

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국



(10) 국제공개번호

WO 2025/105790 A1

2025 년 5 월 22 일 (22.05.2025) WIPO | PCT

(51) 국제특허분류:

G16H 50/20 (2018.01)

A61B 5/00 (2006.01)

G16H 10/60 (2018.01)

G06N 20/00 (2019.01)

G16H 50/70 (2018.01)

경기도 하남시 위례중안로 215, 6409동 503호 (KR).
이광식 (LEE, Kwang-Sig); 01695 서울특별시 노원구
노해로85길 10-55, 514호 (KR).

(21) 국제출원번호:

PCT/KR2024/017753

(74) 대리인: 특허법인시공 (SIGONG IP&LAW FIRM);
06729 서울특별시 서초구 강남대로 275, 9층 (KR).

(22) 국제출원일:

2024 년 11 월 11 일 (11.11.2024)

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의
국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT,
AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH,
CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ,
EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR,
HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH,
KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD,
MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO,
NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW,
SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2023-0158569 2023 년 11 월 15 일 (15.11.2023) KR

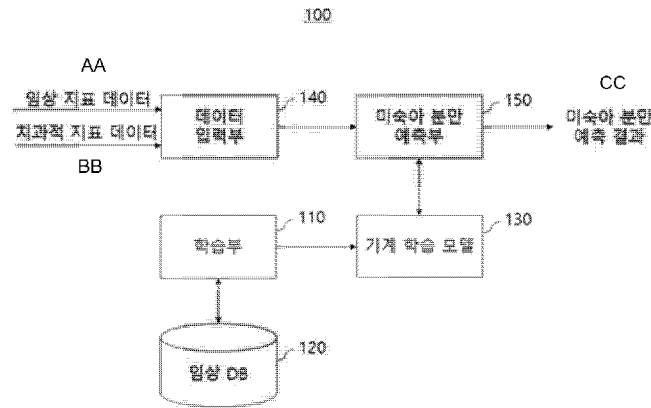
(71) 출원인: 고려대학교 산학협력단 (KOREA UNIVERSITY
RESEARCH AND BUSINESS FOUNDATION)
[KR/KR]; 02841 서울특별시 성북구 안암로 145 (KR).

(72) 발명자: 허주선 (HEO, Ju Sun); 02799 서울특별시
성북구 종암로24가길 80, 103동 1904호 (KR). 안기훈
(AHN, Ki Hoon); 06195 서울특별시 강남구 삼성로 417,
103동 1801호 (KR). 박정수 (PARK, Jung Soo); 13014

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의
역내 권리의 보호를 위하여): ARIPO (BW, CV, GH, GM,
KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ,

(54) Title: SYSTEM AND METHOD FOR PREDICTING PREMATURE CHILD DELIVERY ON BASIS OF
MACHINE-LEARNING UTILIZING CLINICAL INDEX AND DENTAL INDEX

(54) 발명의 명칭: 임상적 지표 및 치과적 지표를 활용하는 기계 학습 기반의 미숙아 분만 예측 시스템 및 방법



110 ... Training unit
120 ... Clinical DB
130 ... Machine-learning model
140 ... Data input unit
150 ... Premature child delivery prediction unit
AA ... Clinical index data
BB ... Dental index data
CC ... Premature child delivery prediction result

(57) Abstract: Disclosed are a system and method for predicting premature child delivery, which predict the probability of premature child delivery on the basis of a machine-learning analysis model by utilizing clinical indices and periodontal indices of prenatal mothers. A device for predicting a premature child delivery by a mother according to the present invention comprises: a data input unit for receiving clinical index data and dental index data of a mother; and a premature child delivery prediction unit for predicting a premature child delivery by the mother by means of a machine-learning model on the basis of the clinical index data and the dental index data, wherein the machine-learning model has been trained to predict premature child delivery on the basis of clinical index data collected for mothers and periodontal index data collected through dental examination and image capturing of the mothers.

(57) 요약서: 기계 학습 분석 모델을 기반으로 산전 산모의 임상 지표 및 치주 지표를 활용하여 미숙아 분만 확률을 예측하는 미숙아 분만 예측 시스템 및 방법이 개시된다. 본 발명에 따른 산모의 미숙아 분만 예측 장치는 산모의 임상 지표 데이터와 치과적 지표 데이터를 입력받는 데이터 입력부; 및 상기 임상 지표 데이터와 상기 치과적 지표 데이터를 기초로 기계 학습 모델-상기 기계 학습 모델은 산모들에 대해 수집된 임상 지표 자료와, 치과적 검진 및 상기 산모들에 대해 영상 촬영을 통해 수집된 치주 지표 자료를 기초로 미숙아 분만을 예측하도록 학습됨-에 의해 산모의 미숙아 분만을 예측하는 미숙아 분만 예측부를 포함한다.

[다음 쪽 계속]

UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

발명의 설명

발명의 명칭: 임상적 지표 및 치과적 지표를 활용하는 기계 학습 기반의 미숙아 분만 예측 시스템 및 방법 기술분야

- [1] 본 발명은 임상적 지표 및 치과적 지표를 활용하는 기계 학습 기반의 미숙아 분만 예측 시스템 및 방법에 관한 것으로, 보다 상세하게는 기계 학습 분석 모델을 기반으로 산전 산모의 임상 지표 및 치주 지표를 활용하여 미숙아 분만 확률을 예측하는 미숙아 분만 예측 시스템 및 방법에 관한 것이다.

배경기술

- [2] 미숙아 분만은 전 세계적으로 그 비율이 점차 증가하고 있어 산전에 미리 미숙아 분만을 예측하고 그에 대한 예방 대책 수립 및 출산 후 관리 계획을 세우는 것은 매우 중요하다. 일반적으로 산모의 미숙아 분만의 위험 요인은 크게 두 가지로 분류할 수 있다. 첫째는 산모/환경적 요인으로, 고령, 낮은 사회경제적 지위, 흡연, 음주, 마약, 우울, 영양, 체질량 지수 등의 요인이다. 둘째는 의학/산과적 요인으로, 이전 미숙아 분만, 다태 임신, 짧은 자궁 경부 길이, 요로 감염, 당뇨, 고혈압, 삼상선 질환, 조기양막파수, 태아 기형 등의 요인이다.
- [3] 다양한 위험 요인에도 불구하고 현재까지 미숙아 분만을 예측하는 모델의 정확도가 낮으며 미숙아 분만과 관련한 예측률이 낮은 수준에 그치고 있다. 이는 기존의 미숙아 분만 예측 모델은 주로 산모측 임상 요인만을 가지고 평가 진행하기 때문이다. 미숙아 분만 예측을 위해, 위험 요인 판단, 자궁 경부 측정, 생화학적 바이오마커 등을 이용하기도 하지만, 연구에 따라 서로 다른 기준과 예측률을 제시하고 있으며 일반화된 모델은 없는 실정이다. 한편, 산모의 치주염과 미숙아 분만에 대해서는 기존 연구들을 통해 그 연관성이 일부 알려져 있긴 하지만, 대부분의 연구들이 후향적 연구로서 연구 결과의 정확성이 떨어지고, 그 이전에 대해 명확히 밝혀진 바가 없다.
- [4] 본 발명은 한국연구재단의 창의도전연구기반 지원 사업(과제고유번호:1345364829, 세부과제번호: 2020R1I1A1A01073697, 연구과제명: 산모의 구강 내 세균 불균형과 장기적 신생아 예후 연관성 분석, 주관기관: 고려대학교 산학협력단, 연구기간: 2020.06.01 ~ 2023.05.31)의 일환으로 수행한 연구로부터 도출된 것이다.
- [5] 한편, 본 발명의 모든 측면에서 과제 제공 주체인 한국 정부의 재산 이익은 없다.

