연구 현황 및 보고

증강지능연구실 황승현 2024-05-31

목차

- KoGES 심뇌혈관 질환
 - 연구 현황
 - 향후 목표
- SIGGRAPH Asia 2024 XR CovidHunter
- 허밍버드

KoGES 심뇌혈관

KoGES 심뇌혈관

• 심뇌혈관 질환 예측 모델 개발 및 주요 인자 찾기

- 전체 데이터 9622 x 32
- 심뇌혈관 질환 있음 9335
- 심뇌혈관 질환 없음 287



진행상황 - 결측값 대치

- 결측값 대치(Imputation)
 - 데이터셋에서 누락된 값을 채우는 과정

- 도메인에 맞게 결측값 대치
 - Domain-based Imputation
 - Expert-driven Imputation
- sequential K-nearest neighbor

결측값 대치 - SKNN

- K-최근접 이웃(KNN) 알고리즘을 사용하여 반복적으로 대치
- 초기 대치: 각 결측값을 K개의 가장 가까운 이웃
- 반복 대치: 초기 대치된 값을 사용하여 다시 KNN 알고리즘을 적용합니다.
- 수렴 검사: 각 반복 단계에서 결측값 대치 결과가 수렴할 때 까지 반복합니다.
- 최종 대치: 모든 결측값이 안정적으로 대치되면 반복을 종료

결측값 대치 - SKNN

```
# Sequential KNNImputer with exception handling for non-numeric columns
def sequential_knn_impute(df, n_neighbors=3):
    imputer = KNNImputer(n_neighbors=n_neighbors)
    df_imputed = df.copy()
    # Iterate over each column
    for column in df.columns:
       if df[column].dtype in [np.float64, np.int64]:
            subset = df_imputed[[column]]
            # Iterate until all NaNs in the specified column are filled
            while subset.isna().sum().sum() > 0:
                try:
                    subset = pd.DataFrame(imputer.fit_transform(subset), columns=[column], index=df.index)
                    df_imputed[column] = subset
                except ValueError:
                    print(f"Skipping column {column} due to incompatible data type.")
                    break
    return df_imputed
```

결측값 대치 - 목록

• AS1_02_GEN: AS1_MARRYA 결혼여부 ◦ 6: 기타 6이거나 99999인 사람이 정상적인 결혼생활일까? AS1_MARRYAETC을 보면 부인 가출, 행방불명, 동거 이렇게 되어있다. ◦ 1(미혼) 또는 4(별거)로 대치 AS1 02 GEN: AS1 JOBB 직업종류 。 이건 어떻게 해야하나 AS1_02_GEN: AS1_EDUA 교육수준 。 이건 어떻게 해야하나 AS1 02 GEN: AS1 INCOME 월평균수입 ㅇ 이건 직업종류, 교육수준과 관련 있지 않을까 ◦ AS1_JOBB, AS1_EDUA 별로 평균으로 대치 • AS1 03 DRSM: AS1 DRINK 음주여부 ○ 없는 거 보니 totalc / GH에서 걸러질 듯 • AS1 03 DRSM: AS1 SMOKEA 흡연여부, AS1 HVSMAM 하루 흡연량 。 이건 어떻게 해야하나 AS1 SMOKEA가 결측이면 AS1 HVSMAM도 결측 ○ AS1_HVSMAM이 결측이면 AS1_SMOKEA가 2, 3 또는 99999 。 AS1 SMOKEA가 2의 평균, 3의 평균으로 대치 AS1_04_ACTIVE: AS1_PHYACTL, AS1_PHYACTM, AS1_PHYACTH 。 이건 어떻게 해야하나 • AS1_05_MEDIC: AS1_HEALTH 전반적인 건강상태 ㅇ 이건 어떻게 해야하나 AS1 05 MEDIC: AS1 TIED 최근 피곤한지 。 이건 어떻게 해야하나 • AS1 13 SLEEP: AS1 SLPAMSF 수면시간 충분한지 ㅇ 이건 어떻게 해야하나 • AS1_16_DIET: AS1_RGMEALFQA 하루 정규식사 횟수 。 대충 3끼 먹는다고 가정..? 。 3끼 다 챙겨먹는 사람이 결측값이 되려나 • 결측값 있는 놈 없애볼까? ○ GH에서 빼면 얼마나 결측값이 사라지는지 관측해야할듯 ■ 이미 GH로 뺸거임... 결측값 있는 행 중에 심뇌혈관질환이 없다면? ■ 결측값 있는 행 1028 ■ 심뇌혈관질환 있는 행 36 전체 0 9726 ■ 전체 1 301 • 도메인 기반으로 값 일부 대치한 후

