

PROPUESTA DE **PROYECTO**

Presentación realizada por Irene Arrabé



CONTENIDO

01

Problema

02

Objetivos

03

Tratamiento de
los datos

04

Modelos

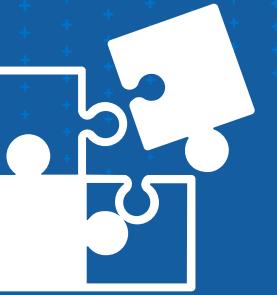
05

Resultados
elegidos

06

Creación de
la app

PROBLEMA



Crear un sistema preciso y
automatizado para
recomendaciones de salud
física.



OBJETIVOS

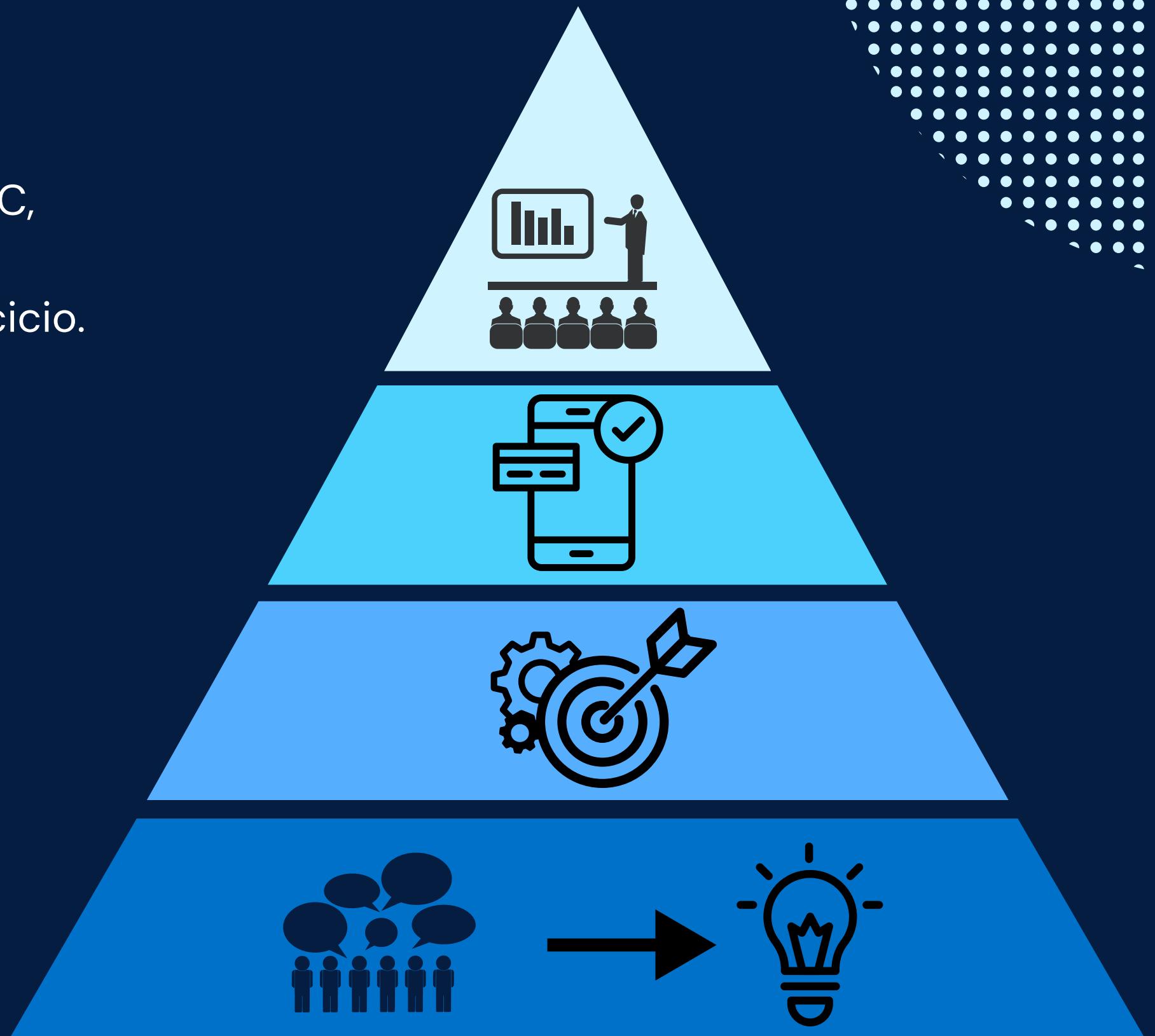
Aplicación basada en ML para calcular IMC, estimar porcentaje de grasa corporal y calorías quemadas en una sesión de ejercicio.