발명의 내용

기술적 과제

- [6] 본 발명은 기계 학습 분석 모델을 기반으로 산전 산모의 임상 지표 및 치주 지표를 활용하여 미숙아 분만 확률을 예측하는 미숙아 분만 예측 시스템 및 방법을 제공하기 위한 것이다.

과제 해결 수단

- [7] 본 발명의 일 양태에 따른 산모의 미숙아 분만 예측 장치는 산모의 임상 지표 데이터와 치과적 지표 데이터를 입력받는 데이터 입력부; 및 상기 임상 지표 데이터와 상기 치과적 지표 데이터를 기초로 기계 학습 모델- 상기 기계 학습 모델은 산모들에 대해 수집된 임상 지표 자료와, 치과적 검진 및 상기 산모들에 대해 영상 촬영을 통해 수집된 치주 지표 자료를 기초로 미숙아 분만을 예측하도록 학습됨 -에 의해 산모의 미숙아 분만을 예측하는 미숙아 분만 예측부를 포함한다.
- [8] 상기 임상 지표 데이터는 산모 연령, 임신 전 체질량 지수, 인공 수정 여부, 과거 조산 병력, 전자간증, 만성 고혈압, 임신성 당뇨, 조기 양막 파수 및 융모양막염 단계 중의 적어도 둘 이상을 포함할 수 있다.
- [9] 상기 치과적 지표 데이터는 치주염 단계, 변형된 치은 지수, 치태 지수 및 우식 경험영구치지수 중의 적어도 둘 이상을 포함할 수 있다.
- [10] 상기 기계 학습 모델은 랜덤 포레스트(Random forest) 모델을 기반으로 상기 임상 지표 데이터와 상기 치과적 지표 데이터의 변수 중요도를 예측하도록 학습된 모델일 수 있다.
- [11] 본 발명의 다른 양태에 따른 산모의 미숙아 분만 예측 방법은 데이터 입력부에 의해, 산모의 임상 지표 데이터와 치과적 지표 데이터를 입력받는 단계; 및 미숙아 분만 예측부에 의해, 상기 임상 지표 데이터와 상기 치과적 지표 데이터를 기초로 기계 학습 모델- 상기 기계 학습 모델은 산모들에 대해 수집된 임상 지표 자료와, 치과적 검진 및 상기 산모들에 대해 영상 촬영을 통해 수집된 치주 지표 자료를 기초로 미숙아 분만을 예측하도록 학습됨 -에 의해 산모의 미숙아 분만을 예측하는 단계를 포함한다.
- [12] 상기 산모의 미숙아 분만을 예측하는 단계는 상기 랜덤 포레스트 모델에 의해, 조산과 관련하여, 임신 전 체질량 지수, 변형 치은 지수, 전자간증, 우식경험영구치 지수, 인공 수정 여부, 과거 조산 병력, 산모 연령, 만성 고혈압, 임신성 당뇨, 조기 양막 파수 및 융모양막염 단계의 변수 중요성 순위를 결정하여 조산 확률을 예측하는 단계를 포함할 수 있다.
- [13] 본 발명의 또 다른 양태에 따르면, 상기 산모의 미숙아 분만 예측 방법을 실행하도록 컴퓨터에 의해 판독 가능한 기록 매체에 기록된 컴퓨터 프로그램이 제공된다.

발명의 효과

- [14] 본 발명의 실시예에 의하면, 기계 학습 분석 모델을 기반으로 산전 산모의 임상 지표 및 치주 지표를 활용하여 미숙아 분만 확률을 예측하는 미숙아 분만 예측 시스템 및 방법이 제공된다.

- [15] 또한, 본 발명의 실시예에 의하면, 산모 측 외에 산모 치주 상태와 관련한 지표를 예측 지표로 추가하고, 이를 기반으로 하는 기계 학습 모델을 도입함으로써 높은 예측 정확도로 미숙아 분만을 예측할 수 있다.
- [16] 또한, 본 발명의 실시예에 의하면, 미숙아 분만 예측에 대한 정확성을 높이고, 미숙아 분만이 예상되는 산모에 대한 개별화된 관리 및 예방 전략을 적용함으로써 미숙아 분만을 감소시키는 효과를 기대할 수 있다.

도면의 간단한 설명

- [17] 도 1은 본 발명의 실시예에 따른 산모의 미숙아 분만 예측 장치의 구성도이다.
- [18] 도 2는 본 발명의 실시예에 따른 산모의 미숙아 분만 예측 방법의 순서도이다.
- [19] 도 3은 다양한 기계 학습 모델에 의해 조산(전체 조산, 자연 조산)과 만삭 확률을 예측한 결과를 나타낸 것이다.
- [20] 도 4는 랜덤 포레스트 기계 학습 모델을 통한 변수 중요도의 산출 결과를 나타낸 것이다.

[21]

발명의 실시를 위한 형태

- [22] 본 발명의 목적 및 효과, 그리고 그것들을 달성하기 위한 기술적 구성들은 첨부되는 도면과 함께 상세하게 후술되는 실시예들을 참조하면 명확해질 것이다. 본 발명을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 그리고 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다.
- [23] 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있다. 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [24] 본 명세서에서 사용되는 '~부'는 적어도 하나의 기능이나 동작을 처리하는 단위로써, 예를 들어 소프트웨어, FPGA 또는 하드웨어 구성요소를 의미할 수 있다. '~부'에서 제공하는 기능은 복수의 구성요소에 의해 분리되어 수행되거나, 다른 추가적인 구성요소와 통합될 수도 있다. 본 명세서의 '~부'는 반드시 소프트웨어 또는 하드웨어에 한정되지 않으며, 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고, 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 이하에서는 도면을 참조하여 본 발명의 실시예에 대해서 구체적으로 설명하기로 한다.
- [25] 본 발명은 전향적 연구 설계에 따라 산전 산모 위험 요인과 치주 요인의 정확한 수집 및 기계 학습 분석을 통한 미숙아 예측 모델을 개발한 것으로, 산전 산모측

위험 요인과 치주 요인을 병합한 미숙아 분만 예측 모델에 관한 것이다. 본 발명의 실시예에 따른 산모의 미숙아 분만 예측 장치 및 방법은 산모들에 대해 수집된 임상 지표 자료와, 치과적 검진 및 산모들에 대해 영상 촬영을 통해 수집된 치주 지표 자료를 기초로 미숙아 분만을 예측하도록 기계 학습 모델을 학습하고, 학습된 기계 학습 모델에 의해 산모의 임상 지표 데이터(예를 들어, 산모 연령, 임신 전 체질량 지수, 인공 수정 여부, 과거 조산 병력, 전자간증, 만성 고혈압, 임신성 당뇨, 조기 양막 파수, 융모양막염 단계 등)와 치과적 지표 데이터(예를 들어, 치주염 단계, 변형된 치은 지수, 치태 지수 및 우식경험영구치지수 등)를 기초로 산모의 미숙아 분만을 예측한다. 실시예에서, 기계 학습 모델은 랜덤 포레스트(Random forest) 모델을 기반으로 임상 지표 데이터와 치과적 지표 데이터의 변수 중요도를 예측하도록 학습된 모델일 수 있다.