AS1_02_GEN, AS1_03_DRSM, AS1_04_ACTIVE, AS1_05_MEDIC, AS1_13_SLEEP, AS1_16_DIET

。 돌린 후 result 데이터프레임에 다시 집어넣기

- AS1_MARRYA 38
- AS1_JOBB 40
- AS1_EDUA 53
- AS1_INCOME 133
- AS1_DRINK 37
- AS1_SMOKEA 84
- AS1_HVSMAM 140
- AS1_PHYACTL 130 AS1_PHYACTM 195 AS1_PHYACTH 156
- AS1_HEALTH 28
- AS1_TIED 99
- AS1_SLPAMSF 50
- AS1_RGMEALFQA 63

결측값 대치 - 목록

• 새로 만든 데이터 고혈압여부: HAS HYPERTENSION ◦ 당뇨여부: HAS DIABETES 고지혈증: HAS HYPERLIPIDEMI LDL: LDL ○ H0MA IR: H0MA IR ◦ 허리둘레: WAIST AVG BMI: BMI 데이터 만들 때 사용한 변수도 결측값을 모두 대치한 다음 사용하자 결측값있는 가공데이터 목록 HOMA IR 280 WAIST AVG 8 BMI 4 RC 3 Al 2 HOMA IR AS1 18 BIOCHEM['AS1 GLU0 TR'] AS1 18 BIOCHEM['AS1 INS0'] WAIST AVG AS1 19 ANTHRO[['AS1 WAIST1', 'AS1 WAIST2', 'AS1 WAIST3']].mean(axis=1) BNI AS1 19 ANTHRO['AS1 HEIGHT'] AS1_19_ANTHRO['AS1_WEIGHT'] RC, AI (AS1 18 BIOCHEM) AS1 18 BIOCHEM['AS1 TCHL TR'] AS1 18 BIOCHEM['AS1 HDL TR'] AS1_18_BIOCHEM['AS1_LDL_TR'] AS1 18 BIOCHEM['AS1 TG TR']

• AS1 18 BIOCHEM, AS1 19 ANTHRO 두 데이터프레임을 통째로 impute

- HOMA_IR 280
 - AS1_18_BIOCHEM['AS1_GLU0_TR']
 - AS1_18_BIOCHEM['AS1_INS0']
- WAIST_AVG 8
- BMI 4
 - AS1_19_ANTHRO['AS1_HEIGHT']
 - AS1_19_ANTHRO['AS1_WEIGHT']
- RC, AI 3, 2
 - AS1_18_BIOCHEM['AS1_TCHL_TR']
 - AS1_18_BIOCHEM['AS1_HDL_TR']
 - AS1_18_BIOCHEM['AS1_LDL_TR']
 - AS1_18_BIOCHEM['AS1_TG_TR']

진행상황 - 데이터 가공

- 고혈압여부: HAS_HYPERTENSION
- 당뇨여부: HAS_DIABETES
- 고지혈증: HAS_HYPERLIPIDEMI
- LDL 콜레스테롤: LDL
- HOMA_IR: HOMA_IR
- 허리둘레: WAIST_AVG
- BMI: BMI
- Remnant cholesterol: RC
- Atherogenic Index: AI

2024-05-31

10

향후 목표

```
Confusion Matrix:

[[50 0]

[50 0]]

[[TP FN]

[FP TN]]

Accuracy: 0.500

Precison: 1.000

Recall : 0.000

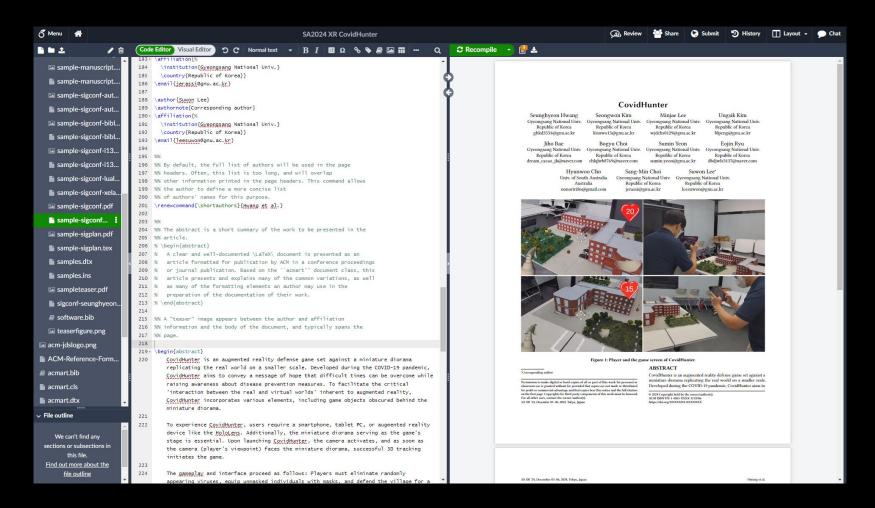
F1 Score: 0.000
```

