



# Fundamentos del aprendizaje automático

(Machine learning)

Joaquín Luque

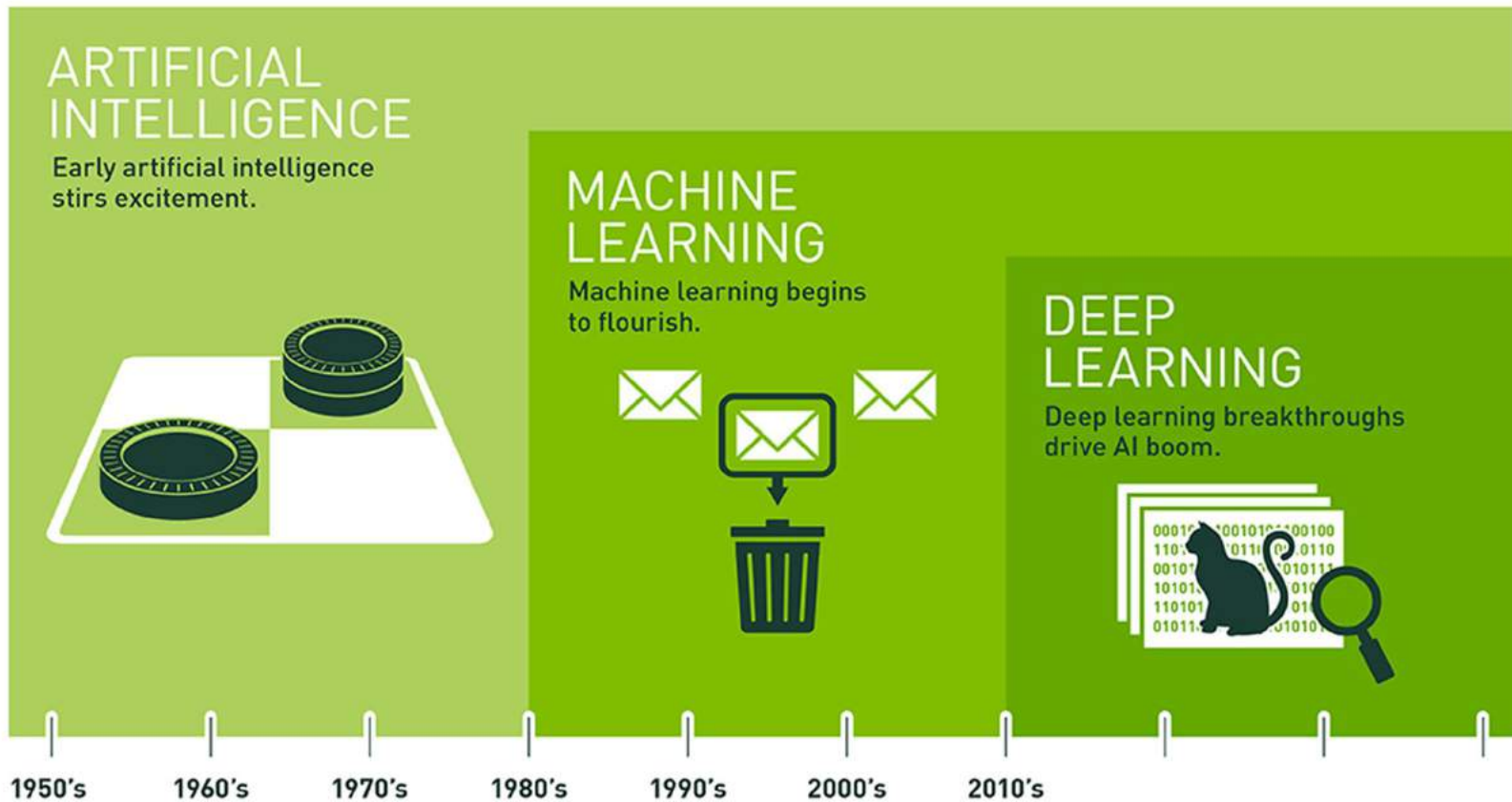
# Contenido

1. Introducción
2. Regresión
  - a) Regresión univariable
  - b) Regresión multivariable
3. Clasificación
  - a) Regresión logística
  - b) Máquinas de vectores soporte (SVM)
    - Forma dual de la optimización (regresión y SVM)
  - c) Funciones Kernel
  - d) Clasificación multiclase
4. Segmentación
5. Reducción de dimensionalidad
6. Deep learning (introducción)

# Definiciones

- Inteligencia artificial
  - Máquinas (programas) que son capaces de imitar (incluso mejorar) comportamientos humanos que etiquetaríamos como “inteligentes”
- Machine learning (aprendizaje automático)
  - Una de las técnicas de la inteligencia artificial
  - La máquina “aprende” a conseguir sus resultados
    - En base a los datos
    - Sin que se expliciten las reglas que debe seguir

