

**Penerapan *Business Intelligence* dalam Analisis Dataset untuk  
Memprediksi Tingkat Obesitas berdasarkan Pola Makan dan Kondisi Fisik  
pada Penduduk Kolombia, Peru, dan Meksiko**



Disusun oleh:

Marta Yulianti / 215314154

Victoria Alysha Fernando S. / 215314158

**Program Studi Informatika  
Fakultas Sains dan Teknologi  
Universitas Sanata Dharma  
Yogyakarta  
2024**

## **A. LATAR BELAKANG**

Gaya hidup menjadi kunci dalam menentukan berat badan seseorang. Pola makan yang tidak seimbang, sementara kurang bergerak secara fisik, dapat menyebabkan penimbunan lemak dan akhirnya mengarah pada obesitas. Faktor lain seperti genetika, tingkat stres, dan kondisi medis juga dapat mempengaruhi berat badan, namun, pola makan dan aktivitas fisik tetap menjadi faktor utama yang dapat diubah.

Obesitas adalah kondisi medis yang serius dan kompleks yang terjadi ketika terlalu banyak lemak disimpan dalam tubuh, melebihi jumlah yang dianggap sehat untuk kesehatan. Obesitas dapat meningkatkan risiko terjadinya berbagai penyakit serius. Selain itu, obesitas juga dapat mengganggu kualitas hidup seseorang dan memperpendek masa hidupnya.

Untuk menggali sejauh mana masalah obesitas di berbagai negara, data statistik tentang estimasi tingkat obesitas pada individu dari negara-negara seperti Meksiko, Peru, dan Kolombia dapat memberikan wawasan yang berharga. Dengan menganalisis pola makan dan kondisi fisik populasi dari negara-negara ini, kita dapat memperoleh pemahaman yang lebih komprehensif tentang faktor-faktor yang memengaruhi tingkat obesitas dalam masyarakat tersebut. Informasi ini dapat menjadi landasan untuk merancang langkah-langkah intervensi yang tepat guna mengatasi masalah obesitas dan meningkatkan kesehatan masyarakat secara keseluruhan. Selain itu, data ini juga dapat menjadi referensi bagi individu-individu yang ingin memprediksi level obesitas mereka sendiri berdasarkan kebiasaan makan dan kondisi fisik.

## **B. RUMUSAN MASALAH**

Berdasarkan penjelasan latar belakang yang telah disampaikan, rumusan masalah dapat dirumuskan sebagai berikut:

- a. Apa saja faktor-faktor yang berkontribusi signifikan terhadap tingkat obesitas, khususnya dalam konteks pola makan dan kondisi fisik individu dari negara-negara seperti Meksiko, Peru, dan Kolombia?
- b. Apakah data statistik tentang pola makan dan kondisi fisik populasi dari negara-negara yang disebutkan dapat dijadikan dasar untuk prediksi tingkat obesitas di masa depan?

### **C. TUJUAN**

Berikut adalah tujuan yang ingin dicapai:

- a. Mengidentifikasi faktor-faktor utama yang memiliki kontribusi signifikan terhadap tingkat obesitas.
- b. Membuat dataset tentang estimasi tingkat obesitas pada individu dari negara-negara seperti Meksiko, Peru, dan Kolombia sebagai referensi yang dapat digunakan untuk prediksi tingkat obesitas di masa depan.

### **D. PEMBAHASAN**

Dataset ini memberikan gambaran tentang tingkat obesitas di Meksiko, Peru, dan Kolombia, menjangkau berbagai rentang usia mulai dari 14 hingga 61 tahun, serta menggali beragam kebiasaan makan dan kondisi fisik yang dimiliki oleh responden. Melalui platform *web*, jawaban responden melalui proses *pre-processing*, menghasilkan 17 atribut yang relevan dan 2111 jawaban (Palechor & Manotas, 2019).

