| **Feature name** | **Mean ± SD** | **Diabetes (*N* = 268)** | **Nondiabetes (*N* = 500)** | ***P* value** |
| --- | --- | --- | --- | --- |
| Age (years) | 33.24 ± 11.8 |  |  |  |
| (<25) |  | 31 (11.8%) | 188 (38.0%) | <0.05 |
| (25–30) |  | 53 (20.0%) | 124 (25.2%) |
| (30–35) |  | 42 (15.5%) | 50 (9.6%) |
| (35–40) |  | 34 (12.9%) | 39 (8.3%) |
| (>40) |  | 108 (40.0%) | 99 (19.9%) |
|  | | | | |
| No. of pregnancies | 3.8 ± 3.4 |  |  |  |
| Never |  | 38 (14.0%) | 73 (14.0%) | <0.05 |
| (1–3) |  | 75 (28.2%) | 238 (45.6%) |
| (4–6) |  | 60 (22.0%) | 115 (25.7%) |
| (>6) |  | 95 (35.8%) | 74 (14.7%) |
|  | | | | |
| Insulin level (u U/ml) | 79.8 ± 15.2 |  |  |  |
| (Less than 200) |  | 221 (83.0%) | 458 (92.0%) | <0.05 |
| (More than 200) |  | 47 (18.0%) | 42 (8.0%) |
|  | | | | |
| BMI (kg/m^2) | 31.9 ± 7.9 |  |  |  |
| Normal |  | 7 (3.0%) | 108 (22.0%) | <0.05 |
| Overweight |  | 122 (46.0%) | 239 (46.9%) |
| Obesity |  | 139 (52.0%) | 153 (31.1%) |
|  | | | | |
| Blood pressure (mm·Hg) | 69 ± 19.4 |  |  |  |
| Low (<65) |  | 46 (18.0%) | 155 (31.4%) | <0.05 |
| Normal (65–85) |  | 173 (64.6%) | 288 (57.2%) |
| High (>85) |  | 49 (17.4%) | 57 (11.4%) |
|  | | | | |
| Glucose (mg/dL) | 120.9 ± 32 |  |  |  |
| Normal (<140) |  | 131 (49.9%) | 438 (88.6%) | <0.05 |
| High (>140) |  | 137 (50.1%) | 62 (11.4%) |
|  | | | | |
| Skin fold (mm) | 20.5 ± 16 |  |  |  |
| (<20) |  | 103 (38.6%) | 235 (47.4%) | <0.05 |
| (20–40) |  | 123 (45.9%) | 217 (43.0%) |
| (>40) |  | 42 (15.5%) | 48 (9.6%) |
|  | | | | |
| Pedigree function | 0.47 ± 0.3 |  |  |  |
| (<0.5) |  | 163 (60.8%) | 319 (63.8%) | <0.05 |
| (0.5–1.0) |  | 87 (32.5%) | 145 (29.0%) |
| (>1.0) |  | 18 (6.7%) | 36 (7.2%) |

**Dan juga penambahan fitur usia berdasarkan range usia**

Berikut adalah poin-poin yang menjelaskan bagaimana setiap fitur dalam Pima Indian Diabetes Dataset memengaruhi risiko diabetes, berdasarkan jurnal yang telah diunggah:

1. **Pregnancies (Jumlah Kehamilan)**: Jumlah kehamilan berkaitan dengan peningkatan risiko diabetes karena selama kehamilan terjadi perubahan metabolisme dan resistensi insulin. Setiap kehamilan dapat menambah stres metabolik pada tubuh, sehingga semakin banyak kehamilan, semakin tinggi risiko mengalami diabetes, terutama jika mengalami diabetes gestasional.
2. **Glucose (Kadar Glukosa)**: Tingkat glukosa puasa dan setelah makan adalah indikator langsung dari diabetes. Kadar glukosa yang tinggi menunjukkan gangguan dalam pemrosesan gula darah oleh tubuh, yang biasanya terjadi akibat resistensi insulin atau ketidakmampuan tubuh memproduksi insulin dalam jumlah yang cukup.
3. **Blood Pressure (Tekanan Darah)**: Hipertensi atau tekanan darah tinggi sering ditemukan bersamaan dengan resistensi insulin, kondisi umum pada penderita diabetes tipe 2. Hipertensi memperburuk risiko komplikasi vaskular yang dapat memperparah kondisi diabetes.
4. **Skin Thickness (Ketebalan Lipatan Kulit)**: Ukuran ketebalan lipatan kulit adalah indikator dari kadar lemak tubuh. Lemak tubuh yang berlebihan berkaitan erat dengan obesitas, yang merupakan faktor risiko utama untuk resistensi insulin dan diabetes.
5. **Insulin**: Tingkat insulin dalam darah mencerminkan kemampuan tubuh dalam memproduksi dan merespons insulin. Kadar insulin yang tidak normal, baik terlalu rendah atau terlalu tinggi, menunjukkan gangguan pada metabolisme glukosa dan bisa menjadi tanda awal diabetes.
6. **BMI (Body Mass Index)**: BMI yang tinggi menunjukkan obesitas, yang merupakan salah satu faktor risiko utama untuk diabetes tipe 2. Obesitas meningkatkan resistensi insulin sehingga tubuh kesulitan mengatur kadar glukosa dengan baik.
7. **Diabetes Pedigree Function**: Fitur ini menunjukkan predisposisi genetik untuk diabetes. Riwayat keluarga yang memiliki diabetes meningkatkan risiko individu karena adanya faktor keturunan yang memengaruhi fungsi insulin atau metabolisme glukosa.
8. **Age (Usia)**: Risiko diabetes meningkat seiring bertambahnya usia. Peningkatan usia membawa perubahan metabolisme dan menurunkan kemampuan tubuh untuk merespons insulin, sehingga meningkatkan risiko diabetes pada orang yang lebih tua.

Setiap fitur ini memainkan peran penting dalam memperkirakan risiko diabetes dengan memberikan gambaran menyeluruh tentang faktor-faktor genetik, fisiologis, dan gaya hidup yang terkait dengan diabetes.

**Tambahan identifikasi Masalah**

**Tambahan use case diagram dan activity diagram**

**Penguji 1**