01 Generar la idea

02 Conseguir un modelo que nos de los resultados esperados

03 Crear una app eficiente y accesible

04 Plantear mejoras futuras



TRATAMIENTO DE LOS DATOS



- Limpia de datos
- one-hot encoding
- label encoding
- Creación de nuevas columnas

Prueba de modelos supervisados y no supervisados:

- RandomForest
- Polinomico
- LinerRegressor
- PCA
- Keras
- Ridge

Feature Engineering

Análisis

Modelos

Resultados

Estudio de la relación entre las variables y elección de las más relevantes

Elección del modelo que mejores aproximaciones haga

Datasets

CLASIFICACIÓN

Age
Gender
Height
Weight
CALC
FAVC
FCVC
NCP
SCC
SMOKE
CH2O
family_history_with_overweight
FAF
TUE
CAEC
MTRANS
NObeyesdad
BMI

KCAL QUEMADAS

Age
Gender
Weight (kg)
Height (m)
Max_BPM
Avg_BPM
Resting_BPM
Session_Duration (hours)
Calories_Burned
Workout_Type
Fat_Percentage
Water_Intake (liters)
Workout_Frequency
Experience_Level
BMI

- ORIGEN: kaggle

	TUE	-0.30	0.05	-0.07	0.10	0.04	0.01	0.06					
NObeyesdad	-0.16	0.09	0.81	0.25	0.07	0.09	-0.15	-0.11					
Male	-0.05	0.62	0.16	-0.27	0.07	0.11	0.19	0.02	-0.07				
BMI	-0.23	0.13	0.93	0.26	0.04	0.14	-0.18	-0.10	0.87	-0.05			
family_with_overweight	-0.20	0.25	0.50	0.04	0.07	0.15	-0.06	0.02	0.36	0.10	0.48		
Alcohol	-0.04	0.13	0.21	0.06	0.07	0.09	-0.09	-0.05	0.17	-0.01	0.17	-0.01	
Andar_bici	-0.05	0.04	-0.10	0.00	0.01	0.02	0.13	0.06	-0.06	0.06	-0.13	0.00	
Picoteo	-0.06	0.10	0.11	0.10	0.15	-0.16	-0.02	0.11	0.08	-0.06	0.08	0.00	
Fumador	-0.10	0.06	0.03	0.01	0.01	-0.03	0.01	0.02	0.03	0.04	-0.00	0.00	
Control_kcal	-0.11	-0.13	0.20	0.07	-0.02	0.01	0.07	-0.01	-0.13	-0.10	-0.18	0.00	
FastFood	-0.06	0.18	0.27	-0.03	-0.01	0.01	-0.11	0.07	0.24	0.06	0.25	0.00	
	Age	Height	Weight	FCVC	NCP	CH2O	FAF	TUE	NObeyesdad	Male	BMI		

The figure is a correlation matrix heatmap for fitness metrics. The x and y axes list the following variables:

