**LAPORAN PROYEK SAINS DATA**

**“*Cervical Cancer Risk Classification*”**

**Dosen Pengampu : EKA MALA SARI ROCHMAN, S.Kom., M.Kom.**

**A picture containing text, sign

Description automatically generated**

Disusun oleh :

**MUHAMMAD MUQTAFIN NUHA**

210411100218

**TEKNIK INFORMATIKA FAKULTAS TEKNIK**

**UNIVERSITAS TRUNOJOYO MADURA**

**BANGKALAN**

**2023**

# PENDAHULUAN

## *Fetal Health Datasets*

Data *Fetal Health* ini berisi 2126 data yang diekstrak dari pemeriksaan Kardiotokogram. Rincian atribut atau fitur yang terdapat pada data *Fetal Health* yaitu:

Table 1 Atribut Fetal Health

| Atribut *Fetal Health* | | |
| --- | --- | --- |
| **No** | **Atribut** | **Deskripsi** |
| 1 | *Baseline value* | Denyut Jantung Janin Dasar (FHR) |
| 2 | *Accelerations* | Jumlah akselerasi per detik |
| 3 | *Fetal movement* | Jumlah gerakan janin per detik |
| 4 | *Uterine contractions* | Jumlah kontraksi rahim per detik |
| 5 | *Light decelerations* | Jumlah LDs per detik |
| 6 | *Severe decelerations* | Jumlah SDs per detik |
| 7 | *Prolongued decelerations* | Jumlah PDs per detik |
| 8 | *Abnormal short term variability* | Persentase waktu dengan variabilitas jangka pendek yang tidak normal |
| 9 | *Mean value of short term variability* | Nilai rata-rata variabilitas jangka pendek |
| 10 | *Percentage of time width abnormal long term variability* | Persentase waktu dengan variabilitas jangka panjang yang tidak normal |
| 11 | *Mean value of long term variability* | Nilai rata-rata variabilitas jangka panjang |
| 12 | *Histrogram width* | Lebar histogram yang dibuat dengan menggunakan semua nilai dari suatu catatan |
| 13 | *Histrogram min* | Histogram nilai minimum |
| 14 | *Histrogram max* | Histogram nilai maksimum |
| 15 | *Histrogram number of peaks* | Jumlah puncak dalam histogram ujian |
| 16 | *Histrogram number of zeroes* | Jumlah angka nol dalam histogram ujian |
| 17 | *Histrogram mode* | Mode histogram |
| 18 | *Histrogram mean* | Rata-rata histogram |
| 19 | *Histrogram median* | Histogram median |
| 20 | *Histrogram variance* | Variasi histogram |
| 21 | *Histrogram tendency* | Kecenderungan histogram |
| 22 | *Fetal health* | kesehatan janin  1 = Normal, 2 = Suspect, 3 = Pathological |

# ANALISIS DATA MENGGUNAKAN POWER BI

1. Preview atau Penyajian Data

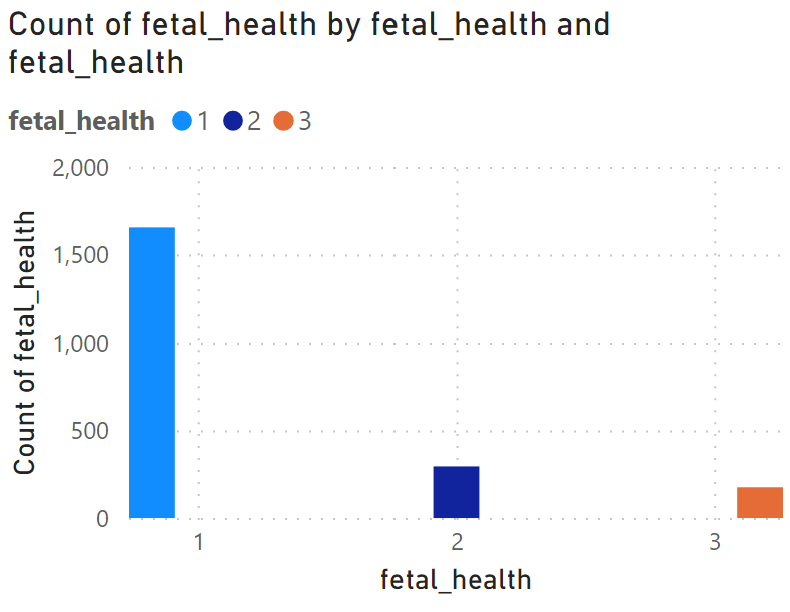
Data *Fetal Health* ini berisi 2126 data yang diekstrak dari pemeriksaan Kardiotokogram, Kardiotokogram (CTG) sendiri bekerja dengan mengirimkan denyut ultrasonik dan membaca responsnya, sehingga menjelaskan denyut jantung janin (FHR), gerakan janin, kontraksi rahim, dan banyak lagi. Berikut credit sumber lengkap Data *Fetal Health* dari Kaggle