# Definiciones

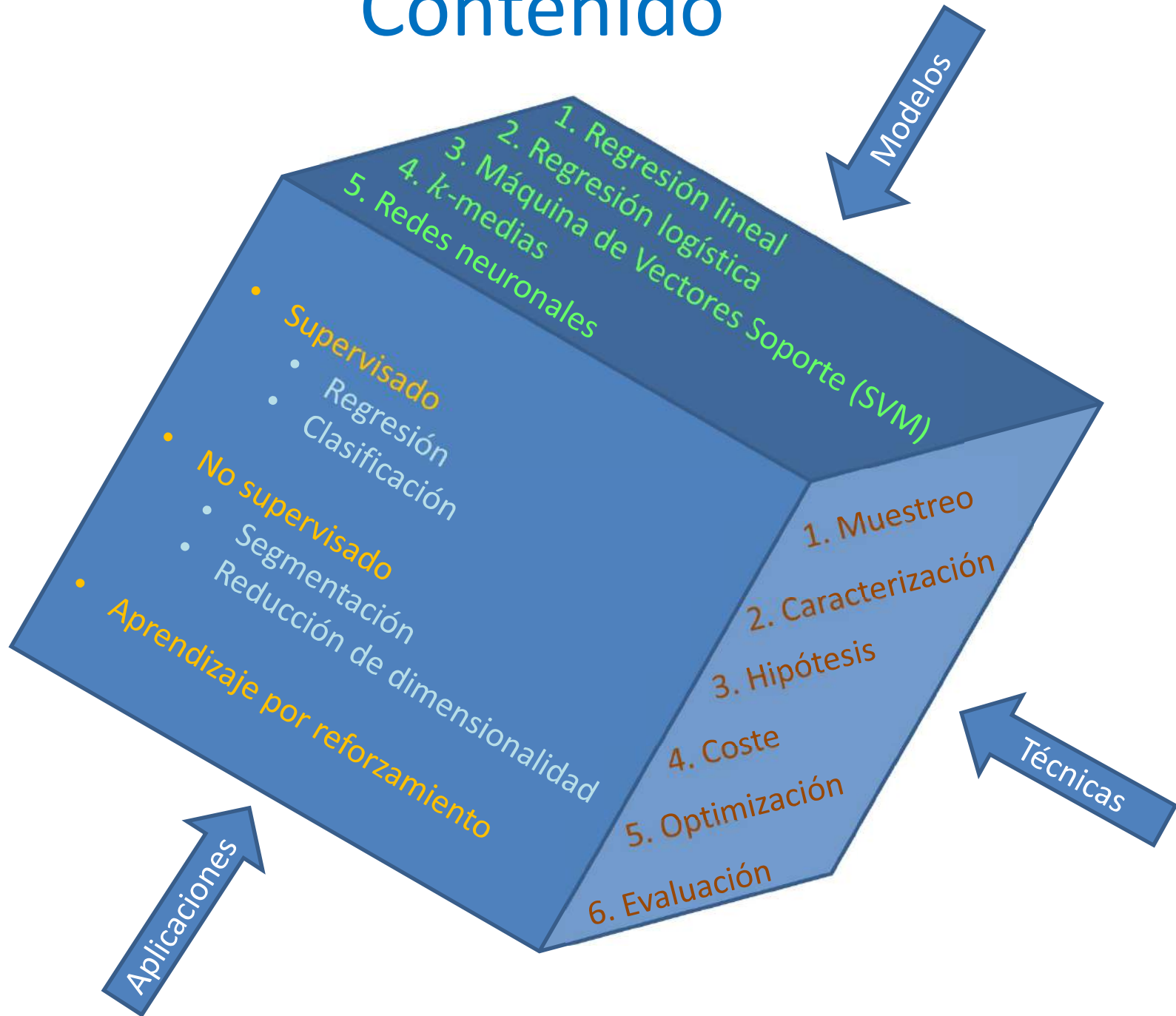


Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

# Contenido



# Contenido



# Contenido

## Aplicaciones

- Aprendizaje (por aplicaciones)
  - Supervisado
    - Regresión
      - Univariable
      - Multivariable
    - Clasificación
  - No supervisado
    - Reducción de dimensionalidad
    - Segmentación
  - Aprendizaje por reforzamiento





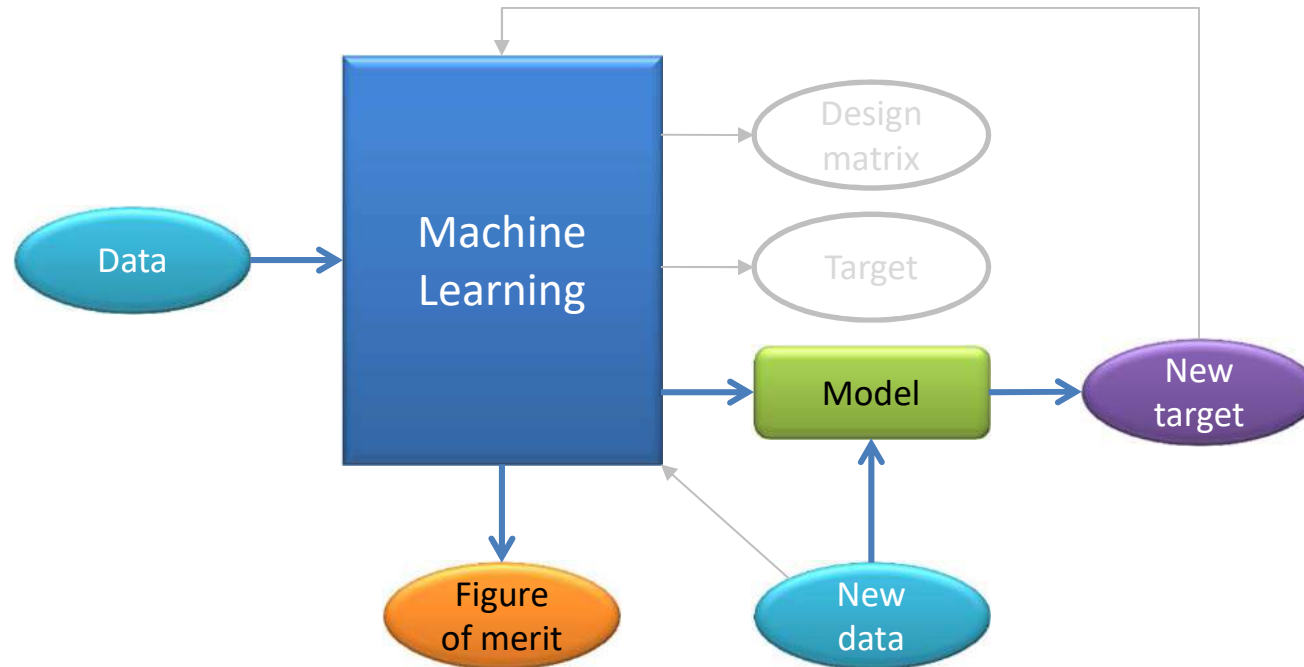
# Estructura del *machine learning*

## Clasificación

Cliente	Ingresos (m€) $x_1^{(i)}$	Edad (años) $x_2^{(i)}$	Vehículo $y^{(i)}$
1	97.17	26.6	S
2	44.67	32.3	N
3	46.64	26.7	N
4	33.84	23.7	N
5	79.35	27.0	S
⋮	⋮	⋮	⋮

