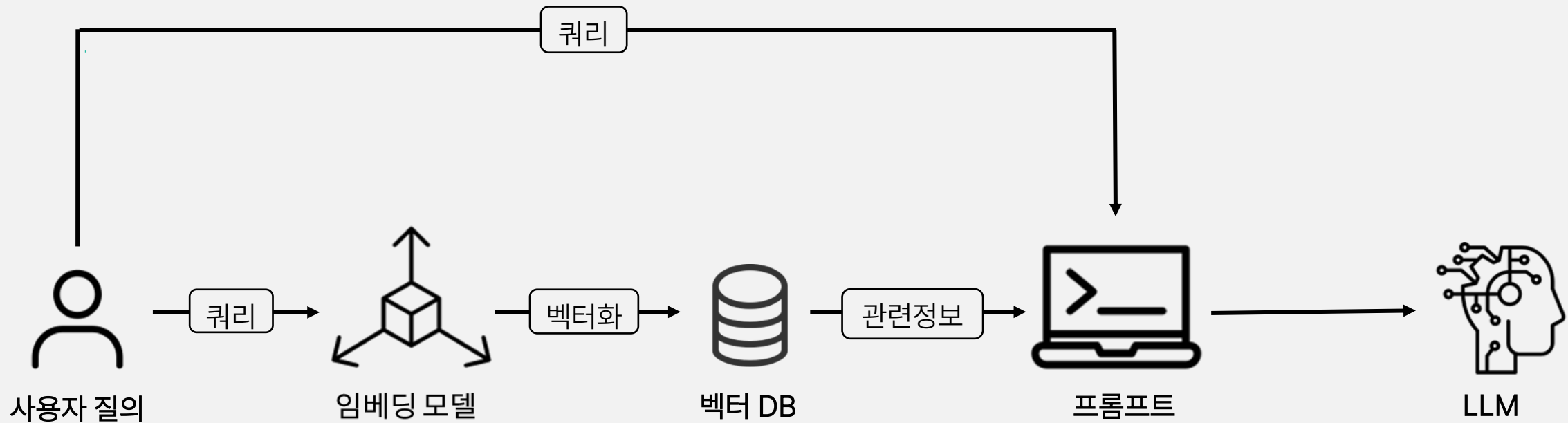
03 사용 모델 및 기술

FSM - 적용



03 사용 모델 및 기술

RAG (Retrieval Augmented Generation) / LLM (Large Language Model)



