2020 년 – 1 학기

졸업프로젝트 1

요구사항 분석서

[개인 프라이버시를 보호하는 협업 학습을 활용한 스마트폰 사용 패턴 분석 및 스트레스 예측]

6 조

201711356 천세진

201612066 김지효

지도교수: 박소영

제출일: 2020 년 4 월 27 일

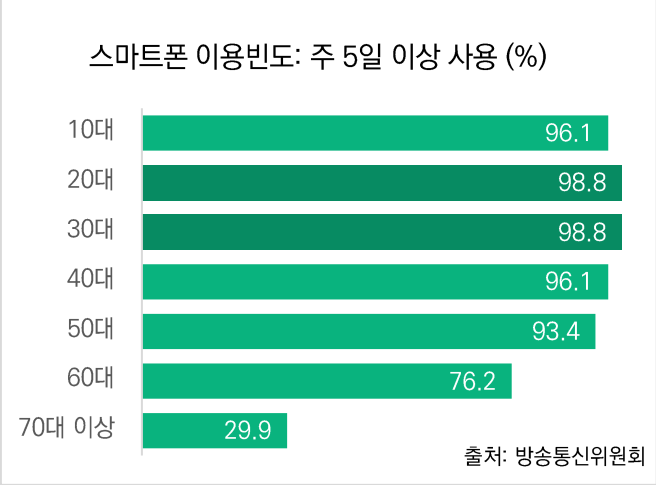
내용

1. 개요

1.1 기획배경

[스마트폰 사용의 증가]

현대인들에게 스마트폰은 일상생활에서 없어서는 안될 필수품으로 자리잡았다. 스마트폰 보유율은 2015 년에 78.8%에서 2019 년에는 91.1%로 꾸준히 증가해 오고 있으며, 63%가 스마트폰이 일상생활에서 필수적인 매체라고 답했다. 주 5 일 이상 스마트폰을 사용하는 인구는 20 대, 30 대에서는 98.8%를 차지할 정도로 현대인의 삶에서 빼놓을 수 없는 필수요소로 자리잡게 되었다.[1]



〈그림 1〉 연령별 스마트폰 이용빈도

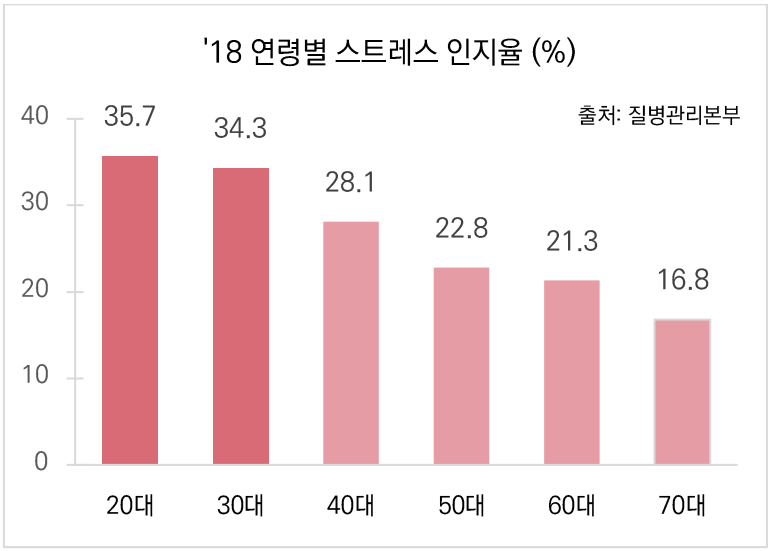
스마트폰에서는 다양한 앱을 기반으로 무궁무진한 기능을 수행할 수 있다. 통신 매체로서 기본적인 기능인 전화, 메시지는 물론이고 동영상 시청, SNS 등 사람마다 스마트폰을 사용하는 이유는 다양하다. 그래서 각 사용자의 스마트폰 사용 기록은 개인의 취향과 특성을 고스란히 드러내기 때문에 스마트폰 사용 패턴을 수집한 자료를 기반으로 각 사용자의 행동 특성 및 심리 상태를 파악할 수 있을 것이라고 생각했다.

