

Informasi Kolom dalam Dataset Diabetes

1. **Pregnancies (Kehamilan):** Kolom ini mencatat jumlah kehamilan yang pernah dialami oleh individu yang terkait dengan data tersebut. Angka nol (0) dalam kolom ini menunjukkan individu tersebut belum pernah hamil, sementara angka lain menunjukkan jumlah kehamilan yang dialami.
2. **Glucose (Glukosa):** Kolom ini menunjukkan kadar glukosa (gula) dalam darah individu dalam satuan mg/dL (miligram per desiliter). Glukosa adalah sumber utama energi bagi tubuh, dan kadar glukosa yang tinggi dapat menunjukkan adanya masalah dengan metabolisme glukosa, seperti resistensi insulin atau diabetes.
3. **BloodPressure (Tekanan Darah):** Kolom ini mencatat tekanan darah sistolik (angka pertama saat mengukur tekanan darah) individu dalam mm Hg (milimeter air raksa). Tekanan darah tinggi (hipertensi) dapat menjadi faktor risiko untuk perkembangan diabetes tipe 2.
4. **SkinThickness (Ketebalan Kulit):** Kolom ini menggambarkan ketebalan lipatan kulit (skinfold) pada trisep (lengan atas) individu dalam satuan mm (milimeter). Pengukuran ketebalan kulit ini dapat memberikan indikasi tentang persentase lemak tubuh dan dapat berhubungan dengan resistensi insulin.
5. **Insulin:** Kolom ini mencatat kadar insulin dalam darah individu dalam satuan $\mu\text{U/mL}$ (mikro unit per mililiter). Insulin adalah hormon yang diproduksi oleh pankreas dan berperan dalam mengatur metabolisme glukosa. Kadar insulin yang rendah atau tinggi dapat menjadi indikasi masalah dalam produksi atau respons insulin, yang berhubungan dengan diabetes.
6. **BMI (Body Mass Index):** Kolom ini menggambarkan Indeks Massa Tubuh (BMI) individu, yang dihitung dengan membagi berat badan dalam kilogram dengan kuadrat tinggi badan dalam meter. BMI digunakan sebagai indikator untuk mengevaluasi apakah individu memiliki berat badan yang sehat. Individu dengan BMI tinggi, terutama dengan penumpukan lemak di sekitar perut, memiliki risiko lebih tinggi untuk mengembangkan diabetes tipe 2.
7. **DiabetesPedigreeFunction:** Kolom ini mencerminkan riwayat keluarga diabetes individu dengan memperhitungkan hubungan genetik antara individu tersebut dengan anggota keluarga lain yang memiliki riwayat diabetes. Skor dalam kolom ini dihitung berdasarkan kompleksitas riwayat keluarga dan memberikan perkiraan risiko genetik individu untuk mengembangkan diabetes.
8. **Age (Usia):** Kolom ini mencatat usia individu dalam tahun. Usia adalah faktor penting dalam risiko terkena diabetes, di mana risiko umumnya meningkat seiring bertambahnya usia. Proses penuaan dan perubahan gaya hidup seiring bertambahnya usia dapat mempengaruhi risiko diabetes.
9. **Outcome:** Kolom ini menunjukkan apakah individu tersebut memiliki diabetes atau tidak. Nilai 1 menunjukkan bahwa individu tersebut positif diabetes, sedangkan nilai 0 menunjukkan individu tersebut tidak memiliki diabetes. Kolom ini merupakan

target yang ingin diprediksi menggunakan model machine learning untuk memprediksi apakah individu memiliki diabetes atau tidak berdasarkan fitur-fitur lainnya.