03 사용 모델 및 기술

RAG / LLM - 데이터

데이터 크롤링



데이터 청킹 & 저장



벡터 DB & 프롬프트



가게 - 1358개

가게명 · 주소 · 전화번호 · 영업시간 · 최소주문금액

메뉴 - 63431개

메뉴명 · 가격

리뷰 - 67823개

별점 · 리뷰 내용

03 사용 모델 및 기술

RAG / LLM - 전처리

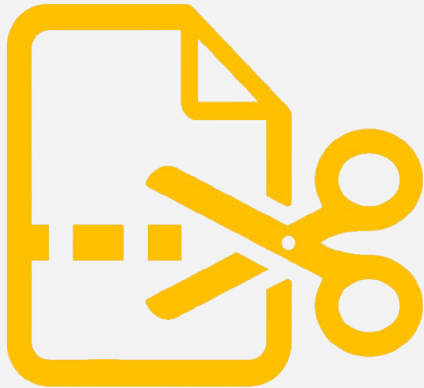
데이터 크롤링



데이터 청킹 & 저장



벡터 DB & 프롬프트



데이터 청킹

가게 · 메뉴 · 리뷰

청크 범위 각각 설정



임베딩 & 검색 효율 향상



데이터 저장

가게 · 메뉴 · 리뷰

각각의 벡터 DB에 저장