[현대인의 스트레스]

[1] 방송통신위원회, 2019 방송매체 이용행태 조사 (2019.12)

최근 현대인의 스트레스가 큰 사회 문제로 대두되고 있다. 스트레스는 긍정적이거나 부정적인 요인에 의해 모두 야기될 수 있으며, 스트레스에 노출되면 아드레날린 분비로 인한 교감신경계 활성화 등 여러가지 신체적 반응이 나타난다. 적절한 양의 스트레스는 개인의 업무 수행 능력 향상 등의 긍정적인 효과를 주기도 하지만, 극심한 스트레스가 지속될 경우 긴장성 두통, 심혈관 질환, 전신통증, 우울증, 심할경우 암 등의 신체적, 정신적 질병을 유발한다.[2]

만 19 세 이상 스트레스인지율은 2008 년 29.2%에서 2018 년 29.1%로 큰 변화가 있지는 않았으나, 2018 년 기준 성인 10 명중 3 명이 스트레스를 '대단히 많이' 또는 '많이' 느끼는 것으로 나타났고, 그림 2 와 같이 20~30 대가 다른 연령 계층보다 스트레스인지율이 높았다.[3] 이렇게 스트레스 인지는 젊은 층에서 더 많이 나타났다.



〈그림 2〉 2018 년 연령별 스트레스 인지율

우리는 사용자의 심리 상태 및 행동 특성이 스마트폰 사용 패턴에 반영되기 때문에 신경망 학습을 통해 스트레스 수준에 따른 스마트폰 사용 패턴을 도출해, 연관성을 설명해낼 수 있을 것이라고 생각한다. 스마트폰 사용 패턴은 연령 별로 아주 다양한 양상을 보일 것으로 예상되므로 우리는 연구대상을 대학생으로 한정해 진행할 예정이다.

---

[2] 이충재, 현대인의 스트레스, 어떻게 극복할 것인가(2018.06)

[3] 질병관리본부, 주간 건강과 질병(2018)

1.2 기술 동향

[완전 동형 암호 (Fully Homomorphic Encryption)]

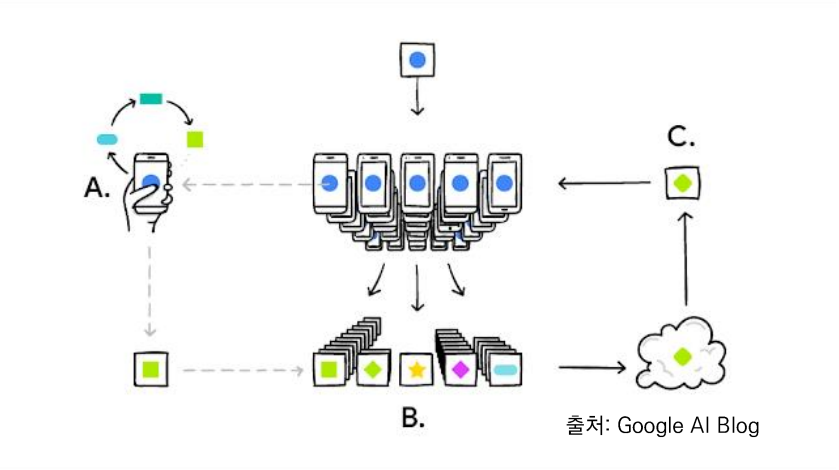
‘동형 암호’란, 평문에 대한 연산을 수행한 후 암호화한 결과(암호문)와 각각의 암호문에 대하여 연산을 수행한 결과가 같은 값을 가지는 암호화 방식이다. 동형 암호화 기법은 일정한 횟수 이상 연산을 수행하면 노이즈(noise) 값이 발생하여 더 연산을 수행할 수 없어, 연산 횟수에 제한이 있는 암호화 방식이다. 이를 해결하기 위하여 암호문의 노이즈를 감소시키는 재부팅(Bootstrapping) 과정을 통해 연산 횟수 제한을 없앨 수 있는 암호화 방식을 완전 동형 암호(Fully Homomorphic Encryption)이라고 한다.

동형 암호는 1978 년에 처음으로 발표되었다. 그러나 초기의 동형 암호는 이론적으로 안전성이 증명되지 않았다. 2009 년 Gentry 가 안전성이 증명된 동형 암호를 발표한 이후, 2011 년 MIT 가 동형 암호 기술을 10 대 Emerging Technology 로 선정하면서 학계 및 산업계의 관심을 끌게되었다.

동형 암호는 암호문을 복호화 하지 않아도 검색, 통계 처리 및 기계학습이 가능하고, 데이터 처리 시 중간 과정에서 복호화 하지 않아도 되므로, 데이터 유출 위험이 감소하는 장점을 가진다. 현재 동형 암호는 생체 인식 분야, 금융 분야 등 여러 분야에서 이용되고 있다.