1. Judul: *Fetal Health Classification*
2. Sumber: Ayres de Campos et al. (2000) SisPorto 2.0 A Program for Automated Analysis of Cardiotocograms. J Matern Fetal Med 5:311-318 ([link](https://onlinelibrary.wiley.com/doi/10.1002/1520-6661(200009/10)9:5%3C311::AID-MFM12%3E3.0.CO;2-9))
3. Download: [Fetal Health Classification](https://www.kaggle.com/datasets/andrewmvd/fetal-health-classification?datasetId=916586&sortBy=relevance)

Untuk table datasets *Fetal Health* dapat dilihat pada lampiran 1.

1. Analisis Persebaran atau Distribusi data

Untuk mengetahui persebaran data *Fetal Health* dilakukan pembuatan diagram batang antara Total *Fetal Health* dengan *Fetal Health*.



Picture 1 Count of fetal health by fetal health

Dari diagram *Fetal Health* diatas diperoleh ringkasan data sebagai berikut:

Table 2 Ringkasan Fetal Health (target)

|  |  |  |
| --- | --- | --- |
| No | *Fetal Health (target)* | Jumlah data |
| 1 | 1 / Normal | 1655 |
| 2 | 2 / Suspect | 295 |
| 3 | 3 / Pathological | 176 |

1. Analisis Paramater

Analisis parameter saya lakukan dengan melakukan beberapa pengecekan, diantaranya:

Table 3 Analisis Parameter

| **No** | **Parameter** | **Data Valid** | **Data Error** | **Data Empty** | **Data Unique** | **Data Distinct** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | *Baseline value* | 100% | 0% | 0% | 1 | 44 |
| 2 | *Accelerations* | 100% | 0% | 0% | 1 | 19 |
| 3 | *Fetal movement* | 100% | 0% | 0% | 53 | 97 |
| 4 | *Uterine contractions* | 100% | 0% | 0% | 0 | 14 |
| 5 | *Light decelerations* | 100% | 0% | 0% | 1 | 15 |
| 6 | *Severe decelerations* | 100% | 0% | 0% | 0 | 1 |
| 7 | *Prolongued decelerations* | 100% | 0% | 0% | 1 | 6 |
| 8 | *Abnormal short term variability* | 100% | 0% | 0% | 2 | 74 |
| 9 | *Mean value of short term variability* | 100% | 0% | 0% | 12 | 55 |
| 10 | *Percentage of time width abnormal long term variability* | 100% | 0% | 0% | 9 | 85 |
| 11 | *Mean value of long term variability* | 100% | 0% | 0% | 66 | 222 |
| 12 | *Histrogram width* | 100% | 0% | 0% | 5 | 146 |
| 13 | *Histrogram min* | 100% | 0% | 0% | 9 | 107 |
| 14 | *Histrogram max* | 100% | 0% | 0% | 2 | 74 |
| 15 | *Histrogram number of peaks* | 100% | 0% | 0% | 2 | 18 |
| 16 | *Histrogram number of zeroes* | 100% | 0% | 0% | 3 | 9 |
| 17 | *Histrogram mode* | 100% | 0% | 0% | 8 | 62 |
| 18 | *Histrogram mean* | 100% | 0% | 0% | 15 | 73 |
| 19 | *Histrogram median* | 100% | 0% | 0% | 17 | 72 |
| 20 | *Histrogram variance* | 100% | 0% | 0% | 34 | 96 |
| 21 | *Histrogram tendency* | 100% | 0% | 0% | 0 | 3 |
| 22 | *Fetal health* | 100% | 0% | 0% | 0 | 3 |

1. Transformasi Data

Untuk transformasi data saya menggunakan parameter *Fetal Health* yang bertipe data string atau teks untuk diubah menjadi angka menggunakan Pelabelan data atau Encoding data.

