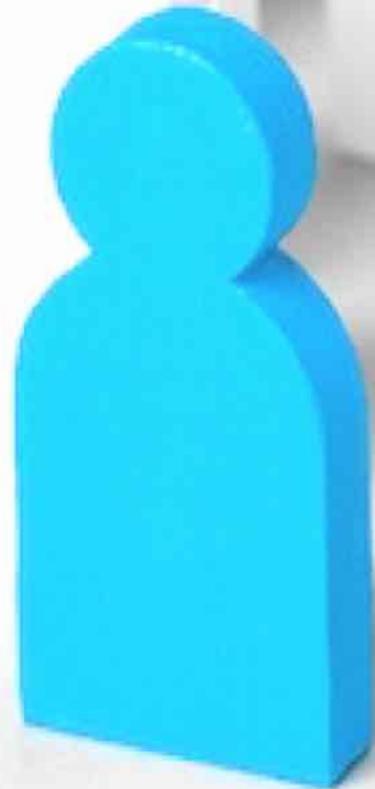


Clasificación temprana de clientes mediante Redes Neuronales

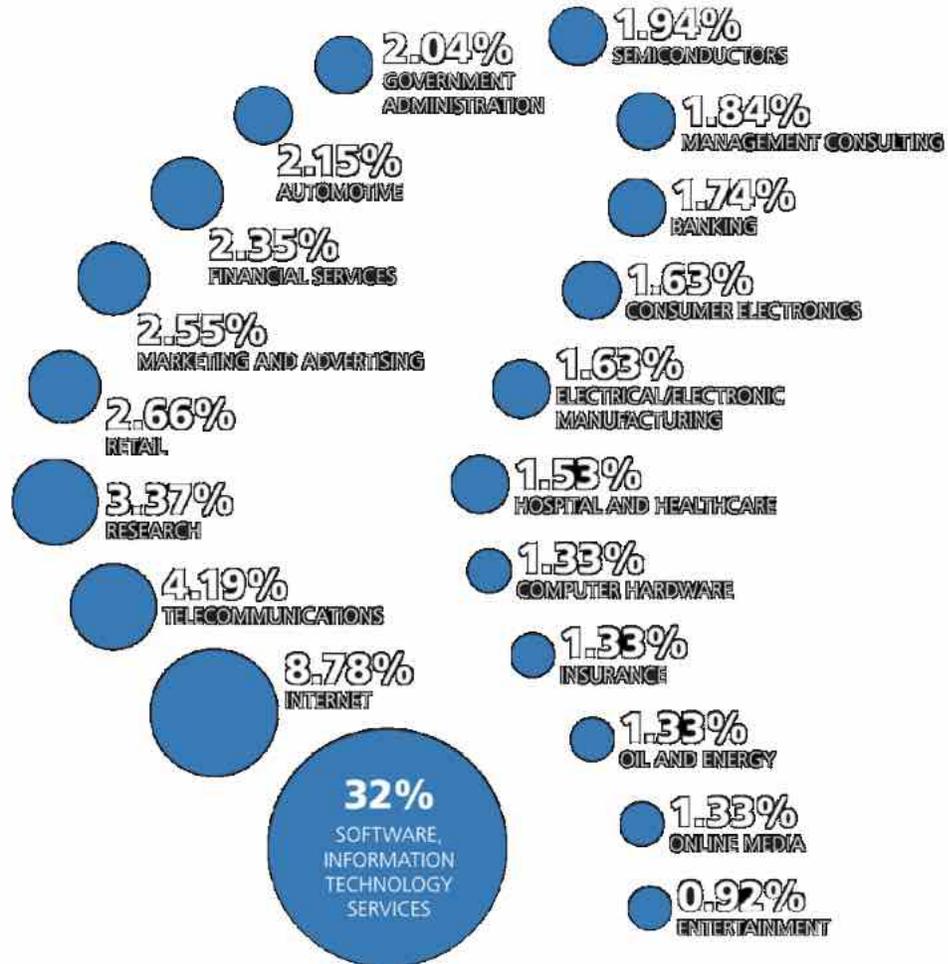


Grupo Caja Rural

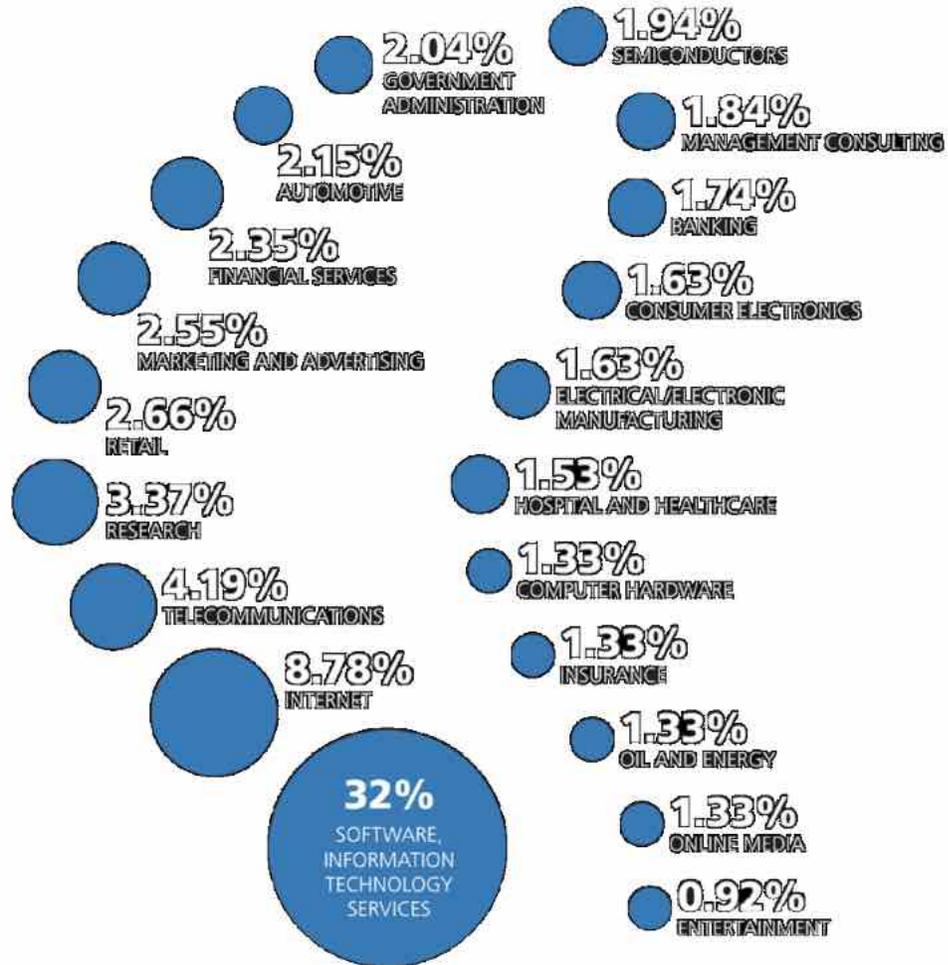


01. INTELIGENCIA ARTIFICIAL EN EL SECTOR BANCARIO

01. Inteligencia artificial en el sector bancario



01. Inteligencia artificial en el sector bancario



Sector bancario