- [26] 도 1은 본 발명의 실시예에 따른 산모의 미숙아 분만 예측 장치의 구성도이다. 도 1을 참조하면, 본 발명의 실시예에 따른 산모의 미숙아 분만 예측 장치(100)는 학습부(110), 임상 데이터베이스(DB; database, 120), 데이터 입력부(140) 및 미숙아 분만 예측부(150)를 포함할 수 있다.
- [27] 학습부(110)는 산모들에 대해 수집된 임상 지표 자료와, 치과적 검진 및 산모들에 대해 영상 촬영을 통해 수집된 치주 지표 자료를 기초로 미숙아 분만을 예측하도록 기계 학습 모델(130)을 학습할 수 있다. 실시예에서, 기계 학습 모델(130)은 학습부(110)에 의해 랜덤 포레스트(Random forest) 모델을 기반으로 임상 지표 데이터와 치과적 지표 데이터의 변수 중요도를 예측하도록 학습될 수 있다.
- [28] 임상 DB(120)는 기계 학습 모델(130)의 학습을 위한 각종 임상 지표 자료와 치주 지표 자료를 저장할 수 있다. 이러한 임상 지표 자료와 치주 지표 자료는 산모들, 특히 고위험군 산모들과 이에 해당하지 않는 산모들에 대해 수집된 임상 지표(예를 들어, 산모 연령, 임신 전 체질량 지수, 인공 수정 여부, 과거 조산 병력, 전자간증, 만성 고혈압, 임신성 당뇨, 조기 양막 파수, 융모양막염 단계 등)와 치주 지표(예를 들어, 치주염 단계, 변형된 치은 지수, 치태 지수, 우식경험영구치지수 등)를 포함할 수 있다.
- [29] 데이터 입력부(140)는 미숙아 분만 예측의 대상인 임신한 산모와 관련된 임상 지표 데이터와 치과적 지표 데이터를 입력받을 수 있다. 실시예에서, 데이터 입력부(140)에 의해 입력되는 임상 지표 데이터는 산모 연령, 임신 전 체질량 지수, 인공 수정 여부, 과거 조산 병력, 전자간증, 만성 고혈압, 임신성 당뇨, 조기 양막 파수, 융모양막염 단계 중의 적어도 둘 이상을 포함할 수 있다. 실시예에서, 데이터 입력부(140)에 의해 입력되는 치과적 지표 데이터는 치주염 단계, 변형된 치은 지수, 치태 지수, 우식경험영구치지수 중의 적어도 둘 이상을 포함할 수 있다.
- [30] 미숙아 분만 예측부(150)는 데이터 입력부(140)에 의해 입력받은 산모의 임상 지표 데이터와 치과적 지표 데이터를 기초로, 학습된 기계 학습 모델(130)에 의해 산모의 미숙아 분만을 예측할 수 있다. 실시예에서, 미숙아 분만 예측부(150)

는 랜덤 포레스트 모델에 의해 예측된 임상 지표 데이터와 치과적 지표 데이터의 변수 중요도를 기초로 미숙아 분만을 예측할 수 있다.

- [31] 도 2는 본 발명의 실시예에 따른 산모의 미숙아 분만 예측 방법의 순서도이다. 도 1 및 도 2를 참조하면, 본 발명의 실시예에 따른 산모의 미숙아 분만 예측 방법은 학습부(110)에 의해 산모들에 대해 수집된 임상 지표 자료와, 치과적 검진 및 산모들에 대해 영상 촬영을 통해 수집된 치주 지표 자료를 기초로 미숙아 분만을 예측하도록 기계 학습 모델(130)을 학습하는 단계(S10)와, 데이터 입력부(140)에 의해 미숙아 분만 예측의 대상인 임신한 산모와 관련된 임상 지표 데이터와 치과적 지표 데이터를 입력받는 단계(S20), 그리고 미숙아 분만 예측부(150)에 의해 산모의 임상 지표 데이터와 치과적 지표 데이터를 기초로 기계 학습 모델(130)에 의해 산모의 미숙아 분만을 예측하는 단계(S30)를 포함할 수 있다.
- [32] 이하에서 본 발명의 실시예에 따른 산모의 미숙아 분만 예측 장치 및 방법의 성능 검증을 위한 실험 내용에 대해 설명한다. 중환자실에 입원한 고위험 산모들을 전향적으로 등록하여 산모들을 대상으로 여러 산과측 임상 지표 자료 뿐 아니라, 치과적 검진 및 파노라마 x-ray 영상 촬영을 통해 치주 지표를 함께 수집하였다. 또한, 산모들로부터 출생된 신생아들을 대상으로, 신생아와 관련한 여러 임상 지표 자료를 수집하였다. 만삭아와 조산아로 군을 분류하여 두 군 간의 임상적, 치과적 지표차이를 분석하였으며, 이를 기반으로 기계 학습을 통한 모델을 구축하여, 미숙아 분만 예측 모델을 구현하고 이의 성능 및 지표 중요도를 측정하였다.
- [33] 도 3은 다양한 기계 학습 모델에 의해 조산(전체 조산, 자연 조산)과 만삭 확률을 예측한 결과를 나타낸 것이다. 도 3에서, 'LR'은 로지스틱 회귀(logistic regression) 모델, 'DT'는 결정 나무(decision tree) 모델, 'NB'는 나이브 베이즈(naive bayes) 모델, 'RF'는 랜덤 포레스트(random forest) 모델, 'SVM'은 서포트 벡터 머신(support vector machine), 'ANN'은 인공신경망(artificial neural network) 모델의 약어이고, 'AUC'는 ROC(receiver operating characteristic curve) 등의 곡선 아래 면적(area under the curve)의 약어이다. 랜덤 포레스트 모델을 이용한 경우, 다른 기계 학습 모델보다 높은 83%의 예측 정확도를 나타냈으며, AUC 특성 또한 다양한 기계 학습 모델 중 최고 수준을 나타내는 것으로 확인되었다.
- [34] 도 4는 랜덤 포레스트 기계 학습 모델을 통한 변수 중요도의 산출 결과를 나타낸 것이다. 랜덤 포레스트 모델은 결정 나무를 기본 모델로 사용하는 앙상블 방법으로, 결정 나무(Decision Tree)를 여러개 만들어서 그 결과들을 종합적으로 고려하여 결론을 도출하는 방법이다. 결정 나무는 중간 지점은 가지를 치는 기준이 되는 독립변수에 대한 검정 조건을 나타내고, 가지는 검정 결과를 나타내며, 최종 지점은 검정 결과에 따른 종속변수 값을 나타내는 나무 구조의 인공지능 모형이다. 랜덤 포레스트 모델은 많은 수의 결정 나무를 훈련 시킨 뒤 이들이 다수결 투표를 하도록 해서 종속변수 값을 예측하는 인공지능 모형이다. 예를 들어, 조산 여부나 출생 체중 등이 랜덤 포레스트 분석의 종속변수로 고려될 수 있다.