- Session_Duration (hours)
- Calories_Burned
- Workout_Type
- Fat_Percentage
- Water_Intake (liters)
- Workout_Frequency (days/week)
- Experience_Level
- BMI
- Male
- Age
- Weight (kg)
- Height (m)
- Max_BPM
- Avg_BPM
- Resting_BPM
- Session_Duration (hours)
- Calories_Burned
- Workout_Type

Key observations from the matrix:

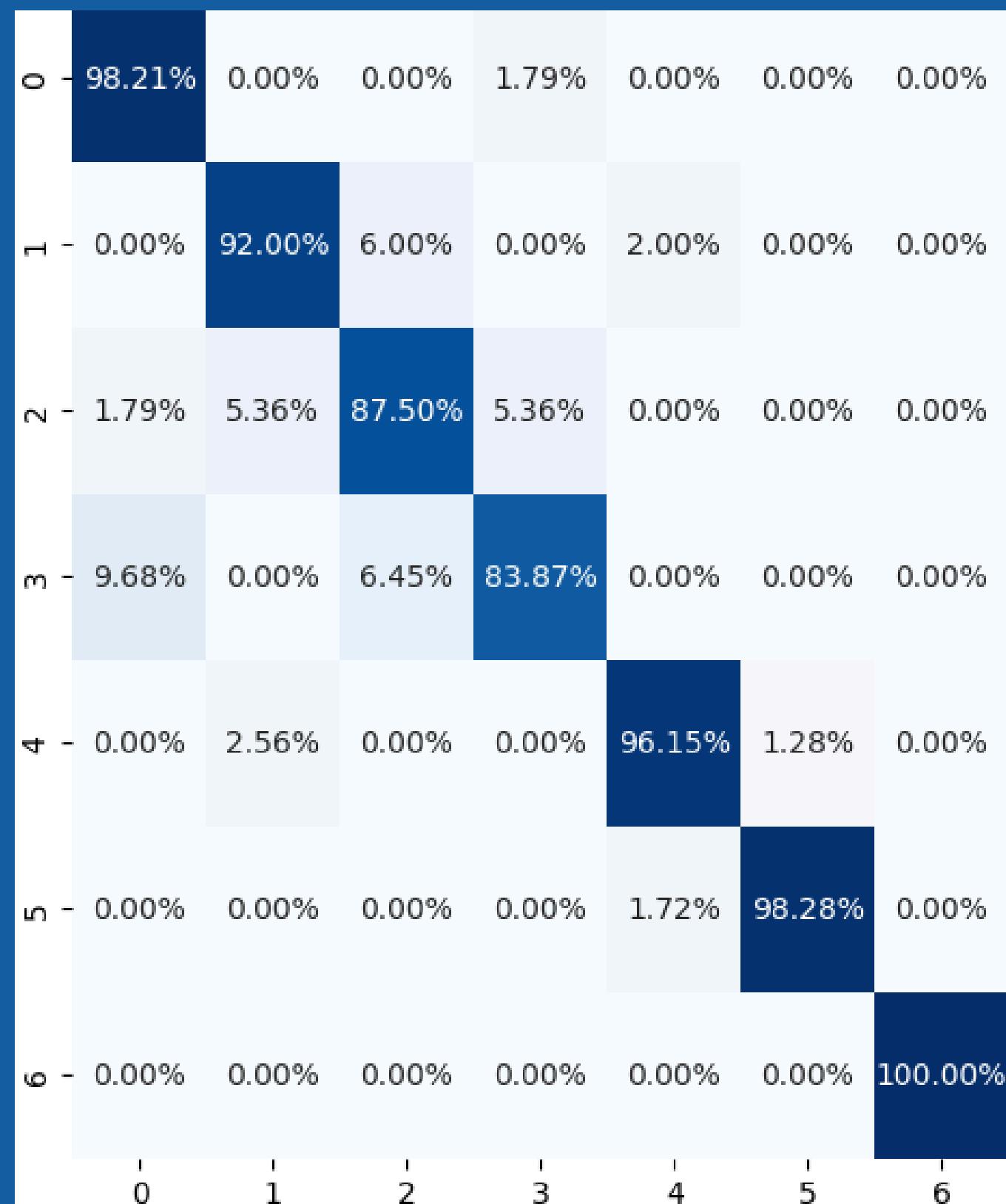
- Calories_Burned** has a strong negative correlation with **Workout_Type** (-0.17) and **Fat_Percentage** (-0.24).
- Workout_Frequency** has a strong positive correlation with **Experience_Level** (0.75) and **Calories_Burned** (0.63).
- Experience_Level** has a strong positive correlation with **Calories_Burned** (0.68).
- BMI** has a very strong negative correlation with **Calories_Burned** (-0.85).
- Male** has a moderate positive correlation with **Calories_Burned** (0.57).
- Age**, **Weight (kg)**, and **Height (m)** show moderate positive correlations with **Calories_Burned**.
- Session_Duration (hours)** has a strong positive correlation with **Calories_Burned** (0.91).