6.75%

01. Inteligencia artificial en el sector bancario



01. Inteligencia artificial en el sector bancario

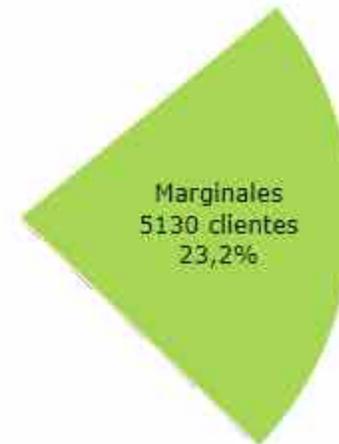


01. Inteligencia artificial en el sector bancario

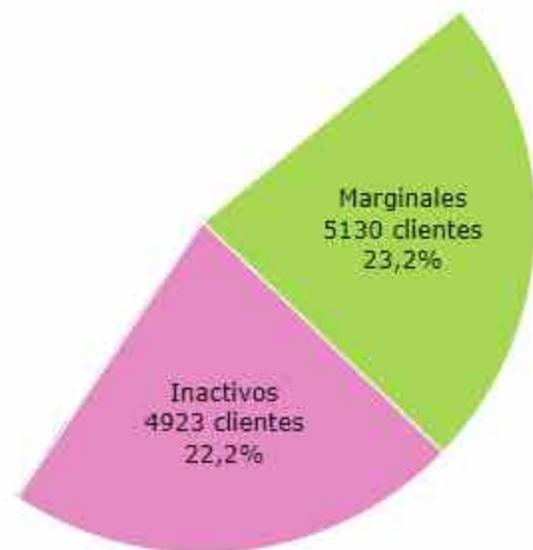


02. ADAPTACIÓN AL MERCADO Y AL CLIENTE

02. Adaptación al mercado y al cliente



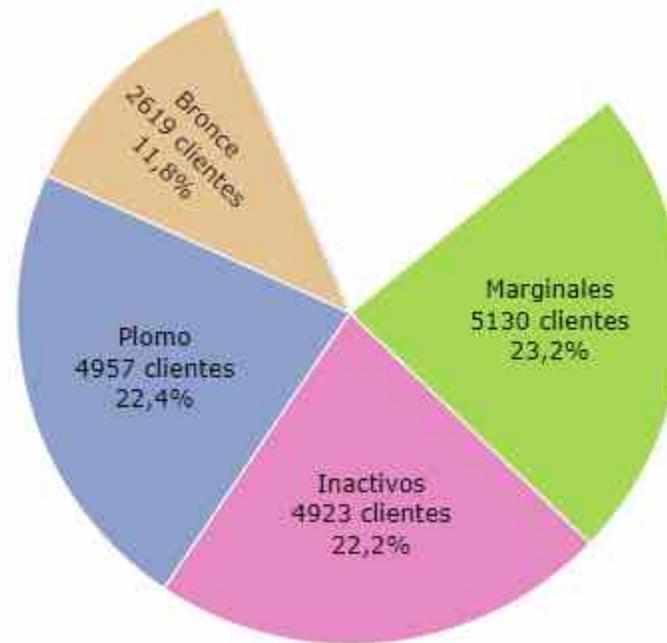