# Estructura del *machine learning*

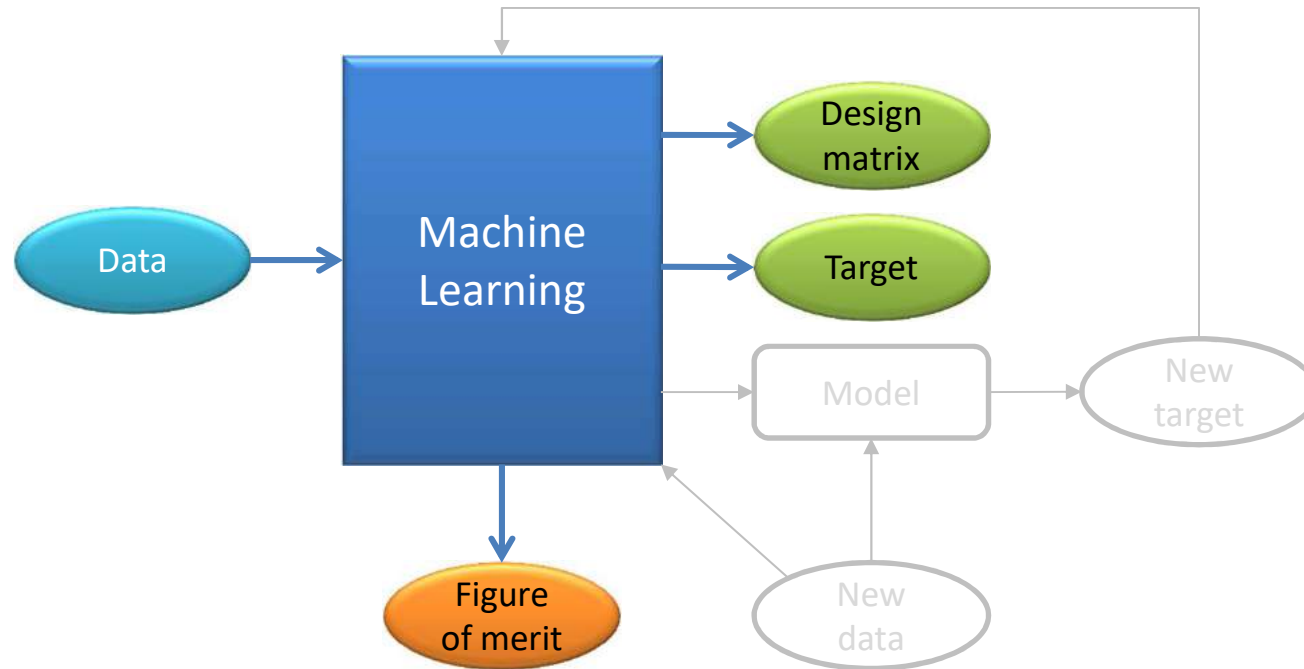
## Regresión y Clasificación





# Estructura del *machine learning*

## Reducción de dimensionalidad



# Estructura del *machine learning*

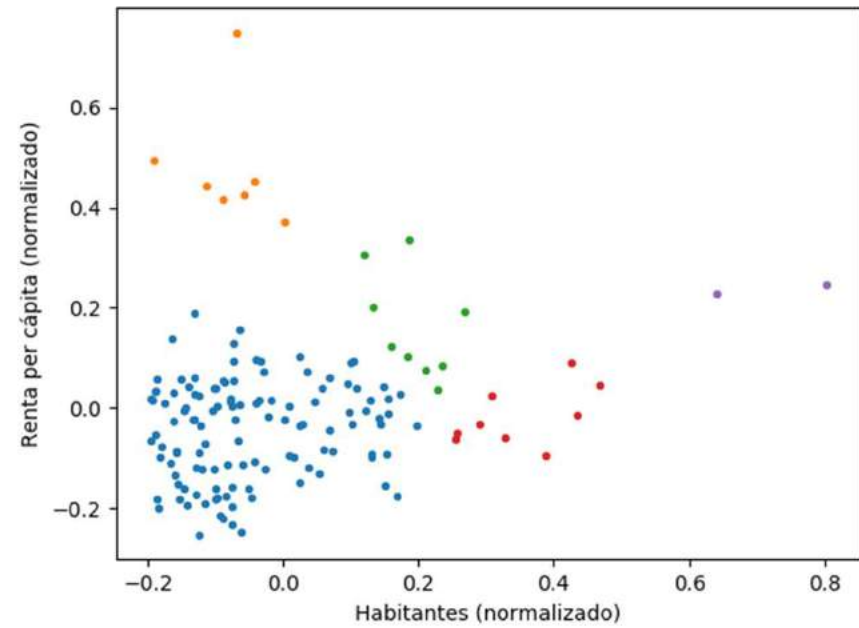
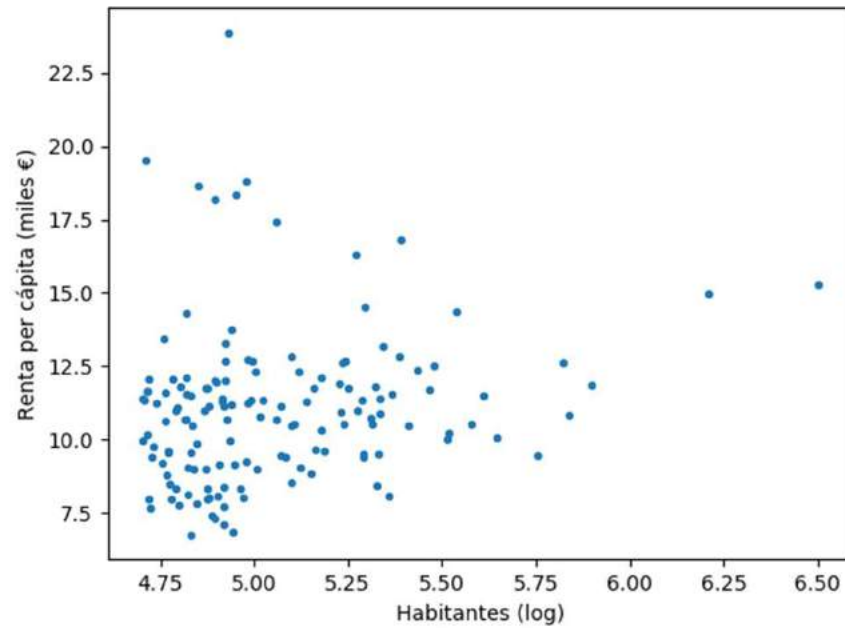
## Segmentación