- [35] 예를 들어, 자료 크기가 100이고, 나무(Tree) 숫자가 100인 경우, 랜덤 포레스트 훈련 과정은 다음과 같다. (1) 참여자가 100명인 (크기가 100인) 훈련자료에 대해 대체를 통한 무작위 선택을 실행해서 참여자가 100명인 (크기가 100인) 하위 자료를 생성한다. (2) 독립변수 가운데 일부 또는 전부를 대체 없이 무작위로 선택하고, 나무를 생성한다. (3) 위 과정을 1000번 실행하여 1000개의 나무를 생성하고 하위 자료들에 포함되지 않은 자료 (Out of Bag) 각 관측값에 대해 다수결 원칙을 적용해서 예측값을 도출한다.
- [36] 기계 학습과 전향적 코호트 데이터를 사용하여 임상 및 치과 예측 변수와 조산 간의 연관성을 테스트하였다. 제왕절개를 통해 단생아 출산한 60명의 여성으로부터 전향적 코호트 데이터를 얻었다. 이 중 30명은 임신 37주 이전에 아이를 분만한 산모로, 조산(PTB; preterm birth)으로 분류되었다. 나머지는 임신 37주(만기출산) 후에 출산한 만삭 산모로 분류되었다. 종속변수는 조산(PTB)과 자연조산(SPTB; spontaneous preterm birth), 만삭 출산으로 하였다. 임상 검사 및 기록에서 고려된 20개의 변수 중 15개의 독립 변수(의학적 요인 10개, 치과 요인 5개)가 기계 학습 분석에 포함되도록 선택되었다. 랜덤 포레스트(RF; Random Forest) 변수 중요도는 PTB/SPTB의 주요 예측 변수를 식별하는데 사용되었다. 예측 변수와 PTB/SPTB 간의 연관성 방향을 분석하기 위해 SHAP(Shapley Additive explain) 값을 계산하였다.
- [37] 랜덤 포레스트 분석 결과, 임신 전 체질량 지수(BMI), 변형 치은 지수(MGI; modified gingival index), 전자간증, 우식경험영구치(DMFT) 지수, 산모 연령은 조산(PTB)에 대한 상위 5개 RF 변수 중요성 순위에 포함되었다. 자연조산(SPTB)의 경우 진통 전 조기 양막 파수(PROM), 임신 전 체질량지수(BMI), 산모 연령, 우식경험영구치(DMFT; decayed missing filled teeth) 지수, 융모양막염(CAM) 단계가 상위 5개 요인에 해당하였다. SHAP 값의 측면에서, PTB/SPTB와 조기 양막 파수(PROM; premature rupture of membrane), 임신 전 체질량지수(BMI), 산모 연령, 변형 치은지수(MGI), 전자간증 및 융모양막염(CAM; chorioamnionitis) 단계와 같은 주요 예측 변수 사이에 일부 긍정적인 상관 관계가 발견되었다. 임상 대응물과 함께 변형 치은지수(MGI)는 PTB/SPTB의 주요 예측변수 중 하나인 것으로 밝혀졌다.
- [38] 임신 37주 이전의 출산으로 정의되는 조산(PTB)은 신생아, 영유아 및 5세 미만 어린이에게 상당한 이환율과 사망률을 초래할 수 있는 가장 흔하고 심각한 임신 합병증 중 하나이다. 조산(PTB)으로 인한 영아 미숙아는 뇌실내 출혈, 미성숙 망막병증, 괴사성 장염, 기관지폐 이형성증, 시각 및 청각 장애, 운동 장애, 신경 발달 장애 등 여러 가지 심각한 주산기 질환을 유발할 수 있다. 이들 중 일부는 신생아기뿐만 아니라 평생 동안 나타나는 평생 합병증이다.
- [39] 조산(PTB)은 크게 두 가지 하위 집합으로 분류될 수 있다. 첫째는 전자간증 또는 불안정 태아 상태와 같은 산모 또는 태아의 징후로 인해 분만이 유도되거나 유아가 사전 진통 없이 제왕절개로 분만되는 의도된 조산이고, 둘째는 진통 전

양막 파열(PROM) 유무에 관계없이 조기 진통이 발생하는 자연적 PTB(SPTB)이다. 자연조산(SPTB)은 전체 조산(PTB)의 약 2/3를 차지하며 예측하기 어렵다. 조산 또는 자연조산에 대해 자주 논의되는 위험 요인은 크게 두 가지 범주로 분류할 수 있다. 첫째는 고령, 낮은 사회 경제적 지위, 흡연, 음주, 약물 사용, 기분 장애 또는 심리적 스트레스를 포함한 산모/환경 요인, 영양 요인 및 임신 전 체질량 지수(BMI)이다. 둘째는 이전 조산, 다태 임신, 짧은 자궁 경부 길이, 요로 감염, 당뇨병, 고혈압 및 갑상선 장애, 조기 양막파수(PROM) 및 태아 기형을 포함한 의학 적/산과적 요인이다.

- [40] 수십 년간의 연구를 통해 확인된 수많은 위험 요소에도 불구하고 PTB에 대한 예측률은 여전히 낮은 상황이다. 요즘 일반적으로 고려되는 PTB 예측 전략에는 위험 평가, 자궁경부 길이(CL)의 연속 측정, 태아 피브로넥틴 및 염증성 사이토카인을 포함한 생화학적 바이오마커 평가가 포함된다. CL 측정은 임신 중기부터 산부인과에서 일상적인 절차가 되었지만, 특히 저위험군에서 임상적 유용성은 여전히 논란의 여지가 남아 있다. 피브로넥틴 스크리닝에 대해서는 권장 사항이 일치하지 않으며, 낮은 비용 효율성과 환자 부담도 고려해야 한다.
- [41] 치주염은 구강에서 발생하는 만성 염증성 질환으로, 전자간증 및 저체중아를 비롯한 PTB 및 임신 합병증과 널리 연관되어 있다. 치주염과 조산(PTB)은 낮은 사회경제적 지위, 높은 BMI, 흡연, 음주, 당뇨병 및 고혈압을 포함한 의학적 요인과 같은 위험 요인을 공유한다. 이론적으로 치주염은 직접적 또는 간접적인 경로를 통해 태반 조직의 염증 상태를 상승시켜 PTB로 이어지는 노동 단계의 부적절한 활성화에 기여할 수 있다.
- [42] 본 발명은 전향적으로 수집된 데이터와 기계 학습 방법을 사용하여 조산(PTB)의 예측 변수로서 치주염 관련 매개 변수의 중요도를 예측하여 이를 기초로 조산의 예측 정확도를 높이는 것을 특징으로 한다. 기계 학습 기반 데이터 분석은 전통적인 선형 회귀 모델과 달리 예측변수와 결과 사이의 차이를 계산하며, "*ceteris paribus*" (다른 모든 변수는 일정하게 유지됨)에서 상대적으로 자유롭고 복잡한 비선형 관계를 모델링할 수 있기 때문에 위험 요소 평가에서 기존 접근 방식보다 우수한 것으로 입증되었다. 따라서 위험 요인 분석에 기계 학습 방법을 사용하면 기존 접근 방식을 사용한 분석의 공백을 메워 질병에 대한 추가 검사 도구를 식별하는 것이 가능해진다. 이에 본 발명은 기계 학습 및 전향적 코호트 데이터를 사용하여 조산의 임상 예측 변수와 치과 예측 변수 간의 연관성을 테스트하였다.
- [43] 전향적 코호트 연구는 고려대학교 안암병원 임상시험심사위원회(IRB)의 승인을 받아 수행되었다. 연구는 조산 산모를 대상으로 한 구강 미생물군집 연구의 일환으로 수행되었다. 단태아 제왕절개를 위해 입원한 임산부를 모집하였으며, 임신 37주 이전에 분만한 환자 30명과 만삭아 30명의 두 그룹으로 구성되었다. 의료 및 산과 데이터는 출산 전후에 수집되었다. 여기에는 산모 연령, 임신 전 체질량 지수(BMI), 임신 중 BMI 변화, 체외 수정 이력, 이전 조산(PTB) 경험, 전자간

중, 만성 고혈압, 임신성 당뇨병(GDM), 조기 양막파수(PROM), 조직학적 용모양막염(CAM), 장염 등이 포함되었다. 조기 진통은 자궁경부가 1cm 이상 확장되고 규칙적인 자궁 수축이 일어나는 것으로 정의된다.