02. Adaptación al mercado y al cliente



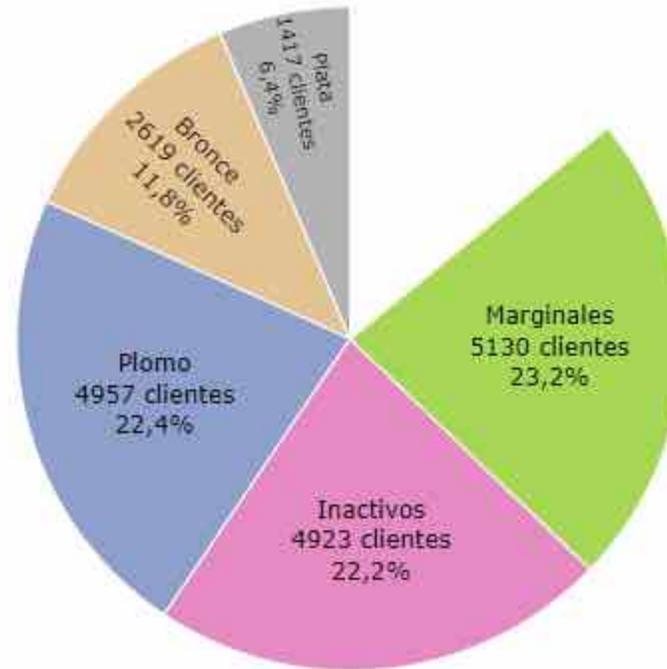
02. Adaptación al mercado y al cliente



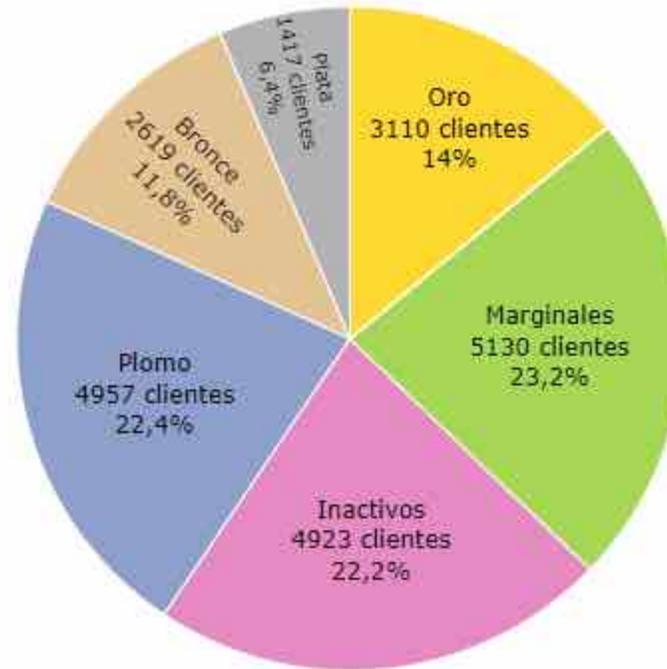
02. Adaptación al mercado y al cliente



02. Adaptación al mercado y al cliente

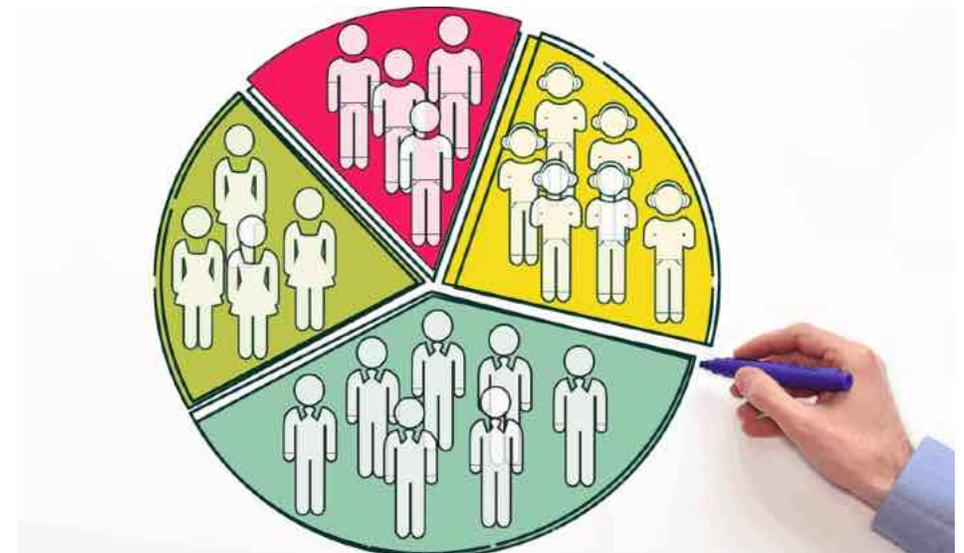


02. Adaptación al mercado y al cliente

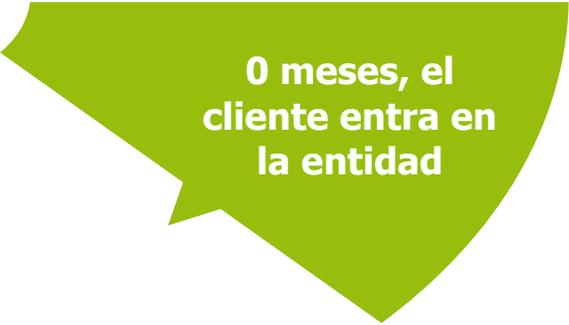