[연합 학습 (Federated Learning)]

인공지능의 기술 수준이 급속도로 향상되면서 이를 기반 기술로 한 활용 범위가 더욱 확산되고 있다. 기존 인공지능 기술은 모든 데이터를 중앙 서버로 모아 모델을 학습시켜 인공지능을 구현한다. 그러나 개인의 데이터를 타인에게 넘겨주는 것이라는 특성 상 각 사용자의 데이터에 담긴 사생활 침해의 문제가 꾸준히 제기되고 있다. 이를 해결하기 위해, 연합 학습 방식에서는 사용자의 데이터를 그들의 스마트폰에서 직접 처리하고, 신경망 모델을 갱신하는 weight 값만 서버로 전달해 성능을 높인다. 〈그림 3〉과 같이 A 의 스마트폰 사용에 따라 모델을 기기 안에서 개인화한다. 이후 많은 사용자의 업데이트가 B 에 취합된다. 이로 인해 개선된 신경망 모델을 전송해 C 에서 다시 사용자 기기의 모델을 갱신한다.



〈그림 3〉 연합학습 과정

연합 학습을 이용하여 얻을 수 있는 이점은 크게 세 가지가 있다. 첫째로, 연합 학습은 사용자의 데이터가 아닌 학습된 모델을 수집하기 때문에 사생활 침해의 소지가 적다. 두번째로, 중앙 컴퓨팅 파워 부하의 비중이 감소한다. 기존에는 중앙에서 수많은 데이터를 학습해야 했지만, 이 역할을 사용자 기기에 분할할 수 있기 때문에 부하량을 감소시킬 수 있다. 마지막으로 표본 데이터의 정확성이 올라가는 이점이 있다. 기존에는 데이터를 수집해야 했는데, 반감이 있는 사용자로부터 데이터를 모으지 못하는 문제가 있었다. 그러나 연합 학습은 데이터가 아닌 학습 모델을 수집하므로 기존보다 반감이 적고, 표본의 편향성도 줄어든다.

1.3 프로젝트 주요 기능 및 특징

우리는 사용자의 개인 프라이버시를 보호하기 위해 여러 기술들을 적용해서 스마트폰 사용 패턴 수집 및 분석을 하고, 이를 통해 그들의 스트레스 수준을 예측하고자 한다.

i. 연합 학습을 통한 신경망 구축

- 구축된 신경망 모델을 각 사용자의 디바이스로 전송한 후 수집된 데이터로 모델을 갱신한다. 이를 통하여 사용자의 개인 프라이버시를 보호하고, 중앙 서버의 부담을 줄인다.

ii. 동형암호 기법을 사용한 데이터 수집

- 사용자의 스마트폰 사용기록은 그들의 민감한 개인정보이므로 데이터 수집 시 암호화를 통해 privacy 를 보존할 수 있어야 한다. 이를 해결하기 위하여 완전 동형 암호 기법을 이용할 것이다.

iii. 스마트폰 사용 패턴으로 스트레스 수준 예측

- 구축된 모델을 기반으로 사용자의 스마트폰 사용 패턴을 수집해 스트레스 수준을 예측한다.