- [44] 참가자들에 대해 출산 후 입원 기간 동안 전체 구강 치주 검사 전에 파노라마 방사선 사진을 촬영하였다. 측정에는 치주 탐침 깊이(PPD; periodontal probing depth), 임상 부착 수준(CAL; clinical attachment level), 변형 치은 지수(MGI; modified gingival index) 및 플라크 지수(PI; plaque index)가 포함되었다. PPD 및 CAL 측정을 바탕으로 이개부 침범과 같은 치료 복잡성 결정 요인을 밝혀낸 방사선학적 검사 외에도 각 환자에 대한 치주염 단계를 결정하였다. 치주염의 중증도는 치주염 분류에 따라 0~3단계로 분류되었다. 0단계는 치주염의 징후가 없는 건강한 치주를 나타낸다.
- [45] 미혼모를 출산한 여성 60명으로부터 전향적 코호트 데이터를 얻었다. 종속변수는 PTB/SPTB으로 하였다. 분석에는 15개의 독립변수가 포함되었다. 로지스틱 회귀(LR), 의사결정 트리(DT), naive bayes(NB), 랜덤 포레스트(RF), 지원 벡터 머신(SVM) 및 인공 신경망(ANN)을 사용하여 PTB/SPTB를 예측하였다. 랜덤 포레스트(RF)는 종속 변수("부트스트랩 집계")에 대해 다수결을 수행하는 의사결정 나무 모델이다. 원본 데이터에 42명의 참가자가 포함되어 있다고 가정하고, 1000개의 의사결정 나무가 있는 랜덤 포레스트를 예로 들어 설명한다. 랜덤 포레스트 훈련 및 테스트에는 두 단계가 포함된다. 먼저, 무작위 추출과 복원을 기반으로 42명의 참여자로부터 새로운 데이터를 생성하고, 이러한 새로운 데이터를 기반으로 의사결정 나무를 생성한다. 원본 데이터의 일부 참여자는 새 데이터에서 제외되며, 이러한 남은 데이터를 Out-of-bag 데이터라고 합니다. 이 과정을 1000번 반복하여 수행한다.
- [46] 둘째, 1000개의 의사결정 나무는 Out-of-bag 데이터의 모든 참가자에 대한 종속 변수를 예측한다. 이들의 과반수 표를 해당 참가자에 대한 최종 예측으로 삼고, Out-of-bag 데이터 중 잘못된 표의 비율로 Out-of-bag 오류를 계산한다. 전체 정보가 포함된 60개 사례의 데이터는 75:25 비율(42개 대 18개 사례)로 훈련 및 검증 세트로 분할되었다. 훈련된 모델의 검증 기준은 정확도(18개 사례 중 올바른 예측 비율)와 수신기 작동 특성 곡선(receiver operating characteristic curve)의 아래 영역(AUC)으로 하였다.
- [47] 랜덤 포레스트의 순열 중요도(permutation importance)가 PTB/SPTB의 주요 예측 변수를 식별하는 데 사용되었다. 예측 변수와 PTB/SPTB 간의 연관성 방향을 분석하기 위해 SHAP(Shapley Additive explain) 값을 계산하였다. RF 순열 중요도는 예측 데이터가 무작위로 섞일 때 정확도의 전반적인 감소를 측정하는 방식으로 선정되었다. 이러한 정확도의 감소는 모델이 예측 변수에 의존하는 정도를 나타낸다. SHAP 값은 예측 변수가 있는 경우와 없는 경우에 기계 학습에서 예측한 PTB/SPTB 확률 간의 차이를 측정하여 산출되었다. 예측 변수의 순위를 도출하

기 위해 RF 순열의 중요도가 사용되었다. 각 예측 변수와 종속 변수 PTB/SPTB 간의 연관성 방향을 평가하기 위해 SHAP 플롯이 생성되었다.

- [48] 다양한 의료 변수 중 전자간증, 조기 양막 파수(PROM), 조기 진통 및 분만 시 진통의 발생률은 PTB 그룹에서 유의하게 더 높게 나타났다($P<0.05$). 산모 연령, 체외 수정 발생률 및 임신성 당뇨병(GDM)은 PTB의 위험 요인으로 알려져 있지만 연구 집단에서는 유의미한 차이를 보이지 않았다. 치주염 단계, DMFT(우식경험영구치) 지수, 치태 지수는 두 군 간 차이가 없었으나, 변형 치은 지수(MGI)는 PTB군에서 유의하게 높은 값을 보였다($P<0.05$).
- [49] 랜덤 포레스트(RF) 모델은 로지스틱 회귀(LR) 모델과 비슷한 성능을 나타내었다. 랜덤 포레스트(RF) 변수 중요도를 기준으로, 임신 전 BMI, 변형 치은 지수(MGI), 전자간증, 우식경험영구치(DMFT) 지수 및 산모 연령이 PTB의 상위 5개 항목에 포함되었다. 조기 양막 파수(PROM), 임신 전 BMI, 산모 연령, 우식경험영구치(DMFT) 지수 및 용모양막염(CAM) 단계의 랜덤 포레스트(RF) 변수 중요도 순위는 SPTB의 상위 5개 항목에 포함되었다.
- [50] 도 4는 랜덤 포레스트의 각 독립 변수에 대한 SHAP 값과 각 독립 변수에 대한 SHAP 값을 나타낸 것이다. 각 독립 변수에 대한 SHAP 값은 기계 학습 분석에 특정 독립변수가 포함될 때 종속변수(PTB 또는 SPTB)의 확률이 감소하거나 증가함을 나타낸다. 전반적으로, 최소 SHAP(음수)의 절대값보다 큰 최대 SHAP(양수)의 절대값은 예측변수와 종속변수 사이의 양의 관계를 나타낸다. 특히 랜덤 포레스트(RF) 순열 중요도의 상위 순위에 있는 변수들 사이에서 일부 긍정적인 연관성이 관찰되었다.
- [51] 본 발명에서는 기계 학습 기법을 이용하여 구축한 조산(PTB) 예측 모델에서 변형 치은 지수(MGI), 우식경험영구치(DMFT) 지수 등 치과적 요인이 주요 예측변수로 작용할 수 있음을 제시하였다. 본 발명이 기존 연구와 다른 점은 잘 알려진 의학적 위험인자 외에 다양한 의학적 배경, 산과적 병력 등 치과적 요인을 추가했다는 점이다. AI나 다변량 회귀분석을 활용한 기존 예측 모델에서는 AUC를 이용해 계산한 모델 성능이 0.61~0.71 범위에 불과하였다. 본 발명에 의하면, 테스트된 5가지 기계 학습 방법(DT, NB, RF, SVM, ANN) 중 랜덤 포레스트(RF) 모델이 조산(PTB) 예측 모델에서 AUC 0.73, 자연조산(SPTB) 예측 모델에서 0.86으로 가장 높은 성능을 보였다.
- [52] RF 변수 중요도 순위는 랜덤 포레스트 모델에서 치과적 요인의 중요성을 검증하는데 활용되었다. 변형 치은 지수(MGI)는 PTB에서 2위, SPTB 모델에서 6위를 차지했다. 이는 산모 연령(5위), 이전 PTB(14위), 전자간증(3위), 만성 고혈압(15위), 임신성 당뇨병(10위) 등 의학적 관점에서 잘 알려진 PTB 위험 요소보다 순위가 높았다. 변형 치은 지수(MGI)는 검사 기간 동안의 치은 염증 수준을 나타내며 만성 염증에 대한 민감성을 나타낸다. 변형 치은 지수(MGI)가 높을수록 숙주의 치은염 민감성을 의미 있게 반영하며, 이는 특히 젊은 성인의 미래 치주염과 강한 상관관계가 있다. 이러한 맥락에서 산전 선별검사나 임신 중 정기 검진에서