02. Adaptación al mercado y al cliente



03. ALGORITMOS PREDICTIVOS

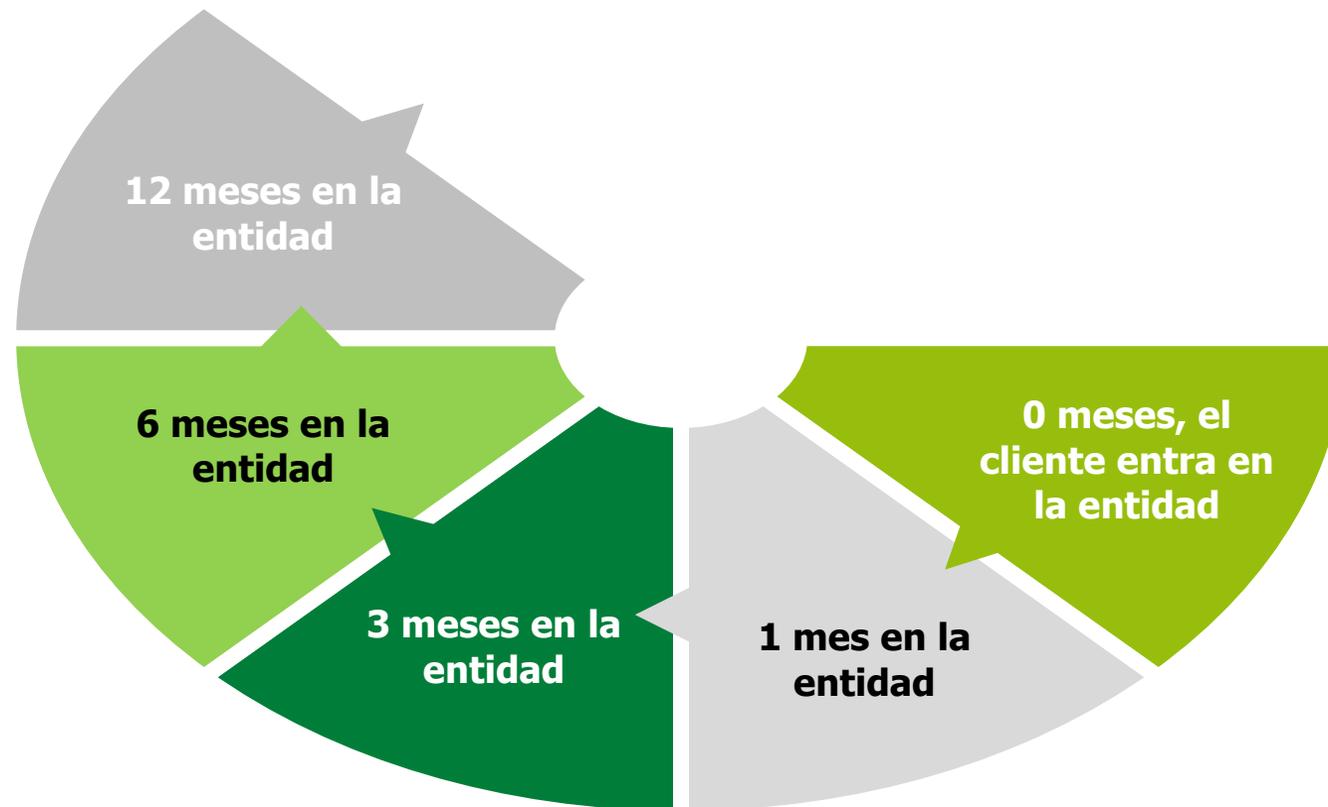
A green callout box with a pointed bottom and a tail pointing towards the bottom-left. It contains the text '0 meses, el cliente entra en la entidad' in white, bold, sans-serif font.

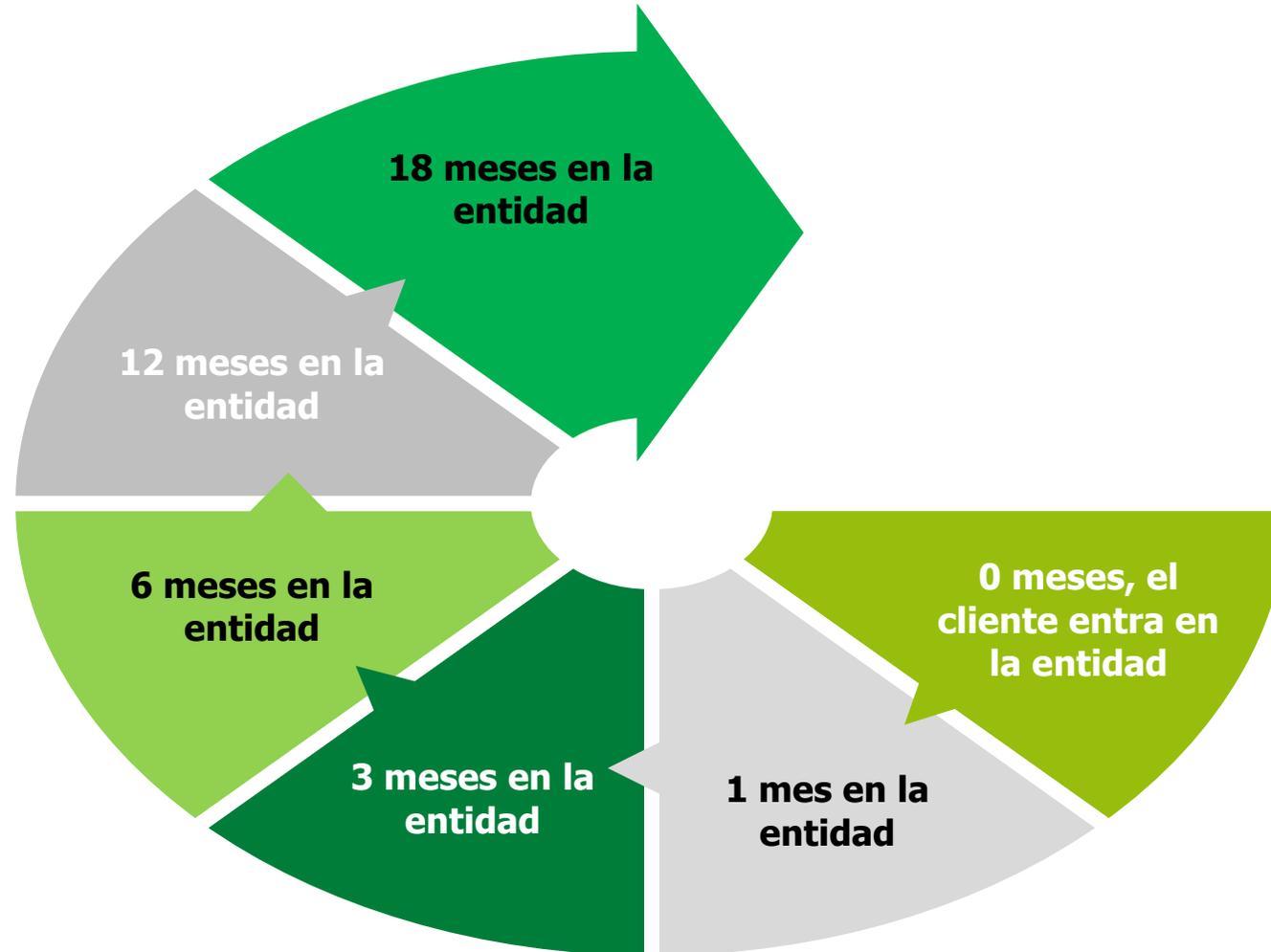
**0 meses, el
cliente entra en
la entidad**