iv. 사용자에게 스트레스 수준 알림

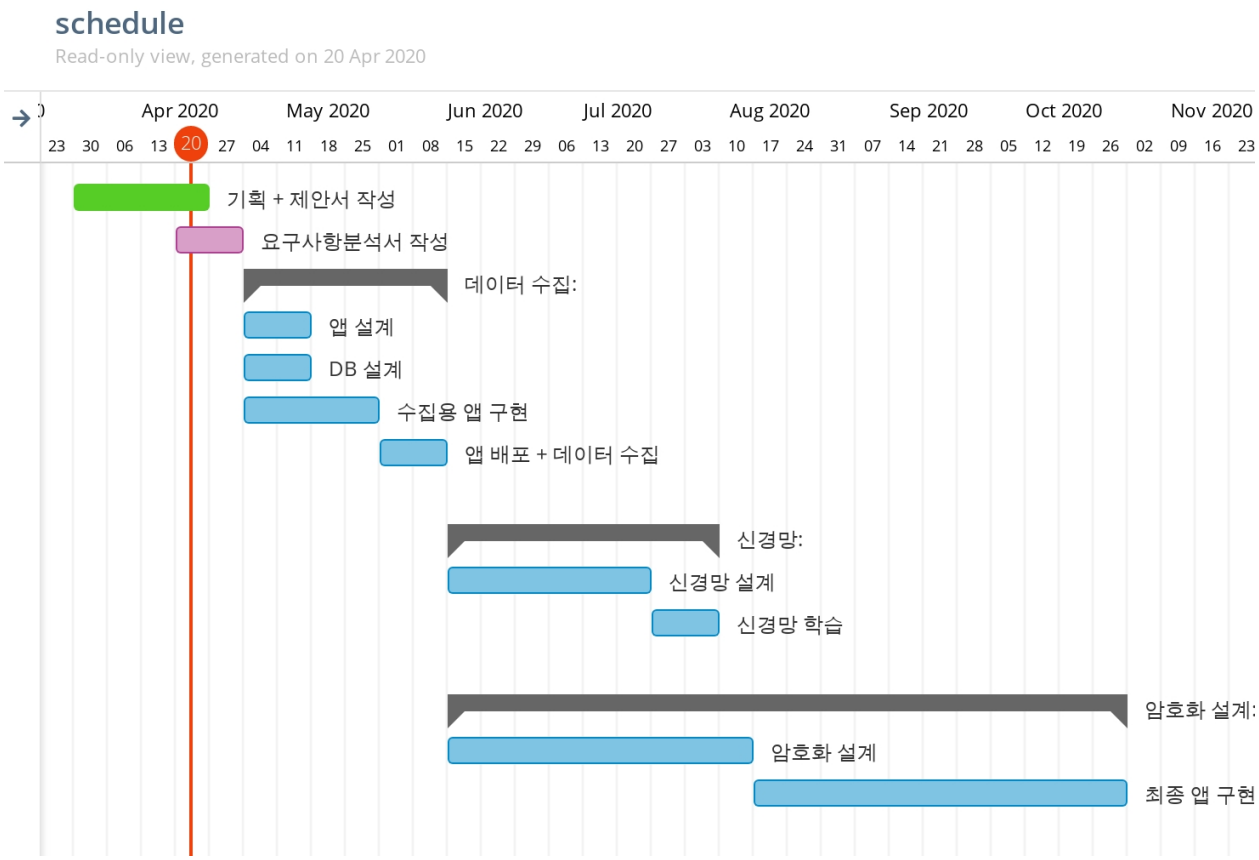
- 매일 사용자에게 스트레스 예측 수준에 대한 알림을 보내 사용자가 자신의 스트레스 지수를 인지할 수 있게 한다.

1.4 조원 구성 및 역할 분담

팀원	소속/학번	역할
	이름: 천세진 (팀장) 소속: 공과대학 컴퓨터공학부 학번: 201711356	- iOS 앱 설계 및 구현 - 암호화 시스템 설계 및 구축 - 신경망 설계 및 구현
	이름: 김지효 소속: 경영대학 기술경영학과 학번: 201612066	- Android 앱 설계 및 구현 - DB 설계 및 구축 - 신경망 설계 및 구현

1.5 일정

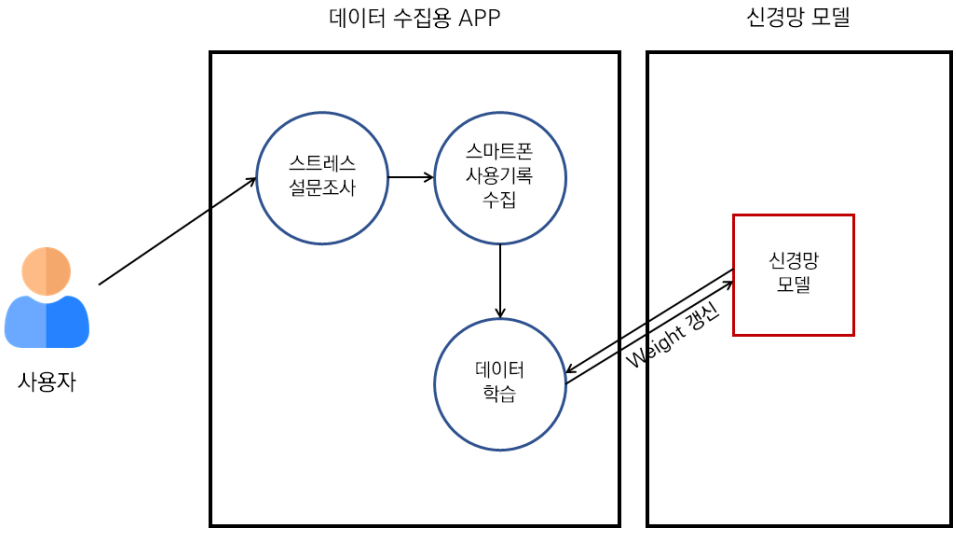
– 프로젝트 추진 일정 (2020.03~ 2020.11)