검색 안정성 향상

03 사용 모델 및 기술

RAG / LLM - 설계

데이터 크롤링



데이터 청킹 & 저장



벡터 DB & 프롬프트



벡터 DB

의도 · 상태

벡터 DB 조합 설계



검색 안정성 & 효율 향상



프롬프트

의도 · 상태

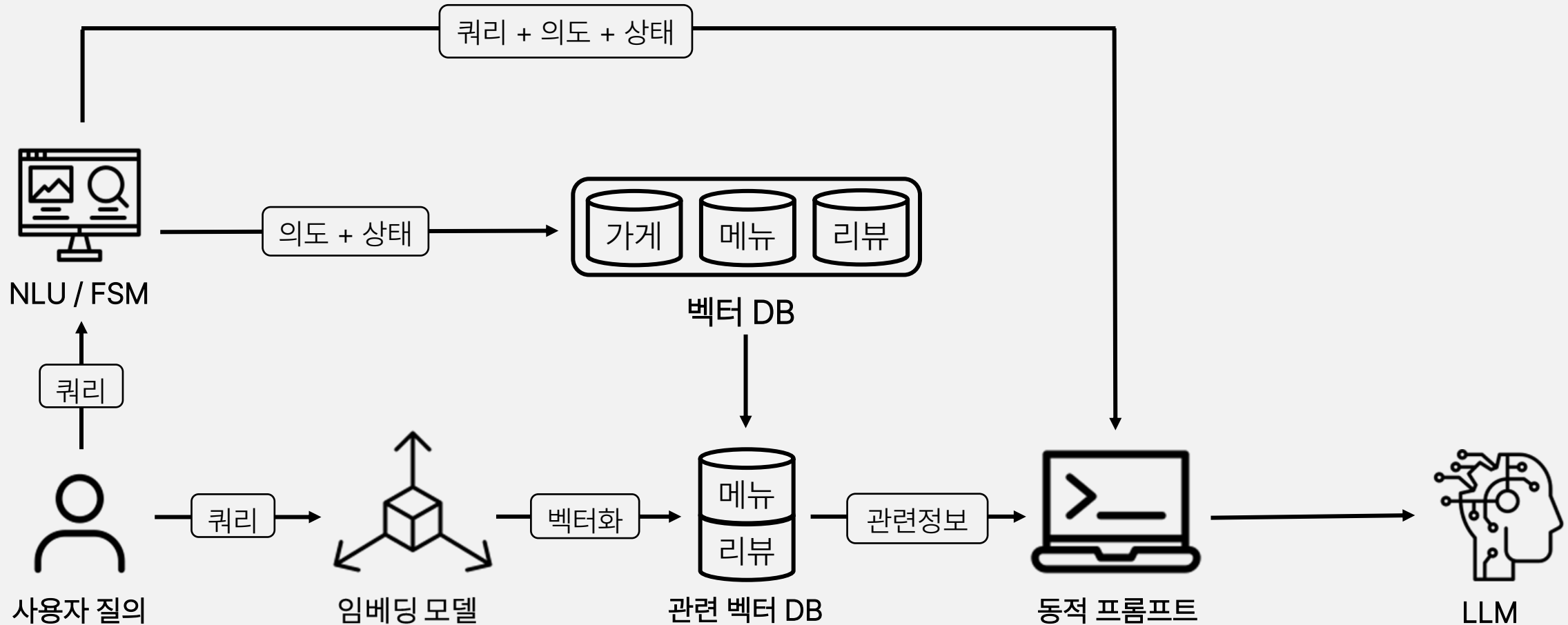
상황별 프롬프트 설정



최적 답변 생성

03 사용 모델 및 기술

RAG / LLM - 아키텍처



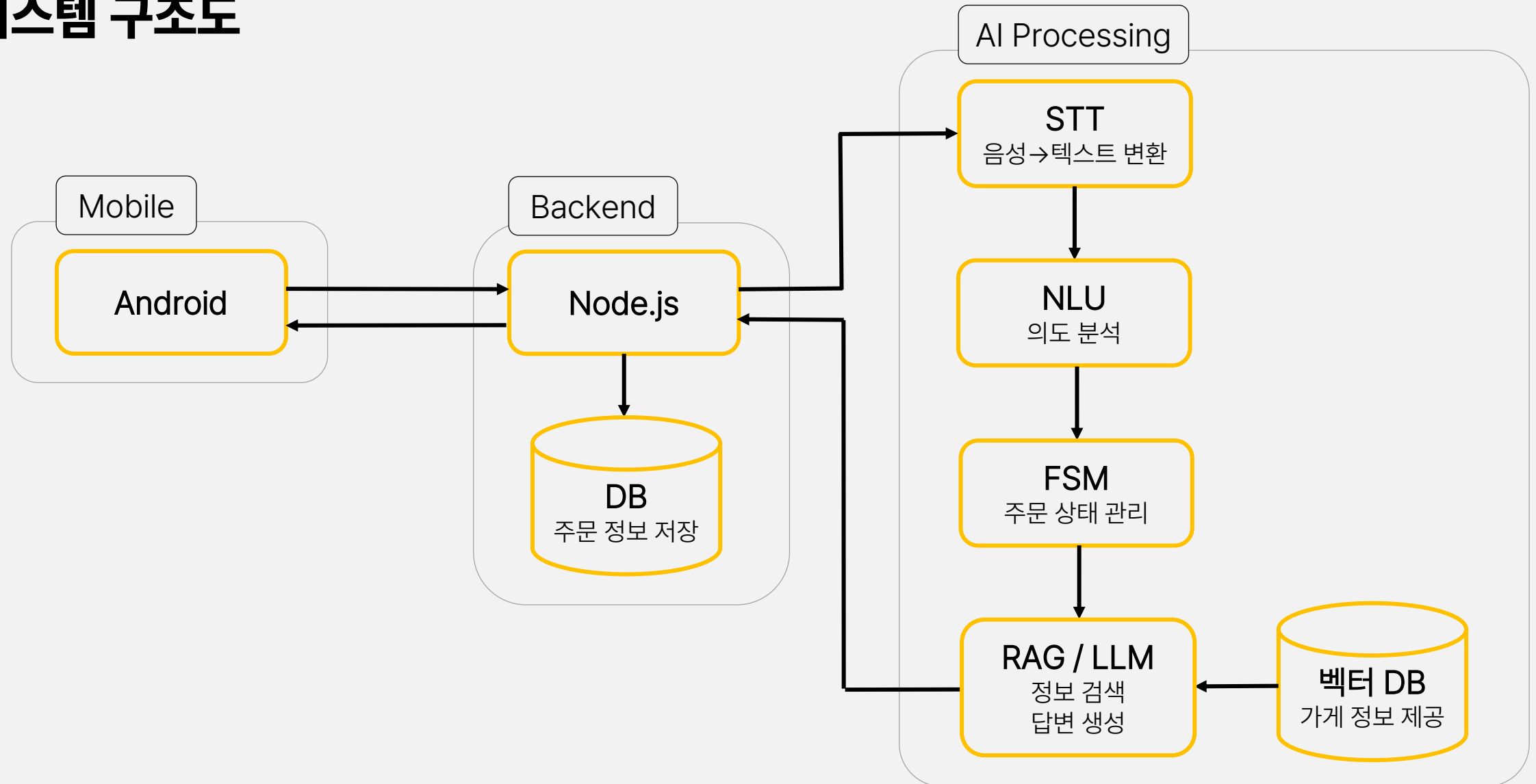
고령층을 위한 음성 기반 주문 서비스

시스템 구현

04

04 시스템 구현

시스템 구조도



고령층을 위한 음성 기반 주문 서비스

기대효과 및 활용방안