높은 변형 치은 지수(MGI), 특히 평균 $MGI \geq 2$ (그림 2a)는 임신 중에 발생할 수 있는 감염에 대한 산모의 감수성이 높거나 진행성 경향이 있음을 나타낸다.

- [53] 본 발명에 따른 전향적 연구에 따라, 치주염 단계, 변형 치은 지수(MGI), PI 및 우식경험영구치(DMFT) 지수를 포함한 여러 치과 매개 변수에 대한 데이터가 수집되었다. 치주염 병기 결정은 질병의 중증도를 나타내며 새로 확립된 치주염 분류를 기반으로 한다. 병기 설정 시스템을 사용하여 평가한 참가자의 치주 상태는 두 그룹 간에 차이가 없었다. 3단계와 4단계에 해당하는 중등도/중증 치주염 사례는 거의 없었으며 유병률은 그룹 간에 차이가 없었다. 치주염(1기 이상)과 건강한 치주(0기) 사이의 유의미한 유병률 차이는 관찰되지 않았다. 이러한 결과는 일반적으로 조산아에서 치주염 발생률이 더 높았던 이전 데이터와 다소 다른 결과이다.
- [54] SPTB 분석 모델에서 치과 매개변수의 높은 AUC(0.86)를 고려할 때, 양막 파열(SPTB)이 있거나 없는 조기 분만은 PTB의 주요 원인으로 간주될 수 있다. 이는 감염, 혈관 장애, 탈락막 노화, 산모-태아 내성 붕괴, 프로그레스테론 작용 감소, 자궁경부 질환 등 다양한 원인이 있는 증후군으로 간주된다. 산모 면역학의 섬세한 균형은 건강한 임신과 분만 과정의 전제 조건이다. 정지 상태에서 전염증성 상태로의 전환은 진통을 활성화하고 인터루킨(IL)-1, IL-6 및 종양 괴사 인자(TNF)- α 와 같은 염증성 사이토카인이 이 과정에 관여한다.
- [55] 이러한 맥락에서 임산부가 BMI, 스트레스, 행동 요인, 유전적 요인, 영양 결핍 등 다양한 요인과 관련된 감염에 취약해지면 분만 과정이 비정상적으로 앞당겨질 수 있다. 본 발명의 연구 결과에 따르면, 높은 변형 치은 지수(MGI)는 산모의 감염 감수성에 대한 선별 도구로 작용할 수 있다. 이는 산전 선별검사에 포함될 수 있는 간단하고 비침습적이며 비용 효과적인 검사일 수 있다. 또한 MGI 점수가 높은 산모의 경우 치태에 병원성 세균이 더 많이 있을 가능성이 높으므로 구강 병원균의 혈행성 전파를 예방하기 위해 산전이나 중기에 치주 치료를 시행하는 것이 바람직하다.
- [56] 의존성 플롯은 우식경험영구치(DMFT) 지수 ≤ 10 범위 내에서 음의 상관 관계를 나타냈으며, 이는 충치에 민감한 환자가 SPTB에 걸릴 확률이 낮음을 의미한다. 치아우식증과 치주염은 구강에서 발생하는 서로 대조되는 두 가지 감염이다. 각 질병과 주로 관련된 박테리아 종은 거의 반대되는 특성을 갖기 때문이다. 이러한 맥락에서 치주염에 민감한 환자는 우식 경향이 낮으므로 우식경험영구치(DMFT) 지수가 낮고 SHAP 값이 높다. 본 발명은 랜덤 포레스트(RF) 변수 중요도 및 SHAP 요약/중속성 플롯과 같은 최첨단 AI 접근 방식을 사용하여 주요 예측 변수를 식별하고 해당 연관성의 방향을 설명할 수 있다. 선형 또는 로지스틱 회귀 분석과 같은 기존 통계 접근 방식에서는 비현실적인 데이터 가정이 방법론, 즉 "다른 모든 변수가 일정하게 유지되는" *ceteris paribus*의 기초가 된다. 대조적으로 SHAP는 모든 현실적인 시나리오를 고려한다.

- [57] SPTB의 세 가지 예측 변수 조기 양막파수(PROM), 임신 전 BMI 및 변형 치은 지수(MGI)가 있다고 가정한다. 참가자에 대한 변형 치은 지수(MGI)의 SHAP 값은 다음 네 가지 시나리오의 평균이다. (1) PROM 제외, 임신 전 BMI 제외; (2) PROM 포함, 임신 전 BMI 제외; (3) PROM 제외, 임신 전 BMI 포함; (4) PROM 포함, 임신 전 BMI 포함. 즉, SHAP 값은 기존 방법에서 무시되는 모든 가능한 하위 그룹 분석 결과를 통합한다. 이 연구의 한계는 표본 크기가 작다는 것이다. 일반적으로 기계 학습 기술은 명시적인 프로그래밍 없이 데이터를 구조화할 수 있으며 대규모 데이터 세트를 사용할 때 예측 성능이 향상되고 신뢰성이 높아진다. 그러나 본 연구에서는 상대적으로 작은 데이터 세트에서도 랜덤 포레스트(RF) 방법을 사용하여 PTB의 경우 73%(AUC), SPTB의 경우 86%(AUC)의 견고한 성능을 얻을 수 있었다. 이전 연구와 달리, 본 발명에서 데이터 세트는 전향적으로 수집되었으며, 치과, 의료, 산과 요인을 포함한 다각적인 매개변수가 포함되었으며, 이에 따라 높은 품질의 데이터로 인해 강력한 모델 구축이 가능하였다.
- [58] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크로컴퓨터, FPGA(Field Programmable Gate Array), PLU(Programmable Logic Unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다.
- [59] 처리 장치는 운영 체제 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술 분야에서 통상의 지식을 가진 자는 처리 장치가 복수 개의 처리 요소(Processing Element) 및/또는 복수 유형의 처리요소를 포함할 수 있음을 이해할 것이다.
- [60] 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor)와 같은, 다른 처리 구성(Processing configuration)도 가능하다. 소프트웨어는 컴퓨터 프로그램(Computer Program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다.
- [61] 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된

방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[62] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다.