MODELOS

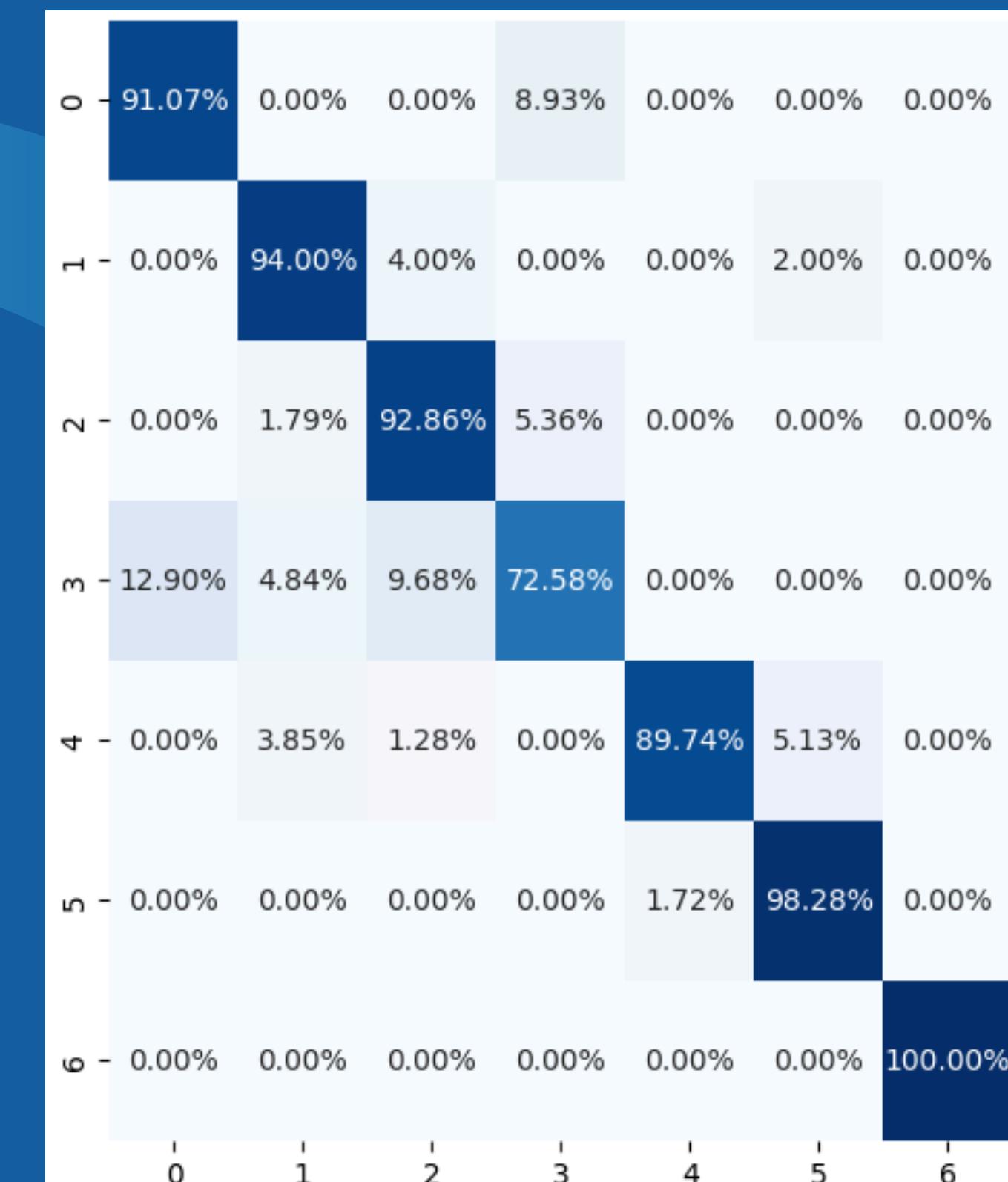
1. Clasificación
2. Regresión



Redes Neuronales



RandomForestClassifier



RESULTADOS DE LAS PREDICCIONES

Redes Neuronales

RandomForest

RESULTADOS DE LAS PREDICCIONES