Ciudad	Nombre	Población $x_1^{(i)}$	Ingresos (€) $x_2^{(i)}$	Segmento $y^{(i)} \in \{1 \dots K\}$
1	Albacete	172,816	10,515	?
2	Alcalá de Guadaíra	75,106	8,332	?
3	Alcalá de Henares	194,310	11,320	?
4	Alcobendas	114,864	17,417	?
5	Alcorcón	168,141	11,917	?
⋮	⋮	⋮	⋮	⋮

No existen datos de entrenamiento  
(clasificación no supervisada)

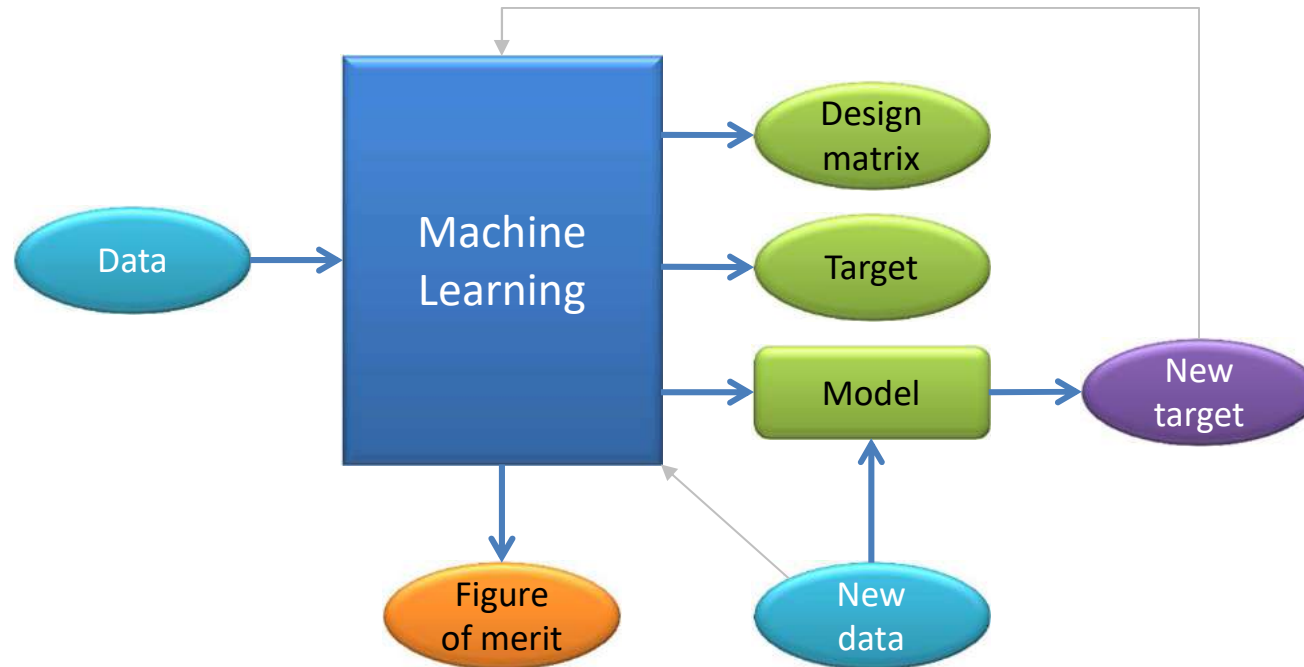
# Estructura del *machine learning*

## Segmentación



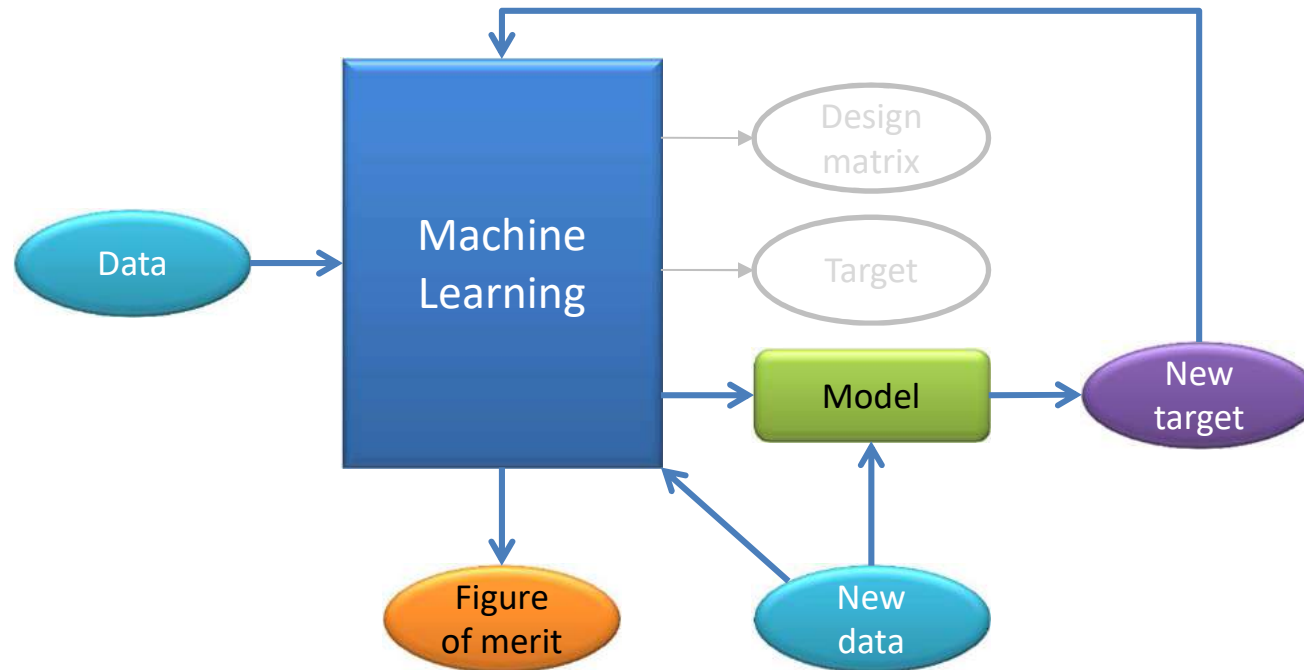
# Estructura del *machine learning*

## Segmentación



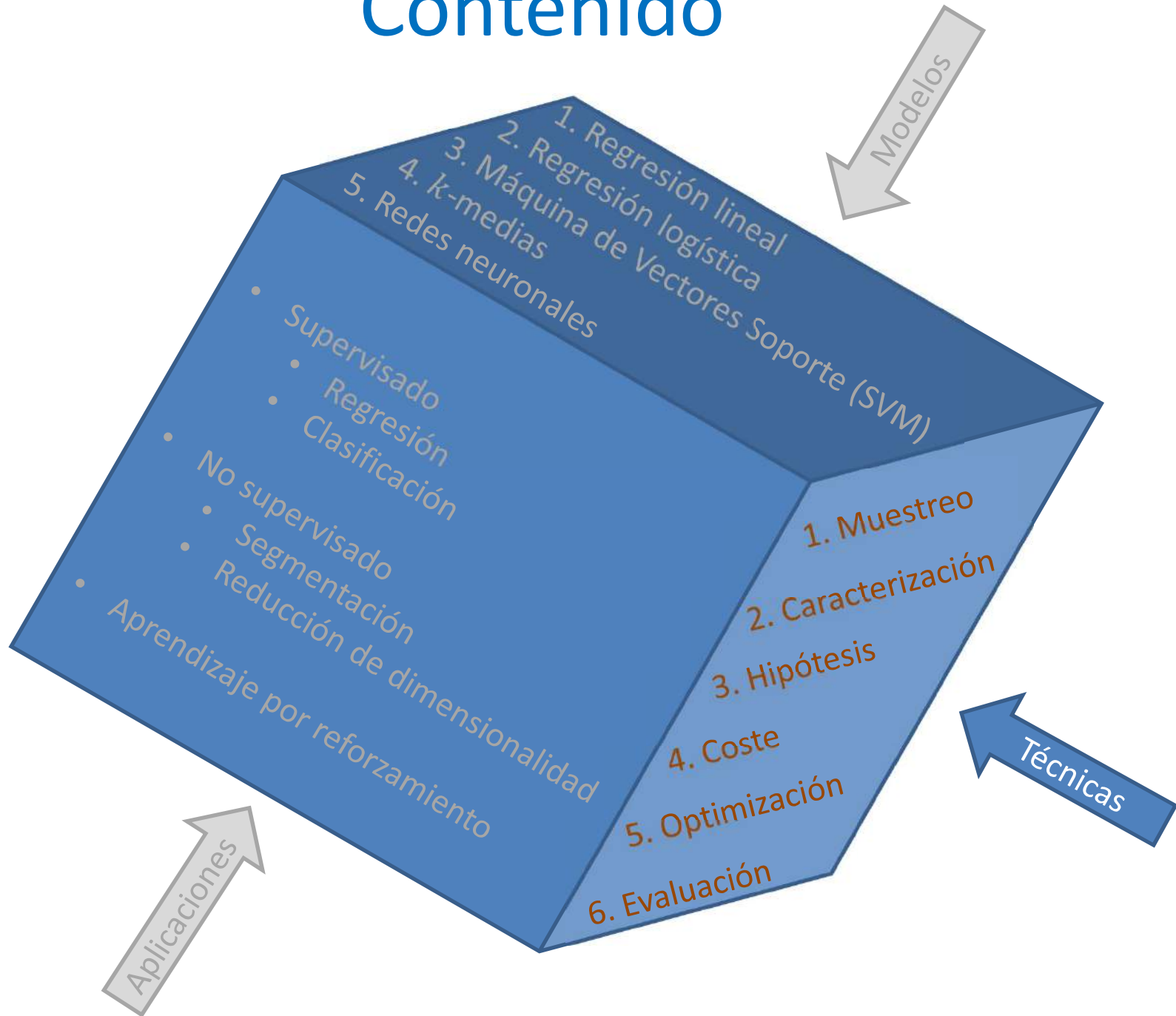
# Estructura del *machine learning*

## Aprendizaje por reforzamiento

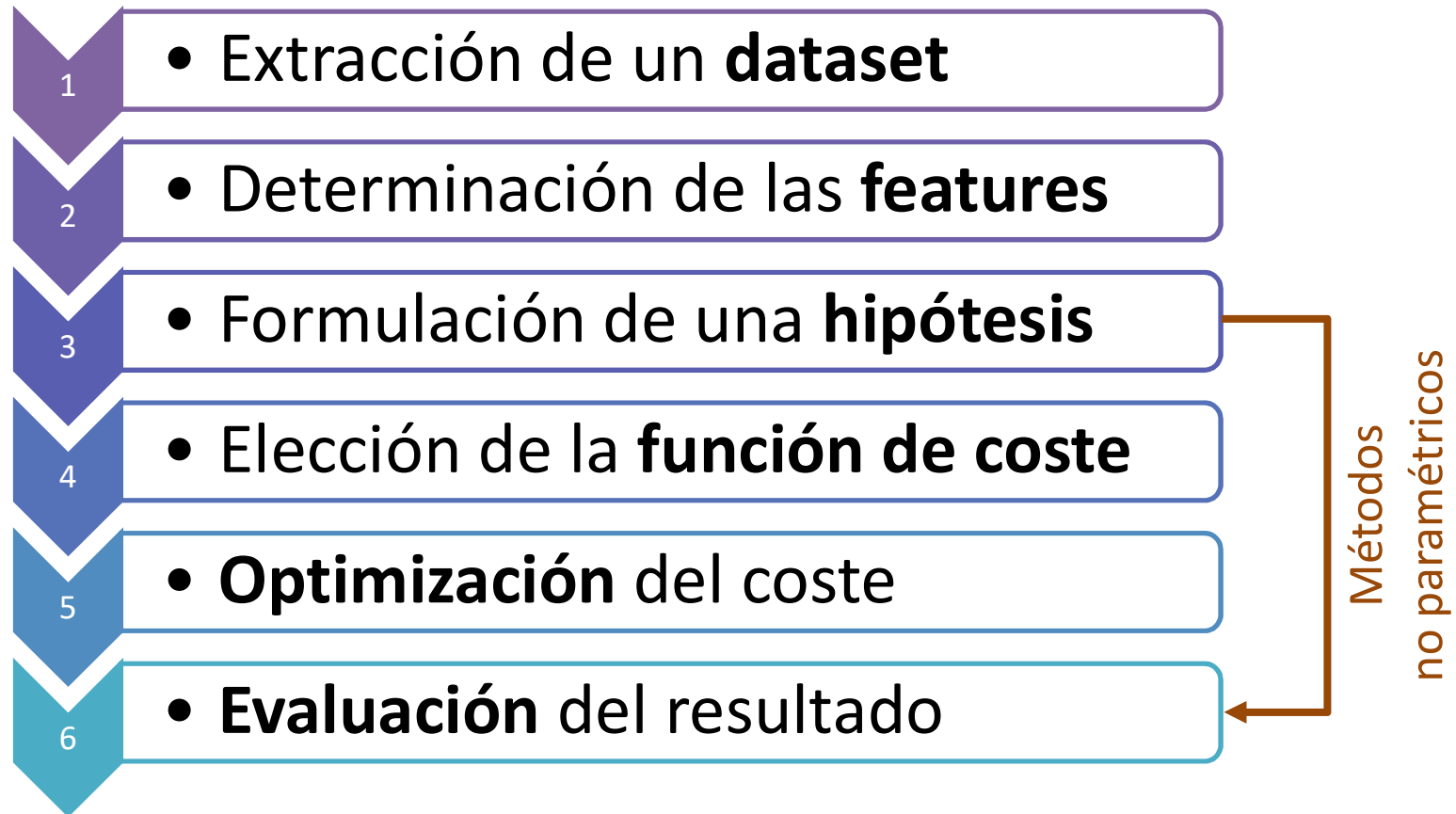




# Contenido



# Pasos del *machine learning*



# Contenido

## Técnicas

- Aprendizaje (por técnicas)