2 기능적 요구사항

2.1 Top Level Use Case Diagram

1. 신경망 모델 구축 단계



i. 스트레스 설문 조사

- 설문조사를 통해 사용자의 스트레스 척도를 파악할 수 있는 기능이다. 이 과정으로 사용자의 스트레스 지수에 대한 데이터 수집이 이루어진다.

ii. 스마트폰 사용기록 수집

- 사용자의 스마트폰 사용 기록을 수집하는 기능이다. 앱 사용 기록 등 사용자의 스마트폰 사용 패턴 분석에 필요한 데이터를 수집한다.

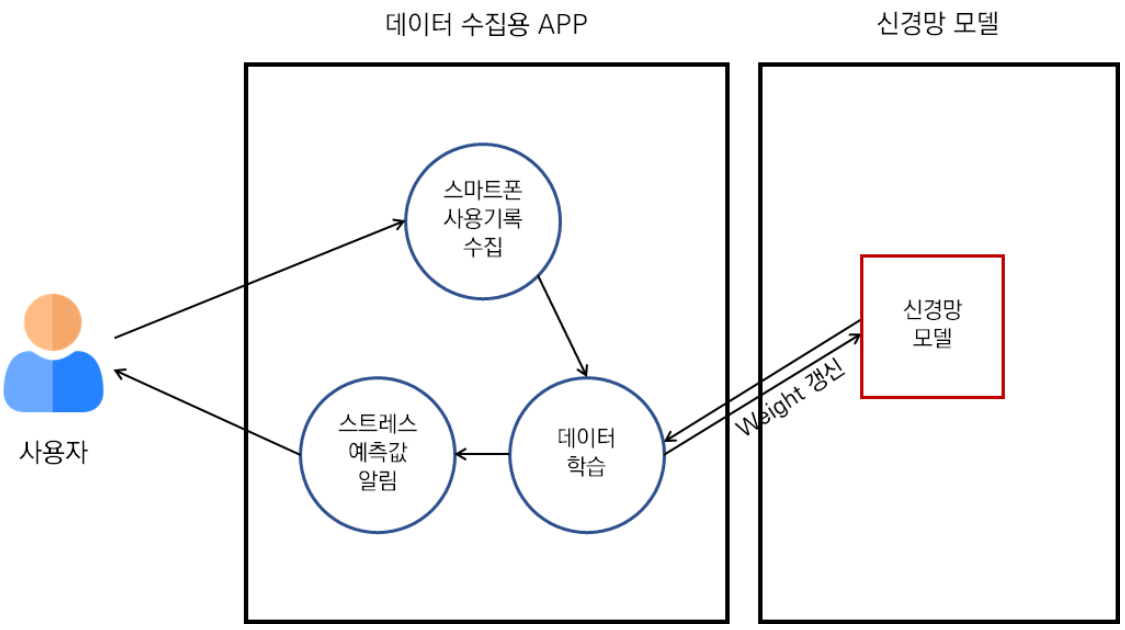
iii. 데이터 학습

- 연합 학습을 통해 수집한 데이터를 학습시킨다. 사용자의 기기에서 데이터 처리 및 모델링이 이루어진다.

iv. 신경망 모델 갱신

- 사용자 기기에서 갱신된 weight 값을 중앙 서버로 모아 신경망 모델의 성능을 개선한다.

2. 최종 구현 앱 단계



i. 스마트폰 사용기록 수집

- 사용자의 스마트폰 사용 기록을 수집하는 기능이다. 앱 사용 기록 등 사용자의 스마트폰 사용 패턴 분석에 필요한 데이터를 수집한다.

ii. 데이터 학습

- 연합 학습을 통해 수집한 데이터를 학습시킨다. 사용자의 기기에서 데이터 처리 및 모델링이 이루어진다.

iii. 신경망 모델 갱신

- 사용자 기기에서 갱신된 weight 값을 중앙 서버로 모아 신경망 모델의 성능을 개선한다.

iv. 스트레스 예측 값 알림

- 신경망을 거쳐 구해진 스트레스 예측 값을 사용자에게 알린다.