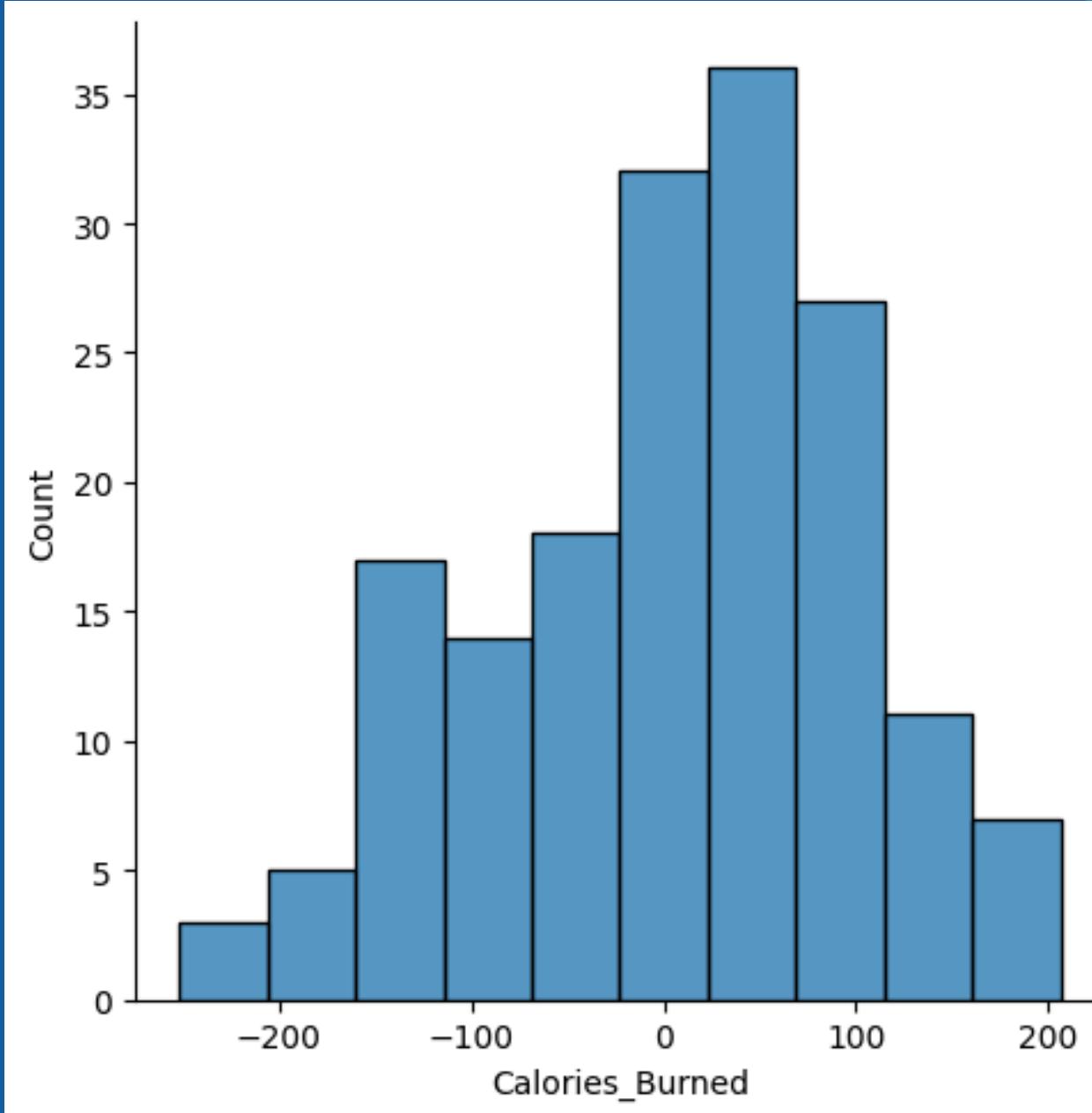
Redes Neuronales

	precision	recall	f1-score
0	0.89	0.98	0.93
1	0.90	0.92	0.91
2	0.88	0.88	0.88
3	0.93	0.84	0.88
4	0.97	0.96	0.97
5	0.98	0.98	0.98
6	1.00	1.00	1.00
accuracy			0.94