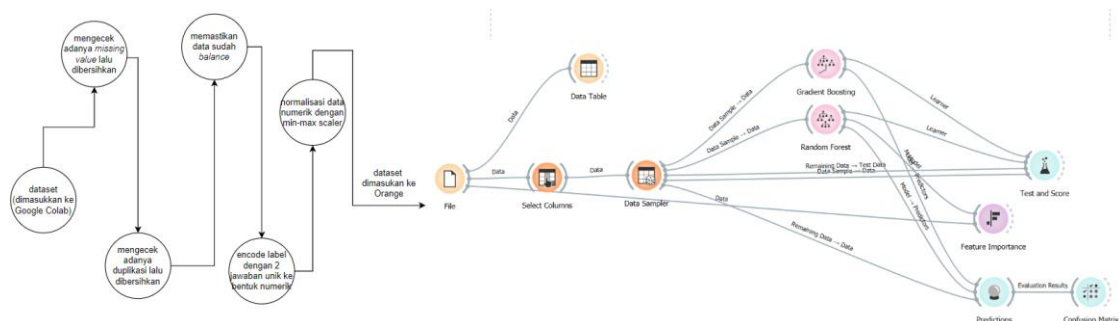
Analisis deskriptif menjadi kunci untuk mengungkap faktor-faktor yang signifikan dalam kontribusinya terhadap tingkat obesitas. Informasi ini bukan hanya berharga bagi lembaga kesehatan dan pemerintah, tetapi juga menjadi landasan strategis

dalam merumuskan langkah-langkah konkret untuk menanggulangi masalah obesitas di ketiga negara tersebut.

Tidak hanya itu, analisis prediktif membantu memprediksi tingkat obesitas berdasarkan pola makan dan fisik (tanpa perlu mengetahui berat dan tinggi). Dengan menerapkan teknik klasifikasi, kita dapat membandingkan kinerja algoritma seperti *Gradient Boosting* dan *Random Forest*. Evaluasi akurasi akan membantu menentukan algoritma yang paling sesuai untuk menganalisis dataset ini. Tingkat akurasi yang tinggi dalam teknik klasifikasi akan menjadi dasar yang kuat untuk pengembangan aplikasi yang dilakukan lembaga kesehatan ataupun pemerintah, untuk secara otomatis mengklasifikasikan tingkat obesitas, dengan dataset ini sebagai acuan atau *training set*. Aplikasi tersebut juga memudahkan warga untuk mengetahui tingkat berat badan mereka. Hal ini sangat membantu bagi mereka yang kesulitan mengakses informasi kesehatan yang relevan.

Dengan demikian, dataset ini tidak hanya menjadi sumber informasi, tetapi juga menjadi landasan untuk mengembangkan solusi konkret yang dapat membantu menangani masalah obesitas secara efektif di Meksiko, Peru, dan Kolombia.

Pengolahan dan menganalisa dataset dilakukan menggunakan 2 aplikasi, yaitu *Google Colab*, dan *Orange*. Berikut diagramnya :



Langkah-kangkah metodologis untuk menerapkan konsep deskriptif serta prediktif dapat dijelaskan sebagai berikut:

## 1. Pengumpulan Data

Data diperoleh dari halaman “UC Irvine Machine Learning Repository”.

Berikut tampilan dataset:

|      | Gender | Age       | Height   | Weight     | family_history_with_overweight | FAVC | FCVC | NCP | CAEC      | SMOKE | CHD      | SCC | FAP      | TUE      | CALC       | NTTRANS               | NOBeyesdad          |
|------|--------|-----------|----------|------------|--------------------------------|------|------|-----|-----------|-------|----------|-----|----------|----------|------------|-----------------------|---------------------|
| 0    | Female | 21.000000 | 1.620000 | 64.000000  | yes                            | no   | 2.0  | 3.0 | Sometimes | no    | 2.000000 | no  | 0.000000 | 1.000000 | no         | Public_Transportation | Normal_Weight       |
| 1    | Female | 21.000000 | 1.520000 | 56.000000  | yes                            | no   | 3.0  | 3.0 | Sometimes | yes   | 3.000000 | yes | 3.000000 | 0.000000 | Sometimes  | Public_Transportation | Normal_Weight       |
| 2    | Male   | 23.000000 | 1.800000 | 77.000000  | yes                            | no   | 2.0  | 3.0 | Sometimes | no    | 2.000000 | no  | 2.000000 | 1.000000 | Frequently | Public_Transportation | Normal_Weight       |
| 3    | Male   | 27.000000 | 1.800000 | 87.000000  | no                             | no   | 3.0  | 3.0 | Sometimes | no    | 2.000000 | no  | 2.000000 | 0.000000 | Frequently | Walking               | Overweight_Level_I  |
| 4    | Male   | 22.000000 | 1.780000 | 89.800000  | no                             | no   | 2.0  | 1.0 | Sometimes | no    | 2.000000 | no  | 0.000000 | 0.000000 | Sometimes  | Public_Transportation | Overweight_Level_II |
| ...  | ...    | ...       | ...      | ...        | ...                            | ...  | ...  | ... | ...       | ...   | ...      | ... | ...      | ...      | ...        | ...                   | ...                 |
| 2106 | Female | 20.976842 | 1.710730 | 131.408528 | yes                            | yes  | 3.0  | 3.0 | Sometimes | no    | 1.728139 | no  | 1.676269 | 0.906247 | Sometimes  | Public_Transportation | Obesity_Type_III    |
| 2107 | Female | 21.982942 | 1.748584 | 133.742943 | yes                            | yes  | 3.0  | 3.0 | Sometimes | no    | 2.005130 | no  | 1.341390 | 0.599270 | Sometimes  | Public_Transportation | Obesity_Type_III    |
| 2108 | Female | 22.524036 | 1.752206 | 133.689352 | yes                            | yes  | 3.0  | 3.0 | Sometimes | no    | 2.054193 | no  | 1.414209 | 0.646288 | Sometimes  | Public_Transportation | Obesity_Type_III    |
| 2109 | Female | 24.361936 | 1.739450 | 133.346641 | yes                            | yes  | 3.0  | 3.0 | Sometimes | no    | 2.852339 | no  | 1.139107 | 0.586035 | Sometimes  | Public_Transportation | Obesity_Type_III    |
| 2110 | Female | 23.664709 | 1.738836 | 133.472641 | yes                            | yes  | 3.0  | 3.0 | Sometimes | no    | 2.863513 | no  | 1.026452 | 0.714137 | Sometimes  | Public_Transportation | Obesity_Type_III    |

## 2. Pemahaman Data

Dataset ini berisi informasi yang digunakan untuk memperkirakan tingkat obesitas individu di Meksiko, Peru, dan Kolombia, berdasarkan pola makan dan kondisi fisik. Menurut (Palechor & Manotas, 2019), dataset ini memberikan label tingkat obesitas berdasarkan perhitungan *Body Mass Index* (BMI) sebagai berikut :

$$BMI = \frac{weight}{height * height}$$

Hasil BMI kemudian disesuaikan dengan standar WHO dan norma-norma kesehatan Meksiko, menghasilkan atribut "NOBeyesdad", yang mencakup kategori tingkat obesitas, mulai dari berat kurang, berat normal, *overweight* tingkat I, *overweight* tingkat II, hingga obesitas tipe I, II, dan III.

Berikut tabel yang memuat deskripsi dari atribut-atribut dataset :

| Atribut | Deskripsi / Pertanyaan | Pilihan Jawaban              |
|---------|------------------------|------------------------------|
| Gender  | Jenis kelamin          | Perempuan<br>Laki - laki     |
| Age     | Usia                   | Nilai numerik                |
| Height  | Tinggi badan           | Nilai numerik dalam meter    |
| Weight  | Berat badan            | Nilai numerik dalam kilogram |