05

05 기대효과 및 활용방안

활용방안



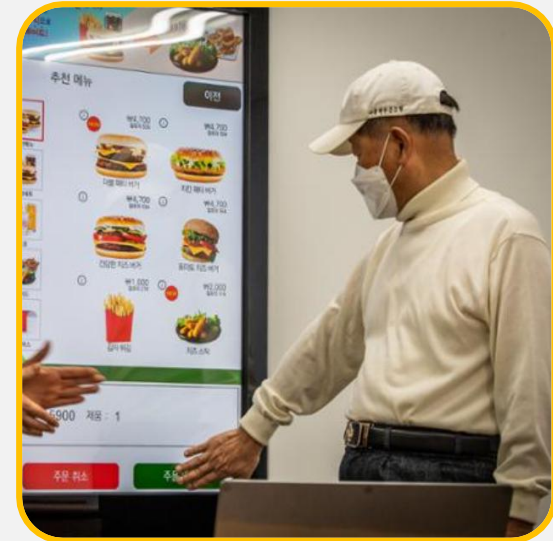
스마트홈 · IoT 음성 비서 연동

배달 앱을 직접 실행하지 않고,
자연스러운 대화로 자동 주문



시각장애인 보조 플랫폼

음성만으로 주문할 수 있기 때문에,
시각장애인 보조용으로 활용 가능



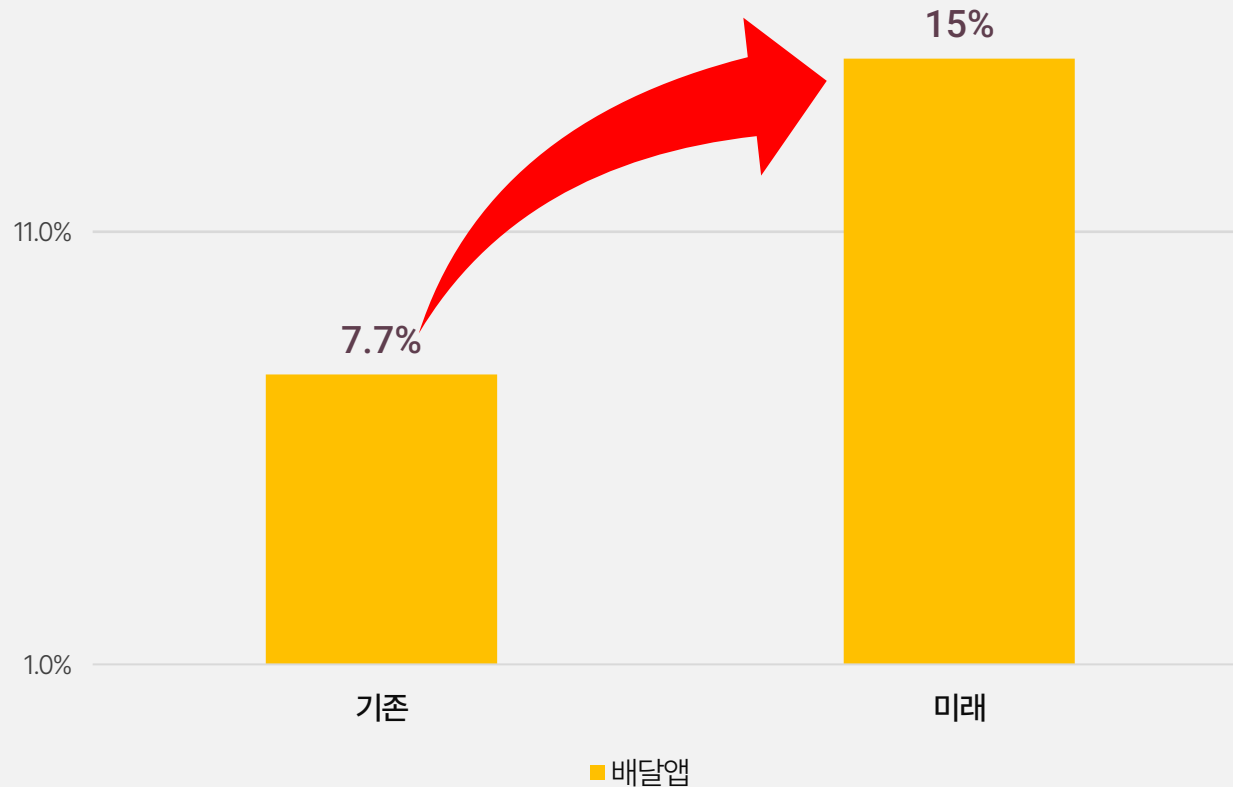
오프라인 매장 활용

키오스크, 주문 태블릿 등
오프라인 매장에 활용 가능

05 기대효과 및 활용방안

기대효과

70세 이상 고령층의
음식 배달 앱 이용 경험률



주문 처리 시간 단축

음성만으로 주문이 가능해지면서
별도 메뉴 탐색, 입력 과정 생략



QnA