  - Muestreo
    - Simple
    - Hold-out
    - Cros-validación
  - Determinación de características (features)
  - Formulación de hipótesis
  - Función de coste
  - Optimización
  - Evaluación

# Contenido

## Técnicas

- Aprendizaje (por técnicas)
  - Muestreo
  - Determinación de características (features)
    - Representación multidimensional
    - Codificación 1-hot
    - Normalización de variables
    - Kernels
    - Análisis de componentes principales
  - Formulación de hipótesis
  - Función de coste
  - Optimización
  - Evaluación

# Contenido

## Técnicas

- Aprendizaje (por técnicas)
  - Muestreo
  - Determinación de características (features)
  - Formulación de hipótesis
    - Lineal
    - Polinómica
    - Logística
  - Función de coste
  - Optimización
  - Evaluación

# Contenido

## Técnicas

- Aprendizaje (por técnicas)
  - Muestreo
  - Determinación de características (features)
  - Formulación de hipótesis
  - Función de coste
    - Cuadrática
    - Logística
    - Bisagra
    - Distorsión
    - Regularización (regresión de arista)
  - Optimización
  - Evaluación

# Contenido

## Técnicas

- Aprendizaje (por técnicas)
  - Muestreo
  - Determinación de características (features)
  - Formulación de hipótesis
  - Función de coste
  - Optimización
    - Ecuación normal
      - Simple
      - Descomposición en Valores Singulares
    - Gradiente descendiente
      - Simple
      - Estocástico
      - Lotes
    - Optimización dual
    - Método del codo
  - Evaluación