|                                    |  |   |
|------------------------------------|--|---|
| family_history<br>_with_overweight | Apakah ada anggota keluarga yang pernah menderita atau sedang menderita kelebihan berat badan?                     | Iya<br>Tidak  |
| FAVC                               | Apakah Anda sering mengonsumsi makanan berkalori tinggi?   | Iya<br>Tidak  |
| FCVC                               | Apakah Anda biasanya makan sayuran dalam makanan Anda?   | Tidak pernah<br>Kadang - kadang   |
| NCP                                | Berapa kali makan utama yang Anda miliki setiap hari?  | Selalu<br>Diantara 1 atau 2<br>3  |
| CAEC                               | Apakah Anda makan makanan di antara waktu makan?   | Tidak<br>Kadang – kadang<br>Sering<br>Selalu                                    |
| SMOKE                              | Apakah Anda merokok?   | Iya<br>Tidak  |
| CH2O                               | Berapa banyak air yang Anda minum setiap hari?   | Kurang dari satu liter<br>Diantara 1 dan 2 L<br>Lebih dari 2 L                  |
| SCC                                | Apakah Anda memantau kalori yang Anda konsumsi setiap hari?  | Yes<br>Tidak  |
| FAF                                | Berapa sering Anda melakukan aktivitas fisik?  | 0 = Tidak pernah<br>1 = 1 atau 2 hari<br>2 = 2 atau 4 hari<br>4 = 4 atau 5 hari |
| TUE                                | Berapa lama Anda menggunakan perangkat teknologi seperti ponsel, permainan video, televisi, komputer, dan lainnya? | 0 = 0 – 2 jam<br>1 = 3 – 5 jam<br>2 = Lebih dari 5 jam                          |
| CALC                               | Berapa sering Anda minum alkohol?  | Tidak minum<br>Kadang – kadang  |

|            |  |  |
|------------|--|--|
|            |  | Sering<br>Selalu   |
| MTRANS     | Transportasi apa yang biasanya Anda gunakan? | Mobil<br>Sepeda motor<br>Sepeda<br>Kendaraan umum<br>Jalan kaki  |
| NObeyesdad | Tingkat obesitas?                            | Insufficient_Weight<br>Normal_Weight<br>Obesity_Type_I<br>Obesity_Type_II<br>Obesity_Type_III<br>Overweight_Level_I<br>Overweight_Level_II |

### 3. *Preprocessing Data*

Proses ini mencakup pembersihan data dari *missing value* dan duplikat. Setelah memastikan dataset terbebas dari nilai yang hilang dan duplikat, langkah selanjutnya adalah mentransformasi atribut kategorikal yang memiliki dua tipe jawaban unik, seperti "pria-wanita" atau "iya-tidak", menggunakan *Label Encoder* untuk mengubahnya menjadi representasi numerik. Kemudian, dilakukan proses normalisasi untuk data numerik menggunakan metode *min-max scaler*. Proses ini penting untuk memastikan bahwa semua fitur memiliki skala yang seragam, sehingga mencegah dominasi fitur tertentu dalam analisis. Tahap pembersihan data hingga normalisasi data dilakukan di aplikasi *google colab*. Setelah itu, dataset (yang baru), dimasukkan ke dalam *Orange*. Karena data *weight* dan *height* hanya digunakan untuk memberi label, dan model akan dilatih untuk memperkirakan tingkat obesitas berdasarkan pola makan dan kondisi fisik, kami menggunakan "*Select Columns*" untuk mengabaikan kolom

*weight* dan *height*. Menggunakan “*Data sampler*”, data dipisahkan menjadi dua subset utama: *training set* (80%) dan *testing set* (20%). Atribut "NObytesdad" dipilih sebagai target dalam proses klasifikasi.

#### 4. Pelatihan Model

Algoritma yang digunakan adalah *Random Forest* dan *Gradient Boosting* (XGBoost). XGBoost merupakan model linier yang sangat efektif dalam menyelesaikan masalah linier. Selain itu, XGBoost juga mendukung beragam proses objektif seperti klasifikasi, regresi, dan peringkat (Christanto et al., 2023). *Random Forest* bekerja dengan cara meningkatkan akurasi klasifikasi melalui penggabungan (aggregation) hasil klasifikasi dari sekelompok pohon keputusan. Dalam *Random Forest*, data diproses dengan menggunakan sejumlah *decision tree* dan proses seleksi. Ketika melakukan klasifikasi dengan *decision tree*, dataset dibagi berdasarkan fitur-fitur yang ada. *Decision tree* membuat prediksi secara acak, yang menghasilkan hasil yang optimal. (Mursianto Ghaita Amany et al., 2021). Pada *Orange*, *gradient boosting* dan *random forest* dihubungkan dengan “*Data Sampler*”, untuk melakukan *training* dengan jumlah yang sudah ditentukan. Parameter yang digunakan yaitu parameter *default* dari aplikasi *Orange*.