RandomForest

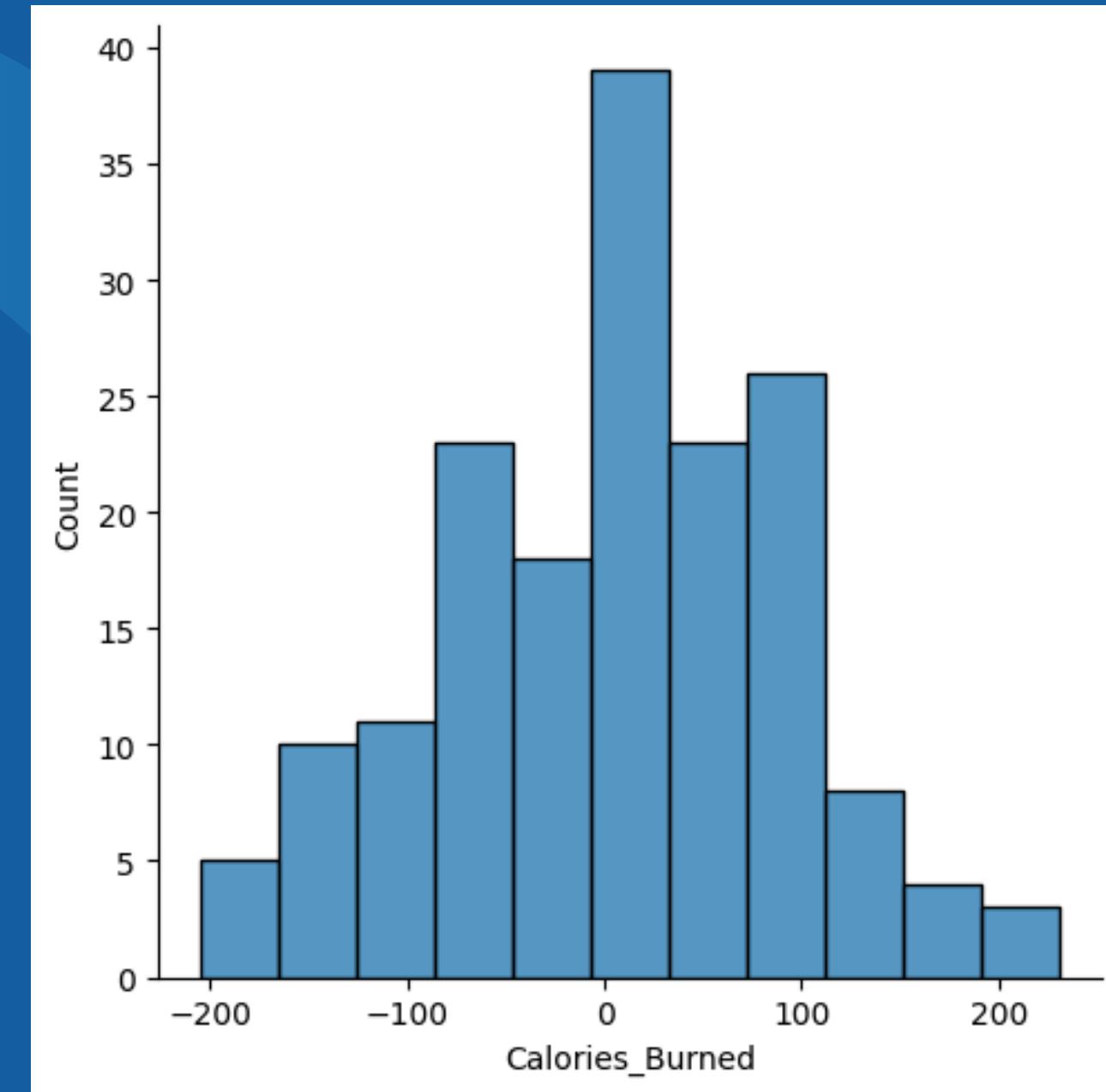
	precision	recall	f1-score
0	0.86	0.91	0.89
1	0.87	0.94	0.90
2	0.85	0.93	0.89
3	0.85	0.73	0.78
4	0.99	0.90	0.94
5	0.92	0.98	0.95
6	1.00	1.00	1.00
accuracy			0.91

