

데이터 셋

헬스케어

1. 헬스케어 분류 데이터 셋
2. 헬스케어 회귀 데이터 셋

- Fetal Health Classification Dataset

- <https://www.kaggle.com/datasets/andrewmvd/fetal-health-classification>
- 유아 및 산모 사망을 예방하기 위한 태아 건강 상태 분류 데이터셋
- 건강 상태는 정상, 경계, 위독으로 분류됨
- 2126개의 데이터



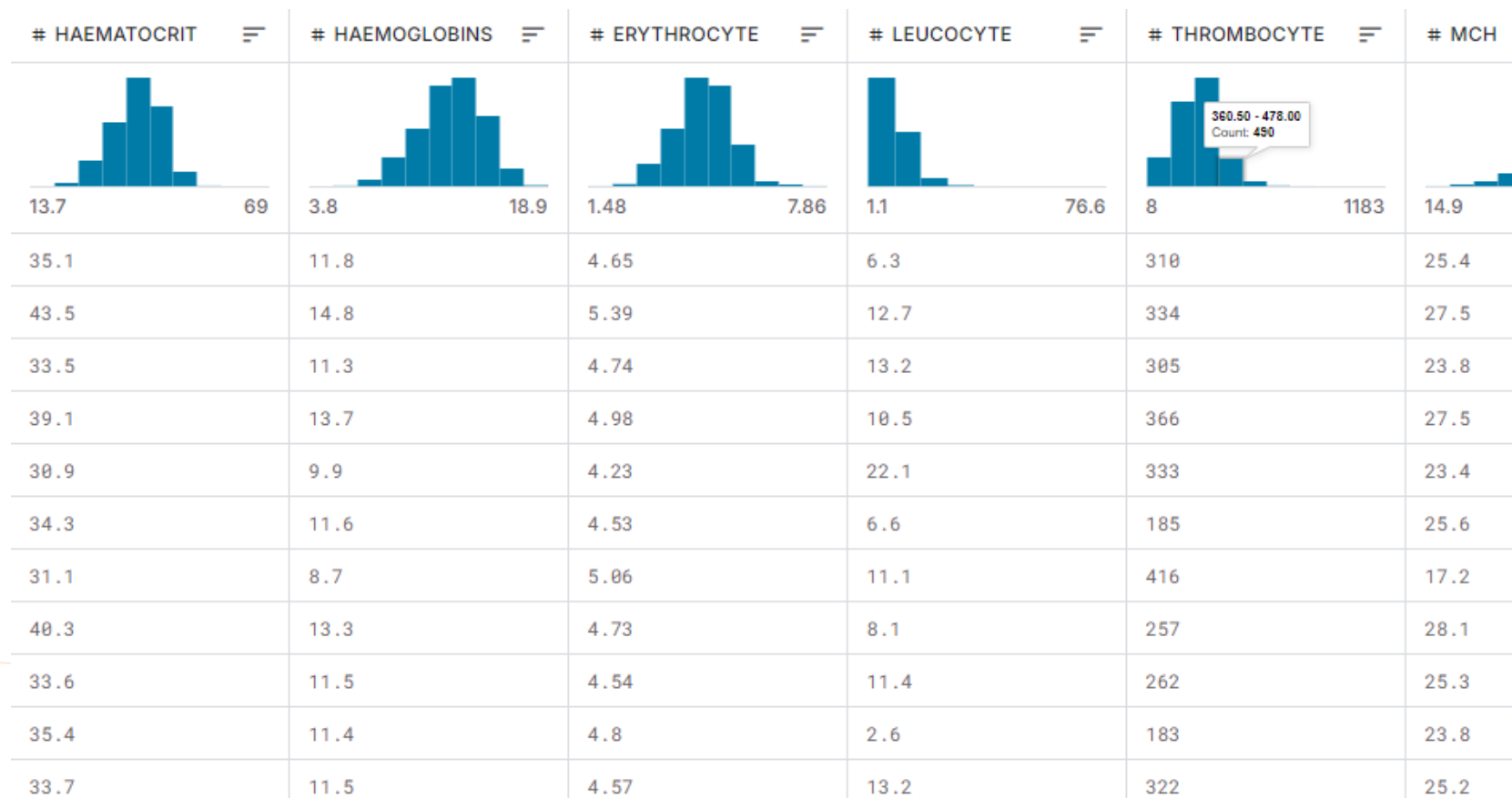
- 특징

- baseline value: 태아의 심박수 기저값 (bpm) (106 ~ 160)
- accelerations: 태아 심박수 가속 횟수 (0 ~ 0.02)
- decelerations: 태아 심박수 감속 횟수 (0 ~ 0.48)
- uterine_contractions: 자궁 수축 횟수 (0 ~ 0.01)
- light_decelerations: 가벼운 감속 (0 ~ 0.01)
- severe_decelerations: 심한 감속 (0 ~ 0)
- prolonged_decelerations: 장기 감속 (0 ~ 0.01)
- abnormal_short_term_variability: 비정상적인 단기 변동성 (12 ~ 87)
- mean_value_of_short_term_variability: 평균 단기 변동성 (0.2 ~ 7)
- percentage_of_time_with_abnormal_long_term_variability: 비정상적인 장기 변동성 시간의 백분율 (0 ~ 91)
- mean_value_of_long_term_variability: 장기 변동성 (0 ~ 50.7)

- 특징 (continue)
 - histogram_width: 태아 심박수 히스토그램 폭 (3 ~ 180)
 - histogram_min: 태아 심박수 히스토그램 최소값 (50 ~ 159)
 - histogram_max: 태아 심박수 히스토그램 최대값 (122 ~ 238)
 - histogram_number_of_peaks: 태아 심박수 히스토그램 피크 수 (0 ~ 18)
 - histogram_number_of_zeroes: 태아 심박수 히스토그램 제로 수 (0 ~ 10)
 - histogram_mode: 태아 심박수 히스토그램 모드 (60 ~ 187)
 - histogram_mean: 태아 심박수 히스토그램 평균 (73 ~ 182)
 - histogram_median: 태아 심박수 히스토그램 중앙값 (77 ~ 186)
 - histogram_variance: 태아 심박수 히스토그램 분산 (0 ~ 269)
 - histogram_tendency: 태아 심박수 히스토그램 경향성 (-1 ~ 1)
 - fetal_health: 태아 건강 (1, 2 ,3)

- Medical Cost Personal Datasets

- <https://www.kaggle.com/datasets/saurabhshahane/patient-treatment-classification/data?select=data-ori.csv>
- 인도네시아 병원에서 수집한 건강 기록 예측 데이터셋
- 4412개의 데이터



- 특징
 - HAEMATOCRIT: 적혈구 용적률 (13.7 ~ 69.0)
 - HAEMOGLOBINS: 혈색소 (3.8 ~ 18.9)
 - ERYTHROCYTE: 적혈구 수 (1.48 ~ 7.86)
 - LEUCOCYTE: 백혈구 수 (1.1 ~ 76.6)
 - THROMBOCYTE: 혈소판 수 (8 ~ 1183)
 - MCH: 평균 적혈구 혈색소량 (14.9 ~ 40.8)
 - MCHC: 평균 적혈구 혈색소 농도 (26 ~ 39)
 - MCV: 평균 적혈구 용적 (54 ~ 116)
 - AGE: 나이 (1 ~ 99)
 - SEX: 성별
 - SOURCE: 입원 필요 여부 (0, 1)

Thank you