[63] 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CDROM, DVD와 같은 광기록 매체(optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[64] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명에 개시된 실시예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

[65]

[66] [부호의 설명]

[67] 100 : 산모의 미숙아 분만 예측 장치

[68] 110 : 학습부

[69] 120 : 임상 데이터베이스

[70] 130 : 기계 학습 모델

[71] 140 : 데이터 입력부

[72] 150 : 미숙아 분만 예측부

청구범위

- [청구항 1] 산모의 임상 지표 데이터와 치과적 지표 데이터를 입력받는 데이터 입력부; 및
상기 임상 지표 데이터와 상기 치과적 지표 데이터를 기초로 기계 학습 모델- 상기 기계 학습 모델은 산모들에 대해 수집된 임상 지표 자료와, 치과적 검진 및 상기 산모들에 대해 영상 촬영을 통해 수집된 치주 지표 자료를 기초로 미숙아 분만을 예측하도록 학습됨 -에 의해 산모의 미숙아 분만을 예측하는 미숙아 분만 예측부를 포함하는, 산모의 미숙아 분만 예측 장치.
- [청구항 2] 청구항 1에 있어서,
상기 임상 지표 데이터는 산모 연령, 임신 전 체질량 지수, 인공 수정 여부, 과거 조산 병력, 전자간증, 만성 고혈압, 임신성 당뇨, 조기 양막 파수 및 융모양막염 단계 중의 적어도 둘 이상을 포함하는, 산모의 미숙아 분만 예측 장치.
- [청구항 3] 청구항 1에 있어서,
상기 치과적 지표 데이터는 치주염 단계, 변형된 치은 지수, 치태 지수 및 우식경험연구치지수 중의 적어도 둘 이상을 포함하는, 산모의 미숙아 분만 예측 장치.
- [청구항 4] 청구항 1에 있어서,
상기 기계 학습 모델은 랜덤 포레스트(Random forest) 모델을 기반으로 상기 임상 지표 데이터와 상기 치과적 지표 데이터의 변수 중요도를 예측하도록 학습된 모델인, 산모의 미숙아 분만 예측 장치.
- [청구항 5] 데이터 입력부에 의해, 산모의 임상 지표 데이터와 치과적 지표 데이터를 입력받는 단계; 및
미숙아 분만 예측부에 의해, 상기 임상 지표 데이터와 상기 치과적 지표 데이터를 기초로 기계 학습 모델- 상기 기계 학습 모델은 산모들에 대해 수집된 임상 지표 자료와, 치과적 검진 및 상기 산모들에 대해 영상 촬영을 통해 수집된 치주 지표 자료를 기초로 미숙아 분만을 예측하도록 학습됨 -에 의해 산모의 미숙아 분만을 예측하는 단계를 포함하는, 산모의 미숙아 분만 예측 방법.
- [청구항 6] 청구항 5에 있어서,
상기 임상 지표 데이터는 산모 연령, 임신 전 체질량 지수, 인공 수정 여부, 과거 조산 병력, 전자간증, 만성 고혈압, 임신성 당뇨, 조기 양막 파수 및 융모양막염 단계 중의 적어도 둘 이상을 포함하는, 산모의 미숙아 분만 예측 방법.
- [청구항 7] 청구항 5에 있어서,

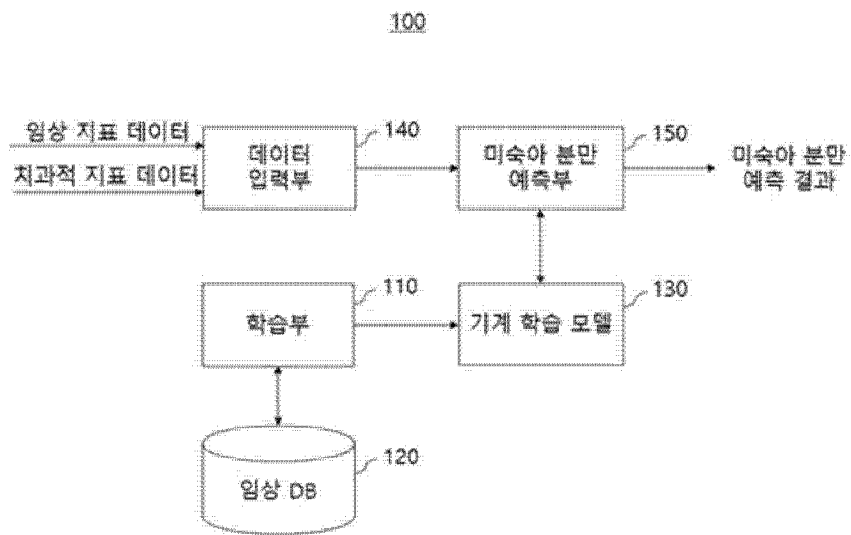
상기 치과적 지표 데이터는 치주염 단계, 변형된 치은 지수, 치태 지수 및 우식경험영구치지수 중의 적어도 둘 이상을 포함하는, 산모의 미숙아 분만 예측 방법.

[청구항 8] 청구항 5에 있어서,
상기 기계 학습 모델은 랜덤 포레스트(Random forest) 모델을 기반으로 상기 임상 지표 데이터와 상기 치과적 지표 데이터의 변수 중요도를 예측하도록 학습된 모델인, 산모의 미숙아 분만 예측 방법.

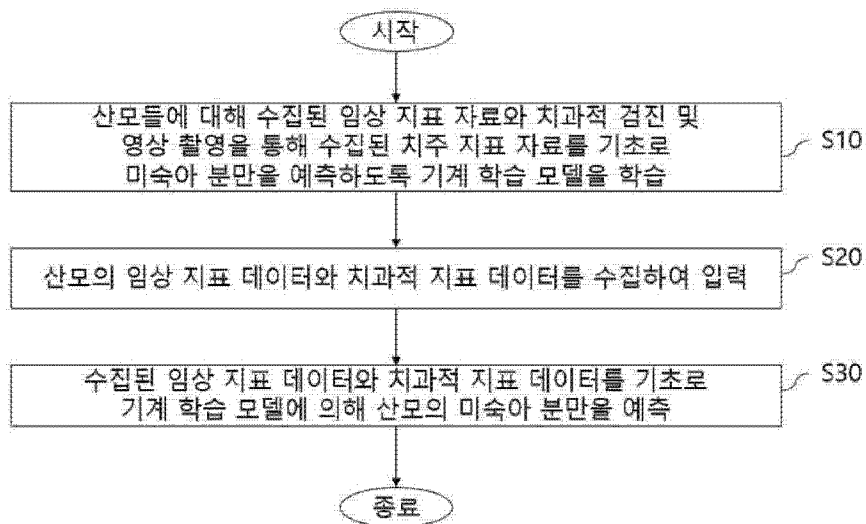
[청구항 9] 청구항 8에 있어서,
상기 산모의 미숙아 분만을 예측하는 단계는:
상기 랜덤 포레스트 모델에 의해, 조산과 관련하여, 임신 전 체질량 지수, 변형 치은 지수, 전자간증, 우식경험영구치 지수, 인공 수정 여부, 과거 조산 병력, 산모 연령, 만성 고혈압, 임신성 당뇨, 조기 양막 파수 및 융모양막염 단계의 변수 중요성 순위를 결정하여 조산 확률을 예측하는 단계를 포함하는, 산모의 미숙아 분만 예측 방법.