- 사망률 데이터 원격 열람
- •모델 성능 개선

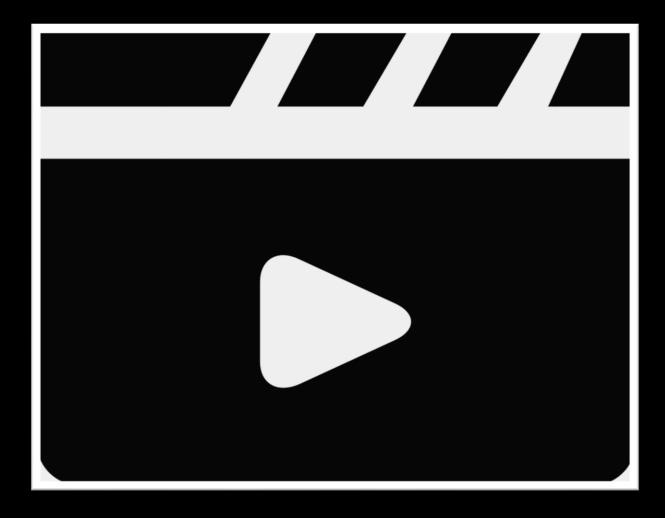
SIGGRAPH ASIA 2024 XR

CovidHunter

제출 준비 중



제출 예정일: 6.3.





HUMMINGBIRD





I AI VOICE MODEL VITS





학술대회 발표 예정

- 제70회 한국컴퓨터정보학회 하계학술대회
- 장소: 제주대학교
- 07-11 07-13일시:
- Deadline: 06-17
- 참석여부 미정

Proceedings of KSCI Conference 2024 (논문 문항: 2 또는 4 페이지)

허밍 버드: AI 음성 기술을 활용한 웰다잉 서비스

전윤찬", 김세희", 황승현^{여발표자 표기)} '중앙대학교 전자계산학과 ⁰경상국립대학교 컴퓨터공학과 e-mail:윤찬이 이메일", 526sally@swu,ac,kr", ghkd3531@gnu,ac,kr⁰

Humming Bird: Well-Dying Services Using Al Voice Technology

윤찬 영문이름", Se-Hee Kim", Seunghyeon Hwang^C 'Dept. of Computer Science, Korea-Computer University(유차이 학과) Dept. of Software Convergence, Seoul Women's University ^ODept. of Computer Science and Engineering, Gyeongsang National University

본 논문에서는 고성능 컴퓨팅 시스템의 성능 항상을 위한 효율적인 통격 작업부하 균등화 정책을 제안한다. 이 정책은 시스템 자원인 CPU와 메모리를 효율적으로 사용하여 고성능 컴퓨팅 시스템의 처리량을 최대화하고 각 작업의 수행시간을 최소화한다. 또한 이 경책은 수행중인 작업의 메모리 요구량과 각 노드의 부하 상태를 과 안하여 작업을 통적으로 할당하다. 이

• Neyword : 인공지능(Artificial Intelligence), 웹 다임(Well Dying), 대화(Conversation) + 키워드 생각나는

Introduction

디지털 기술이 발전하면서 고인을 추모하는 방법에 많 은 변화가 등장했다. 대표적으로 웹이나 페타버스를 이 용한 디지털 추모관이 있다. 그러나 해당 서비스는 시 간전인 요소에 치를되 하게가 인기 때문에 디지털 초모 관에 대한 새로운 때러다임이 필요하다.

이용한 AI 음성 대화형 추모 시스템을 제안한다. 교인 가하고 있음을 인식했다. 은 다양한 형태로 본인의 삶을 기록하는 과정에서, 유

죽은 A!를 통해 음성으로 고인과 상호작용하여 건강하 게 슬픔을 극복하여 웹 다잉(Well-Dying)을 경험하는 것을 목표로 한다.

II. Preliminaries

잘 사는 것을 중요하게 여기는 웰 빙(Well-Being)문 _ 고인을 추모하는 가상 공간의 개념은 지난 수년간 다양 화 뿐만 아니라 본인의 삶을 정리하고 죽음을 준비하는 - 한 학문적 및 실무적 맥락에서 탕구되어 왔다. 이 주제 웰 다잉(Well-Dying)문화가 확산되고 있다. 본 프로젝 - 에 대한 최초의 학문적 논의 중 하나는 1997년 사회복 트는 교인은 본인의 삶을 기록하고, 해당 기록된 데이 지학 교수인 Carla Sofka의 논문[*]으로 거슬러 올라 터와 유족이 보유한 고인의 음성 및 텍스트 데이터를 간다. 그녀는 추모를 위한 온라인 플랫폼의 사용이 중