# Contenido

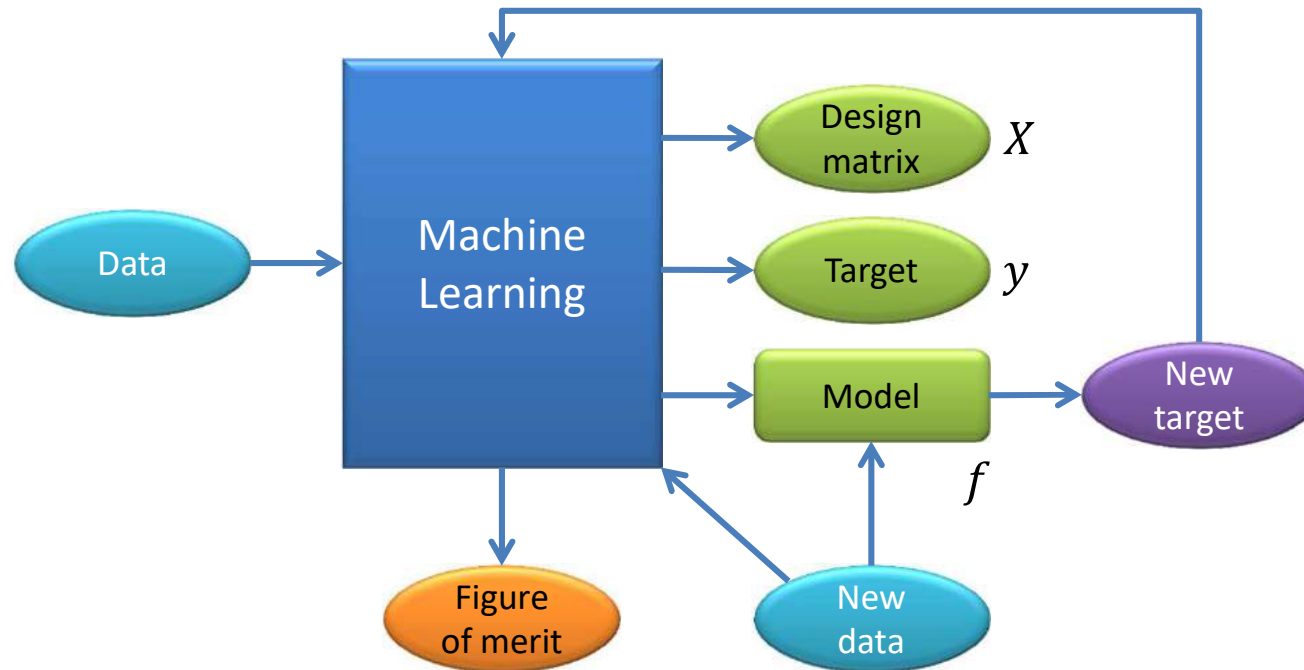
## Técnicas

- Aprendizaje (por técnicas)
  - Muestreo
  - Determinación de características (features)
  - Formulación de hipótesis
  - Función de coste
  - Optimización
  - Evaluación
    - Bootstrap
    - Compromiso sesgo-varianza
    - Curva de aprendizaje
    - Matriz de confusión
    - Medidas de prestaciones
    - Silueta
    - Índice de Rand