Table 4 Transformasi Data

|  |  |  |
| --- | --- | --- |
| No | *Fetal Health (target)* | Jumlah data |
| 1 | Normal | 1655 |
| 2 | Suspect | 295 |
| 3 | Pathological | 176 |

Didapatkan dengan menggunakan Pelabelan data atau Encoding data dengan menggunakan pengurutan dari jumlah data sebagai berikut:

Table 5 Hasil Transformasi Data

|  |  |  |  |
| --- | --- | --- | --- |
| No | *Fetal Health (target)* | Jumlah data | Nilai baru |
| 1 | Normal | 1655 | 1 |
| 2 | Suspect | 295 | 2 |
| 3 | Pathological | 176 | 3 |

1. Analisis Missing Value

Seperti yang ditampilkan pada Tabel 3 Analisis Paramater, didapatkan bahwa tidak ada Data Empty atau Missing Value pada setiap parameter yang ada.

# LAMPIRAN 1

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | *Baseline value* | *Accelerations* | *Fetal movement* | *Uterine contractions* | *Light decelerations* | *Severe decelerations* | *Prolongued decelerations* | *Abnormal short term variability* | *Mean value of short term variability* | *Percentage of time width abnormal long term variability* | *Mean value of long term variability* |
| 1 | 120 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 0,5 | 43 | 2,4 |
| 2 | 132 | 0,006 | 0 | 0,006 | 0,003 | 0 | 0 | 17 | 2,1 | 0 | 10,4 |
| 3 | 133 | 0,003 | 0 | 0,008 | 0,003 | 0 | 0 | 16 | 2,1 | 0 | 13,4 |
| 4 | 134 | 0,003 | 0 | 0,008 | 0,003 | 0 | 0 | 16 | 2,4 | 0 | 23 |
| 5 | 132 | 0,007 | 0 | 0,008 | 0 | 0 | 0 | 16 | 2,4 | 0 | 19,9 |
| 6 | 134 | 0,001 | 0 | 0,01 | 0,009 | 0 | 0,002 | 26 | 5,9 | 0 | 0 |
| 7 | 134 | 0,001 | 0 | 0,013 | 0,008 | 0 | 0,003 | 29 | 6,3 | 0 | 0 |
| 8 | 122 | 0 | 0 | 0 | 0 | 0 | 0 | 83 | 0,5 | 6 | 15,6 |
| 9 | 122 | 0 | 0 | 0,002 | 0 | 0 | 0 | 84 | 0,5 | 5 | 13,6 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | *Histrogram width* | *Histrogram min* | *Histrogram max* | *Histrogram number of peaks* | *Histrogram number of zeroes* | *Histrogram mode* | *Histrogram mean* | *Histrogram median* | *Histrogram variance* | *Histrogram tendency* | *Fetal health* |
| 1 | 64 | 62 | 126 | 2 | 0 | 120 | 137 | 121 | 73 | 1 | 2 |
| 2 | 130 | 68 | 198 | 6 | 1 | 141 | 136 | 140 | 12 | 0 | 1 |
| 3 | 130 | 68 | 198 | 5 | 1 | 141 | 135 | 138 | 13 | 0 | 1 |
| 4 | 117 | 53 | 170 | 11 | 0 | 137 | 134 | 137 | 13 | 1 | 1 |
| 5 | 117 | 53 | 170 | 9 | 0 | 137 | 136 | 138 | 11 | 1 | 1 |
| 6 | 150 | 50 | 200 | 5 | 3 | 76 | 107 | 107 | 170 | 0 | 3 |
| 7 | 150 | 50 | 200 | 6 | 3 | 71 | 107 | 106 | 215 | 0 | 3 |
| 8 | 68 | 62 | 130 | 0 | 0 | 122 | 122 | 123 | 3 | 1 | 3 |
| 9 | 68 | 62 | 130 | 0 | 0 | 122 | 122 | 123 | 3 | 1 | 3 |