#### 5. Evaluasi

Evaluasi model dilakukan untuk menilai seberapa baik kemampuan model XGBoost dan *Random Forest* dalam memprediksi nilai target data uji dengan akurasi yang tinggi. Langkah awalnya adalah menguji kedua model tersebut menggunakan data uji sebagai masukan (Intan Permata & Esther Sorta Mauli Nababan, 2023). Evaluasi disajikan dalam metrik evaluasi, dengan deskripsi setiap penilaian sebagai berikut :



1. AUC menunjukkan seberapa baik model dapat membedakan kelas positif dan negatif. Semakin mendekati 1, model semakin baik.
2. Akurasi klasifikasi menghitung persentase prediksi yang benar dari keseluruhan prediksi.
3. F1 score memberi keseimbangan antara presisi dan recall, dengan nilai maksimal 1.
4. Precision menunjukkan proporsi positif yang benar dari seluruh hasil positif yang diprediksi model.
5. Recall menunjukkan proporsi positif yang benar dari seluruh data positif sebenarnya.
6. MCC mengukur hubungan antara prediksi dan hasil aktual, antara -1 hingga 1. Semakin mendekati 1, semakin sesuai.

Pada *Orange*, *widget* untuk evaluasi kinerja model itu sendiri yaitu “*Test and Score*” dihubungkan dengan algoritma dan “*Data Sampler*” untuk melakukan pengujian pada data *testing* yang sudah ditentukan jumlahnya. Untuk mengevaluasi algoritma untuk prediksi, dihubungkan ke “*Predictions*”. Berikut hasil evaluasi model dalam melakukan klasifikasi :

| Model             | AUC   | CA    | F1    | Prec  | Recall | MCC   |
|-------------------|-------|-------|-------|-------|--------|-------|
| Random Forest     | 0.973 | 0.851 | 0.852 | 0.854 | 0.851  | 0.825 |
| Gradient Boosting | 0.974 | 0.832 | 0.830 | 0.831 | 0.832  | 0.803 |

Hasil evaluasi model untuk prediksi :

| Model             | AUC   | CA    | F1    | Prec  | Recall | MCC   |
|-------------------|-------|-------|-------|-------|--------|-------|
| Gradient Boosting | 0.974 | 0.832 | 0.830 | 0.831 | 0.832  | 0.803 |
| Random Forest     | 0.974 | 0.851 | 0.851 | 0.855 | 0.851  | 0.826 |

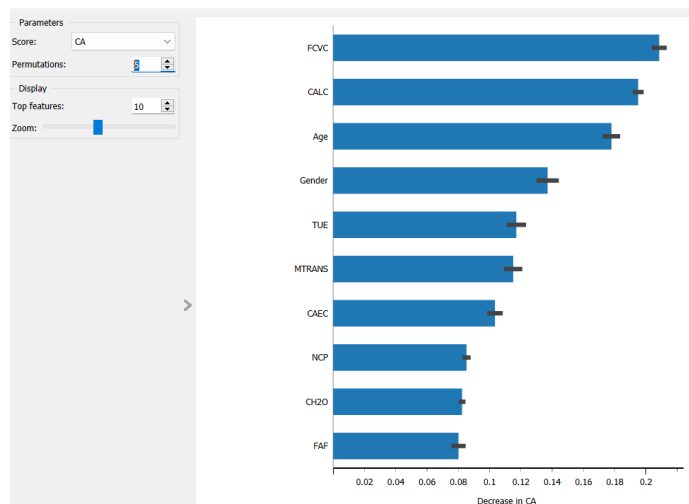
Kemudian, menggunakan “*Confusion Matrix*” untuk melihat kategori yang salah diklasifikasikan, sehingga dapat mengetahui letak kesalahan pada