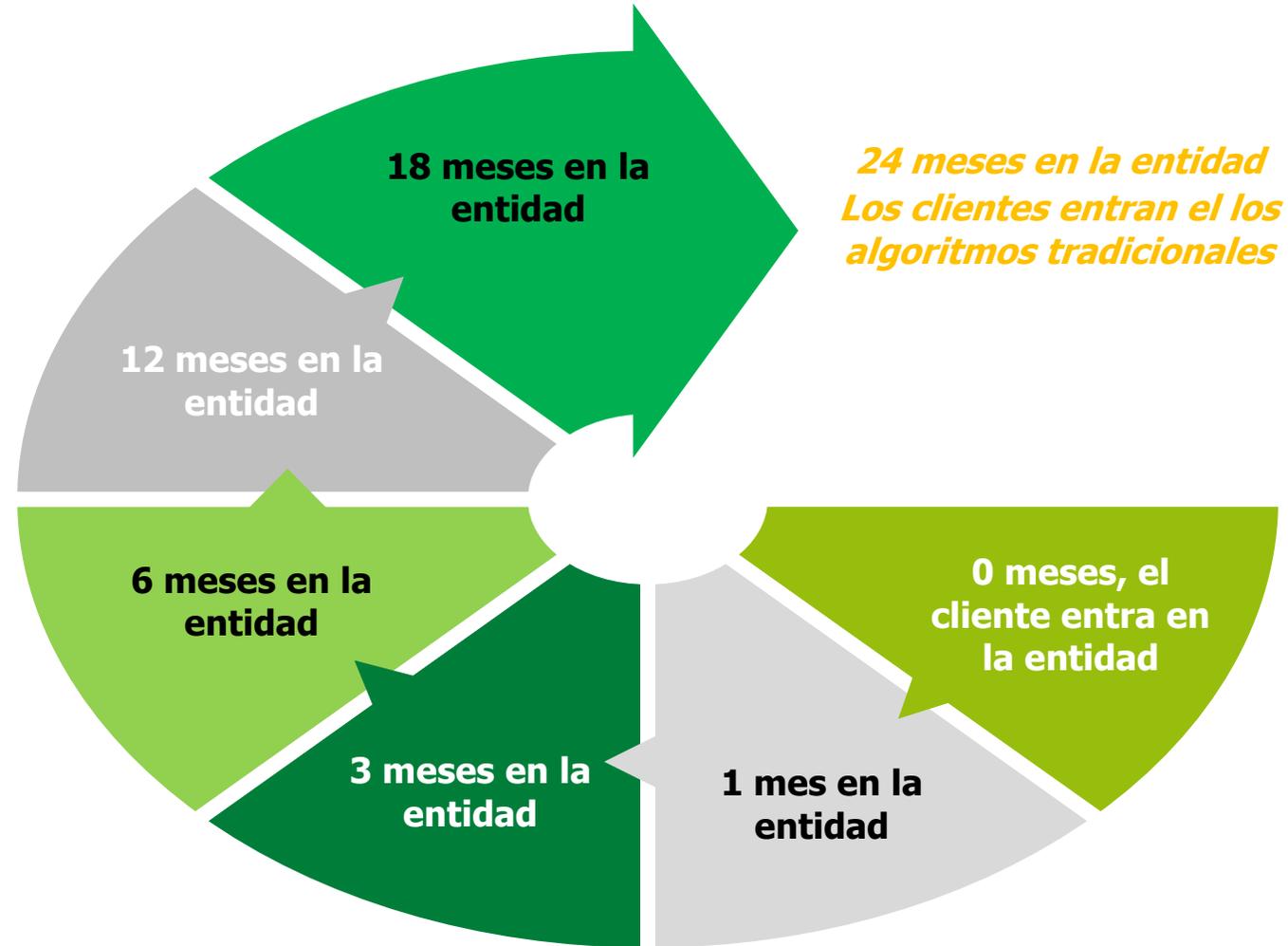












Variables económicas



Ventajas

Usaremos las series temporales de las variables económicas disponibles de los clientes.

Conoceremos las evoluciones y valores de sus variables económicas y movimientos en cuentas, teniendo una visualización de la evolución del cliente con nosotros.

Variables económicas



Ventajas

Usaremos las series temporales de las variables económicas disponibles de los clientes.

Conoceremos las evoluciones y valores de sus variables económicas y movimientos en cuentas, teniendo una visualización de la evolución del cliente con nosotros.

Variables comerciales



Ventajas

Este set de variables describirá la política comercial actual de cada entidad frente a estos nuevos clientes.

Tendremos variables como el nivel de clientes en cada grupo, la política de expansión, la movilidad de los clientes, etc.

Variables económicas



Ventajas

Usaremos las series temporales de las variables económicas disponibles de los clientes.

Conoceremos las evoluciones y valores de sus variables económicas y movimientos en cuentas, teniendo una visualización de la evolución del cliente con nosotros.

Variables comerciales



Ventajas

Este set de variables describirá la política comercial actual de cada entidad frente a estos nuevos clientes.

Tendremos variables como el nivel de clientes en cada grupo, la política de expansión, la movilidad de los clientes, etc.