[청구항 10] 청구항 5의 미숙아 분만 예측 방법을 실행하기 위한 프로그램이 기록된 컴퓨터로 판독 가능한 비일시적 기록 매체.

[도1]



[도2]



[도3]

모델	전체 조산 vs. 만삭		자연 조산 vs. 만삭	
	Accuracy	AUC	Accuracy	AUC
LR	0.68	0.72	0.81	0.86
DT	0.64	0.64	0.76	0.78
NB	0.59	0.78	0.78	0.87
RF	0.65	0.73	0.83	0.86
SVM	0.47	0.40	0.66	0.62
ANN	0.58	0.63	0.70	0.71

LR, logistic regression; DT, decision tree; NB, naïve Bayes; RF, random forest; SVM, support vector machine; ANN, artificial neural network; AUC, area under the curve

[도4]

규칙 91,
10.01.2025

Variables	SHAP values					
	PTB vs. FTB			SPTB vs. FTB		
	Min	Max	Mean	Min	Max	Mean
Maternal age	-0.0805	0.0990	-0.0036	-0.0551	0.0756	0.0023
Pre-pregnancy BMI	-0.1640	0.1997	0.0071	-0.0572	0.1038	0.0042
In vitro fertilization	-0.0445	0.0191	0.0005	-0.0609	0.0272	0.0023
Prior preterm birth	-0.0173	0.0132	-0.0006	-0.0054	0.0350	-0.0006
Preeclampsia	-0.0578	0.1312	0.0000	-0.0214	0.0053	0.0012
Chronic hypertension	-0.0800	0.0091	0.0028	-0.0248	0.0077	0.0017
Gestational diabetes	-0.0350	0.0352	-0.0001	-0.0080	0.0020	0.0006
PROM	-0.0577	0.2256	-0.0029	-0.0989	0.3005	-0.0041
CAM	-0.0184	0.1057	-0.0030	-0.0266	0.1637	-0.0044
CAM stage	-0.0192	0.1176	-0.0041	-0.0300	0.1617	-0.0085
Periodontitis stage	-0.0268	0.0323	0.0010	-0.0189	0.0161	0.0025
MGI	-0.1190	0.1274	-0.0005	-0.0681	0.1306	0.0023
MGI ≥ 2	-0.0752	0.0889	-0.0032	-0.0192	0.0283	-0.0015
Plaque index ≥ 2	-0.0637	0.0620	0.0001	-0.0073	0.0119	-0.0007
DMFT-index	-0.0909	0.0837	-0.0004	-0.0839	0.1171	0.0024

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2024/017753

A. CLASSIFICATION OF SUBJECT MATTER**G16H 50/20**(2018.01)i; **G16H 10/60**(2018.01)i; **G16H 50/70**(2018.01)i; **A61B 5/00**(2006.01)i; **G06N 20/00**(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G16H 50/20(2018.01); A61B 5/00(2006.01); G16H 50/50(2018.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above

Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & keywords: 치주염(periodontitis), 조산(premature birth), 학습(learning), 랜덤 포레스트(random forest)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LEE, Kwang-Sig et al. Determinants of Spontaneous Preterm Labor and Birth Including Gastroesophageal Reflux Disease and Periodontitis. Journal of Korean Medical Science (JKMS). pp. 1-11, 03 March 2020. [Retrieved on 20 December 2024]. Retrieved from <jkms.org/DOIx.php?id=10.3346/jkms.2020.35.e105>. See abstract; and page 3.	1-10
A	최은정 등. 임신부의 치주 질환 활성도와 조산과의 상관관계에 관한 연구. 대한치주과학회지. vol. 30, no. 1, pp. 111-120, 31 March 2000 (CHOI, Eun-Jeong et al. Relationship between Preterm Low Birth Weight and Periodontal Disease Activity in Pregnancy. The journal of Korean academy of periodontology.) [Retrieved on 20 December 2024]. Retrieved from <jpis.org/pdf/10.5051/jkape.2000.30.1.111>. See pages 111-116.	1-10
A	KR 10-2023-0083447 A (KOREA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION) 12 June 2023 (2023-06-12) See claims 1-3; and figures 1-3.	1-10

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

14 February 2025

Date of mailing of the international search report

14 February 2025

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2023-0150145 A (EWHA UNIVERSITY - INDUSTRY COLLABORATION FOUNDATION et al.) 30 October 2023 (2023-10-30) See claims 1-10; and figures 1-5.	1-10
A	KR 10-2021-0016912 A (UNIST(ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY)) 17 February 2021 (2021-02-17) See claims 1-16; and figures 1-6.	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2024/017753

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
KR	10-2023-0083447	A	12 June 2023	US	2023-0178241	A1	08 June 2023
KR	10-2023-0150145	A	30 October 2023	None			
KR	10-2021-0016912	A	17 February 2021	KR	10-2288592	B1	11 August 2021

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G16H 50/20(2018.01)i; G16H 10/60(2018.01)i; G16H 50/70(2018.01)i; A61B 5/00(2006.01)i; G06N 20/00(2019.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G16H 50/20(2018.01); A61B 5/00(2006.01); G16H 50/50(2018.01)

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 치주염(periodontitis), 조산(premature birth), 학습(learning), 랜덤 포레스트(random forest)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	KWANG-SIG LEE 등, Determinants of Spontaneous Preterm Labor and Birth Including Gastroesophageal Reflux Disease and Periodontitis, Journal of Korean Medical Science (JKMS), pp. 1-11, 2020.03.03 [검색일: 2024-12(Dec)-20], 출처 <jkms.org/DOIx.php?id=10.3346/jkms.2020.35.e105> 초록; 및 페이지 3	1-10
A	최은정 등, 임신부의 치주 질환 활성화도와 조산과의 상관관계에 관한 연구, 대한치주과학회지, 30권, 1호, pp. 111-120, 2000.03.31 [검색일: 2024-12(Dec)-20], 출처 <jpis.org/pdf/10.5051/jkape.2000.30.1.111> 페이지 111-116	1-10
A	KR 10-2023-0083447 A (고려대학교 산학협력단) 2023.06.12 청구항 1-3; 및 도면 1-3	1-10
A	KR 10-2023-0150145 A (이화여자대학교 산학협력단 등) 2023.10.30 청구항 1-10; 및 도면 1-5	1-10
A	KR 10-2021-0016912 A (울산과학기술원) 2021.02.17 청구항 1-16; 및 도면 1-6	1-10

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“D” 본 국제출원에서 출원인이 인용한 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“&” 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일

2025년02월14일(14.02.2025)

국제조사보고서 발송일

2025년02월14일(14.02.2025)

ISA/KR의 명칭 및 우편주소

대한민국 특허청
(35208) 대전광역시 서구 청사로 189, 4동 (둔산동,
정부대전청사)

팩스 번호 +82-42-481-8578

심사관

이강하

전화번호 +82-42-481-5687

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2023-0083447 A	2023/06/12	US 2023-0178241 A1	2023/06/08
KR 10-2023-0150145 A	2023/10/30	없음	
KR 10-2021-0016912 A	2021/02/17	KR 10-2288592 B1	2021/08/11