model dalam memprediksi. Berikut hasil *confusion matrix* dari algoritma *Gradient Boosting* yang memiliki nilai akurasi tertinggi :

|        |             | Predicted   |           |            |            |            |           |           | Σ   |
|--------|-------------|-------------|-----------|------------|------------|------------|-----------|-----------|-----|
|        |             | Insuffic... | Normal... | Obesity... | Obesity... | Obesity... | Overwe... | Overwe... |     |
| Actual | Insuffic... | 46          | 4         | 0          | 0          | 0          | 2         | 0         | 52  |
|        | Normal...   | 8           | 29        | 1          | 0          | 0          | 4         | 5         | 47  |
|        | Obesity...  | 0           | 4         | 59         | 5          | 0          | 0         | 2         | 70  |
|        | Obesity...  | 0           | 1         | 2          | 61         | 0          | 0         | 2         | 66  |
|        | Obesity...  | 0           | 0         | 0          | 0          | 78         | 0         | 1         | 79  |
|        | Overwe...   | 0           | 4         | 11         | 0          | 0          | 35        | 2         | 52  |
|        | Overwe...   | 0           | 2         | 4          | 4          | 0          | 3         | 43        | 56  |
| Σ      |             | 54          | 44        | 77         | 70         | 78         | 44        | 55        | 422 |

Berdasarkan nilai akurasi klasifikasi atau CA pada kedua tugas, *Random Forest* menonjol dengan akurasi tertinggi. Oleh karena itu, menggunakan algoritma *Random Forest*, dataset dapat dijadikan acuan dalam pengembangan perangkat lunak untuk memperkirakan tingkat obesitas berdasarkan pola makan dan kondisi fisik di masa mendatang.

Tidak hanya itu, dengan menggunakan algoritma ini, kita dapat mengidentifikasi faktor-faktor yang memiliki pengaruh signifikan pada tingkat obesitas. Di platform *Orange*, kita dapat mengetahui faktor-faktor tersebut dengan menghubungkan algoritma serta data sampler dengan fitur '*Feature Importance*'. Berikut adalah urutan 10 faktor yang paling berpengaruh terhadap tingkat obesitas:



## E. KESIMPULAN DAN SARAN

Kinerja model *Random Forest* terbukti lebih unggul daripada *Gradient Boosting* dalam mengklasifikasikan tingkat obesitas. *Random Forest* memberikan hasil kinerja yang konsisten dengan nilai yang tinggi, menunjukkan akurasi yang sangat baik dan dapat diandalkan dalam mengklasifikasikan tingkat obesitas.

Dengan menggunakan skor akurasi klasifikasi dari *Random Forest* sebagai parameter, kami berhasil mengidentifikasi 10 faktor yang memiliki kontribusi signifikan terhadap tingkat obesitas, yaitu (secara berurutan): seberapa sering mengonsumsi sayur (FCVC), seberapa sering mengonsumsi alkohol (CALC), umur (Age), jenis kelamin (Gender), seberapa sering menggunakan alat teknologi (TUE), transportasi yang biasa digunakan (MTRANS), apakah mengonsumsi makanan antar jam makan (NCP), berapa kali makan dalam sehari (NCP), seberapa banyak mengonsumsi air dalam sehari (CH2O), dan seberapa sering melakukan aktivitas fisik (FAF).

Dengan mengetahui faktor-faktor signifikan tersebut, pemerintah dan lembaga kesehatan dapat merancang program yang lebih detail, terfokus, dan terstruktur untuk mengatasi masalah obesitas di Meksiko, Peru, dan Kolombia.

Model *Random Forest* yang memiliki nilai akurasi tinggi, sebesar 0.851, menunjukkan bahwa dataset dapat digunakan sebagai training set atau acuan untuk pembuatan perangkat lunak yang akan mendeteksi dan mengklasifikasikan tingkat berat badan secara langsung di masa depan. Keberadaan perangkat lunak tersebut akan memudahkan warga dalam mengakses informasi kesehatan mereka.