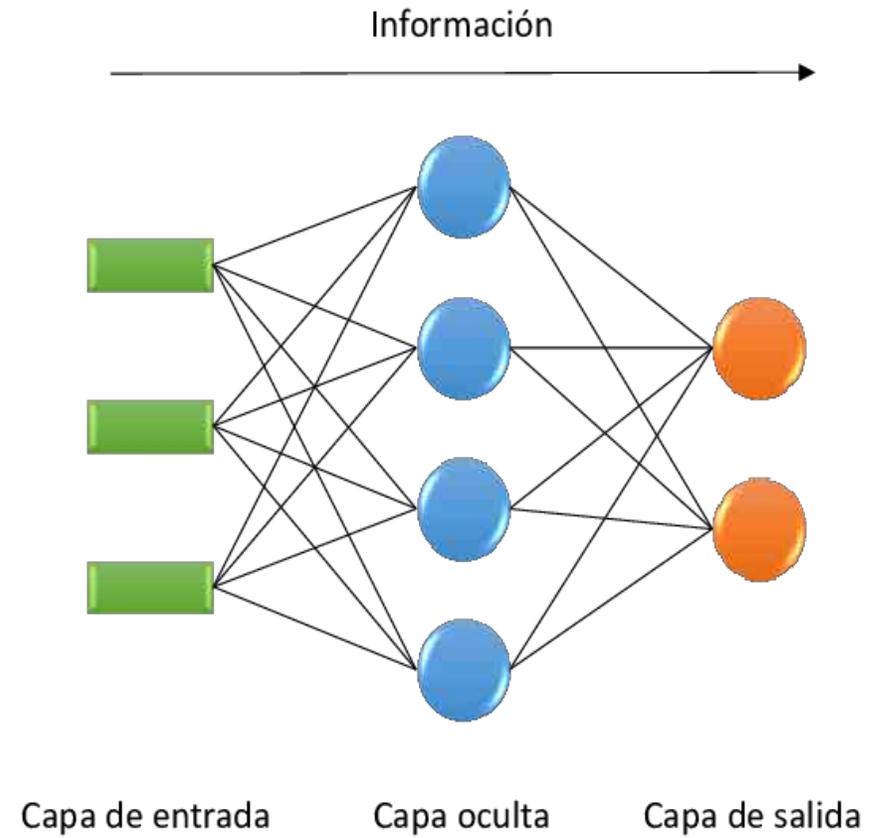
Variables geográficas

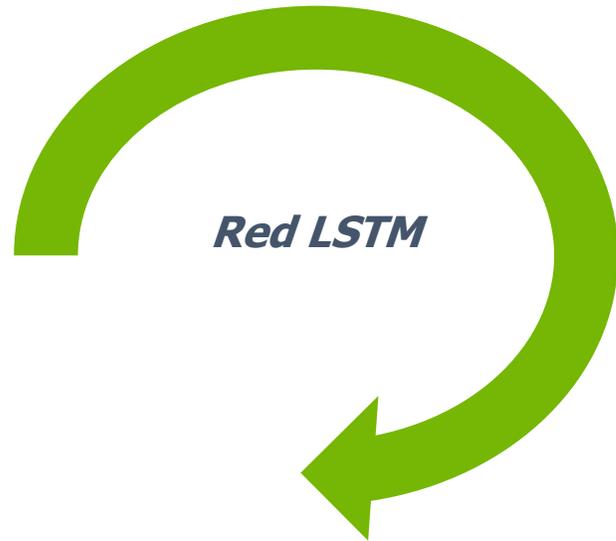


Ventajas

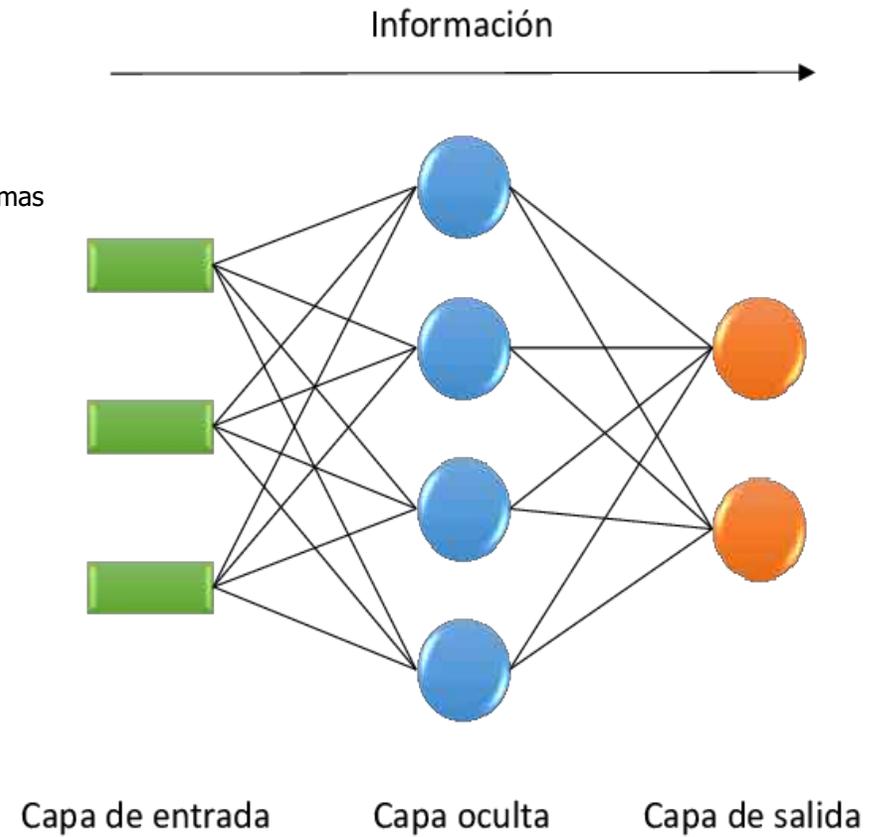
Estas variables describen el poder adquisitivo y demográfico del cliente según su domicilio actual.