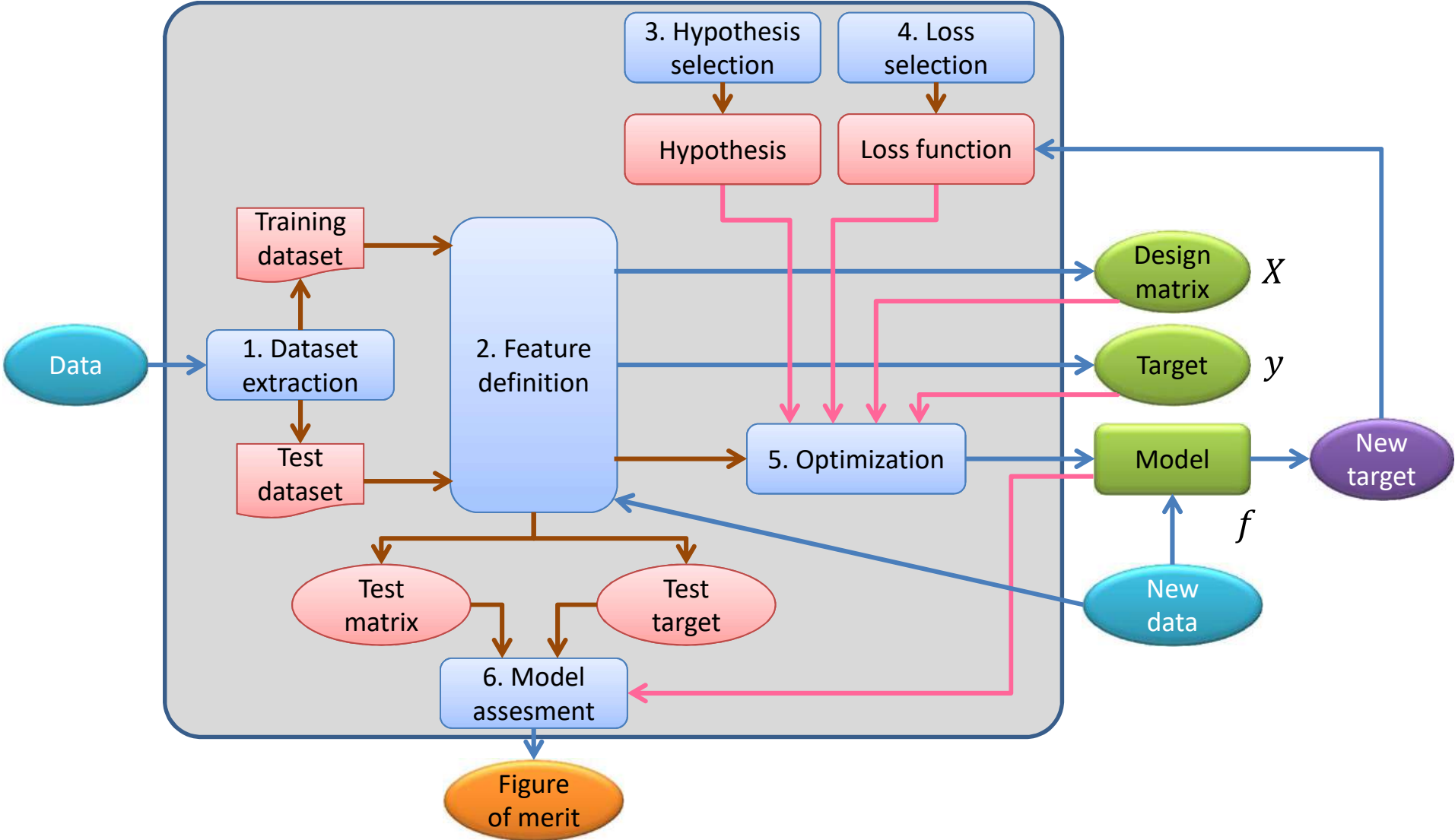
# Estructura del *machine learning*

## Modelo completo



# Estructura del *machine learning*

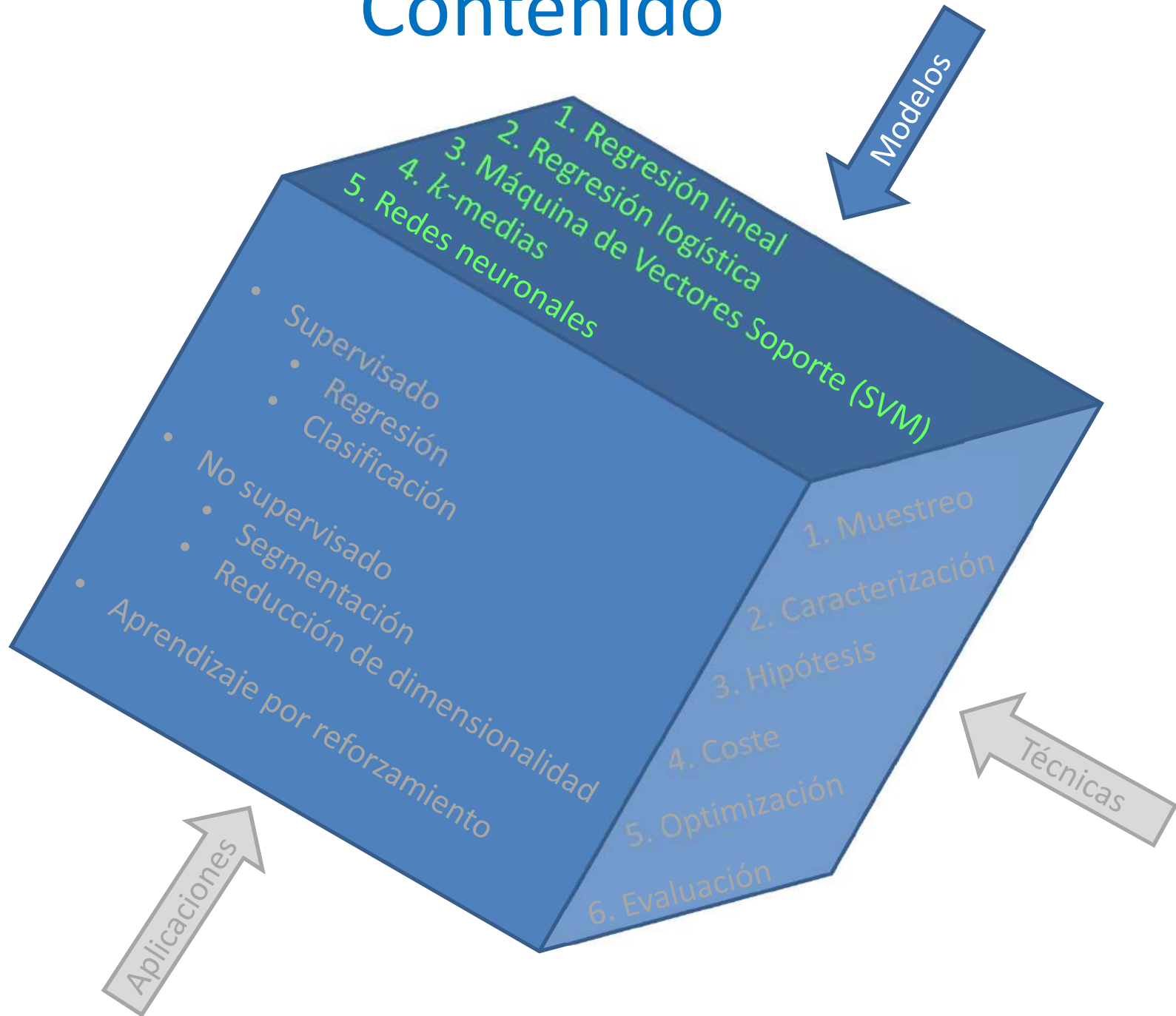
## Detalles



# Niveles del *machine learning*

			$f$	$\varphi$	$K$	$X$
Rule-based systems		$y = f(X)$	Manual	-	-	Manual
Machine Learning	Basic ML	$y = f(X)$	Automatic	-	-	Manual
	Feature-based ML	$y = f[\varphi(X)]$	Automatic	Manual	-	Raw
	Kernel-based ML	$y = f[K(X)]$	Automatic	Automatic	Manual	Raw
	Deep Learning	$y = f(\varphi(X))$	Automatic	Automatic	-	Raw

# Contenido



# Contenido

## Modelos

- Aprendizaje (por modelos)
  - Regresión lineal
  - Regresión logística
  - Máquina de Vectores Soporte (SVM)
  - $k$ -medias
  - Redes neuronales