Kami merekomendasikan untuk melengkapi dataset dengan atribut lain yang berpotensi memengaruhi berat badan, seperti tingkat stres dan durasi istirahat. Dengan demikian, kita dapat memiliki gambaran yang lebih holistik tentang faktor-faktor yang

berperan dalam masalah obesitas. Selanjutnya, kami menyarankan untuk eksplorasi penggunaan model yang dimodifikasi atau kombinasi untuk meningkatkan akurasi prediksi. Dengan pendekatan ini, kita dapat meningkatkan kinerja model dalam memprediksi tingkat obesitas dengan lebih akurat.

## **F. REFERENSI**

- Christanto, H., Rahmad, J., Sinurat, S. H., Hamonangan Sitompul, D. R., Sitomorang, A., Ziegel, D. J., & Indra, E. (2023). Analisis Perbandingan Decision Tree, Support Vector Machine, dan Xgboost dalam Mengklasifikasi Review Hotel Trip Advisor. *Jurnal Teknologi Informatika Dan Komputer*, 9(1), 306–319. <https://doi.org/10.37012/jtik.v9i1.1429>
- Mursianto Ghaita Amany, Falih Isma'il Muhammad, Irfan Muhammad, Sakinah Tiara, & Prasvita Desta Sandya. (2021). Perbandingan Metode Klasifikasi Random Forest dan XGBoost Serta Implementasi Teknik SMOTE pada Kasus Prediksi Hujan. *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, 2(2).
- Palechor, F. M., & Manotas, A. de la H. (2019). Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico. *Data in Brief*, 25. <https://doi.org/10.1016/j.dib.2019.104344>
- Sumber dataset : Estimation of Obesity Levels Based On Eating Habits and Physical Condition . (2019). UCI Machine Learning Repository. <https://doi.org/10.24432/C5H31Z>.

## **G. LAMPIRAN**

# PENGARUH POLA MAKAN DAN KONDISI FISIK PADA OBESITAS

## Problem

Obesitas adalah tantangan kesehatan yang serius dan kompleks. Ketika terlalu banyak lemak disimpan dalam tubuh, kesehatan kita dapat terganggu dan masa hidup kita dapat dipersingkat. Gaya hidup kita, termasuk pola makan dan aktivitas fisik, memiliki dampak besar pada berat badan kita, yang bisa menjadi penyebab utama obesitas. Oleh karena itu, penting bagi kita semua untuk meningkatkan kesadaran akan pentingnya gaya hidup sehat dan mengambil langkah-langkah untuk menjaga kesehatan tubuh kita.



## Tujuan Analisa

- Mengidentifikasi faktor-faktor yang signifikan dalam menentukan tingkat obesitas.
- Membangun model prediktif untuk memperkirakan tingkat obesitas berdasarkan pola makan dan kondisi fisik.
- Membantu lembaga kesehatan dan pemerintah (terutama negara Meksiko, Peru, dan Kolombia) dalam merancang strategi intervensi yang efektif untuk mengatasi masalah obesitas
- Membantu Masyarakat dalam Peningkatan Kesadaran dan Akses Informasi Kesehatan:

## Solusi

- Pengumpulan dataset di Meksiko, Peru, dan Kolombia mengenai pola makan, aktivitas fisik, dan faktor-faktor lain yang terkait dengan obesitas.
- Analisis data untuk mengidentifikasi hubungan antara variabel yang diukur dan tingkat obesitas.
- Pengembangan model prediktif menggunakan teknik klasifikasi seperti Gradient Boosting dan Random Forest.
- Evaluasi model untuk memilih algoritma terbaik dan menentukan faktor-faktor yang paling berpengaruh.
- Penerapan hasil analisis untuk mendukung pembuatan kebijakan dan program intervensi oleh pemerintah dan lembaga kesehatan.
- Dataset dan model dapat menjadi acuan pengembangan perangkat lunak dalam prediksi tingkat obesitas, serta memudahkan warga dalam mengakses informasi kesehatan mereka
- Dengan menggabungkan elemen-elemen ini dalam sebuah poster atau video animasi, informasi akan disampaikan secara jelas dan menarik, sehingga dapat memperkuat pemahaman tentang pentingnya penanganan obesitas di wilayah tersebut.