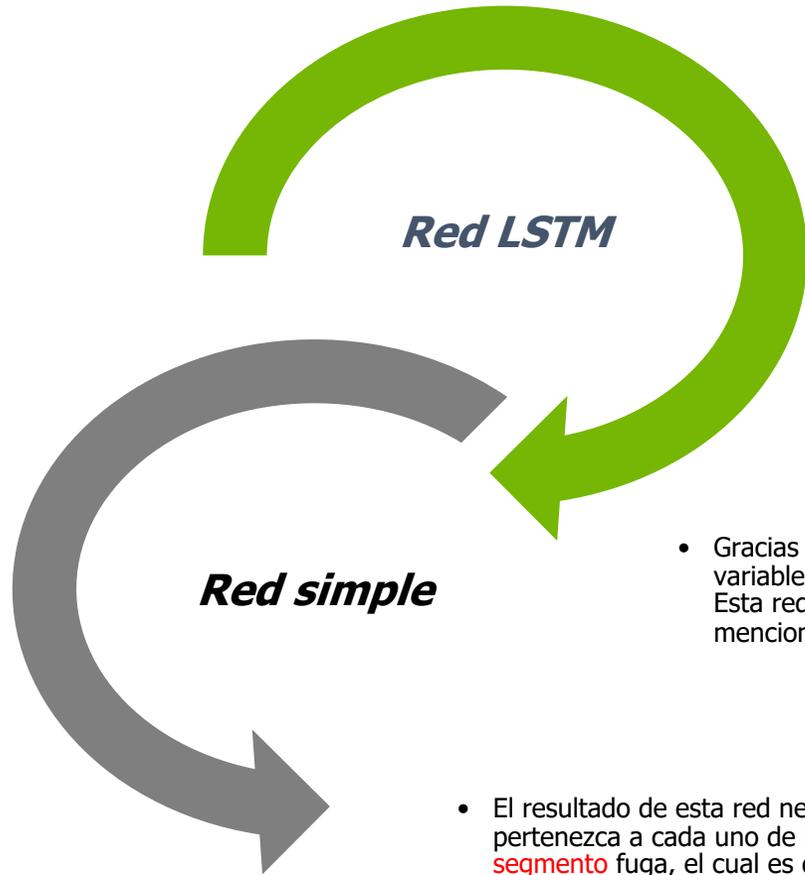
Variables como el precio de la vivienda, nivel de estudios, etc. por sección censal alimentan esta parte. También se cuenta con información demográfica dada por el cliente.





- Una red neuronal donde entran las series temporales de las variables económicas. El resultado obtenido es la estimación de las mismas en el próximo periodo.

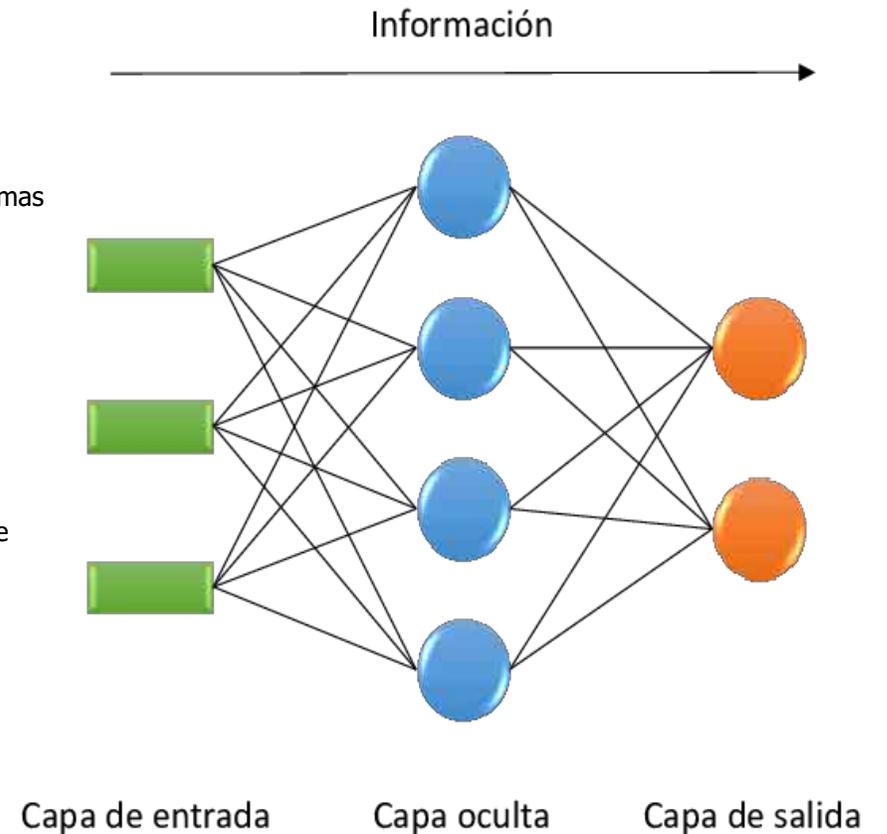




- Una red neuronal donde entran las series temporales de las variables económicas. El resultado obtenido es la estimación de las mismas en el próximo periodo.

- Gracias a la red anterior, podemos unificar el número de variables que tendrán entrada en esta nueva red neuronal. Esta red se ha entrenado con todas las variables anteriormente mencionadas para ver el comportamiento de los clientes.

- El resultado de esta red neuronal es la probabilidad estimada de que el cliente pertenezca a cada uno de los segmentos en el próximo periodo, incluyendo el **segmento** fuga, el cual es que el cliente abandone la entidad.



```

library(keras)

load("~/entrenamiento/datos/ent_6_lstm.rda")

model <- keras_model_sequential() %>%
  bidirectional(layer_lstm(units=72, input_shape = c(numb_col/6, 6),
    return_sequences=TRUE, dropout = 0.4, use_bias=TRUE)) %>%
  bidirectional(layer_gru(units = 32, dropout = 0.2, recurrent_dropout = 0.5,
    return_sequences = TRUE)) %>%
  layer_lstm(units=42, use_bias=TRUE, dropout = 0.4) %>%
  layer_dense(units=12, activation="linear")

# Compiling the model
model %>% compile(
  loss = "logcosh",
  optimizer = "sgd",
  metrics = "mean_squared_error"
)