RandomForestRegressor



R2 score 0.8648127440459994
R2 score train 0.9197284739778846
MAE 79.00778240629157

PolynomialFeatures (degree=2)



R2 score 0.89325012398219960189
R2 score train 0.8825634354384349
MAE 68.94984126559174

RESULTADOS DE LAS PREDICCIONES

RandomForest

	kcal	Kcal_real
8	496.0	454.0
9	1118.0	1107.0
10	737.0	576.0
11	507.0	433.0
12	997.0	911.0
13	1010.0	864.0
14	733.0	621.0

MAE: 88.28 kcal

RMSE: 109.79 kcal

R² Score: 0.85

Polynomial_2

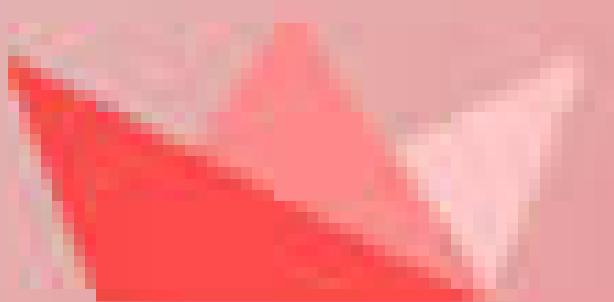
	kcal	Kcal_real
8	441.0	454.0
9	1148.0	1107.0
10	699.0	576.0
11	507.0	433.0
12	1015.0	911.0
13	994.0	864.0
14	693.0	621.0

MAE: 85.37 kcal

RMSE: 107.15 kcal

R² Score: 0.86

CREACIÓN DE LA APP



Streamlit

Proceso

- 1.- Cargar los modelos**
- 2.- Diseñar la app**
- 3.- Pedir los datos al usuario**
- 4.- Hacer un llamado a las funciones y modelos predefinidos para obtener los resultados**

*my healthy
maki*