```

```

ct.R x red_6_lstm.R x red_6_dense.R x
Source on Save Run
library(keras)

load("~/entrenamiento/datos/ent_6_dense.rda")

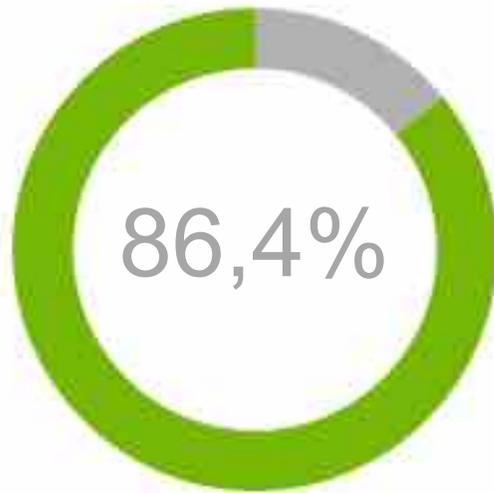
model <- keras_model_sequential() %>%
  layer_dense(units = 15, activation = "relu", input_shape = c(num_col)) %>%
  layer_dropout(rate = 0.4) %>%
  layer_dense(units = 25, activation = "relu") %>%
  layer_dropout(rate = 0.4) %>%
  layer_dense(units = 32, activation = "relu") %>%
  layer_dropout(rate = 0.4) %>%
  layer_dense(units = 26, activation = "relu") %>%
  layer_dropout(rate = 0.4) %>%
  layer_dense(units = 20, activation = "sigmoid")

# Compiling the model
model %>% compile(
  loss = "binary_crossentropy",
  optimizer = "adam",
  metrics = "accuracy"
)

# Fitting loop
model %>% fit(
  x = x_train,
  y = y_train,
  batch_size = 300,
  epochs = 500.

```

03. Algoritmos predictivos



Última fecha de predicción realizada para : **Noviembre de 2018** Muestre sólo clientes con cambio de grupo predicho [XLSX](#) [CSV](#)

Mostrar registros Buscar:

ID Interno	Fecha de alta	Mes de predicción	Fecha de validez	Grupo	Grupo predicho	Porcentaje de predicción	Porcentaje de fuga
<input type="text" value="Todo"/>	<input type="text" value="Todo"/>	<input type="text" value="Todo"/>	<input type="text" value="Todo"/>	<input type="text" value="Todo"/>	<input type="text" value="Todo"/>	<input type="text" value="Todo"/>	<input type="text" value="Todo"/>
	Agosto 2018	6	Febrero 2019	Bronce	Plomo	40,1%	1,3%
	Agosto 2018	6	Febrero 2019	Inactivos	Plomo	48,1%	1,9%
	Agosto 2018	6	Febrero 2019	Inactivos	Plomo	47,6%	2,1%
	Agosto 2018	6	Febrero 2019	Inactivos	Plomo	43,9%	2,7%
	Agosto 2018	6	Febrero 2019	Plomo	Bronce	34,4%	2,1%

Mostrando registros del 1 al 5 de un total de 1.243 registros

Anterior **1** 2 3 4 5 ... 268 Siguiente



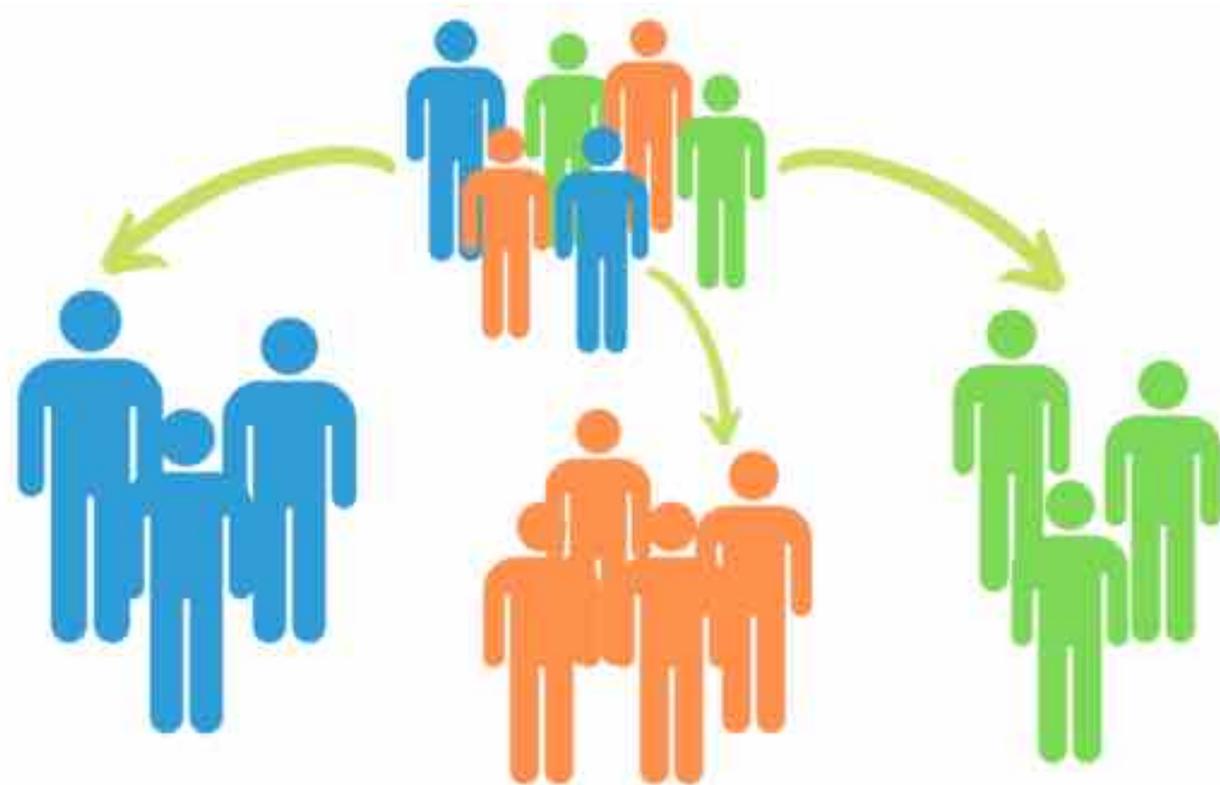
Predicción completa para cliente 

Oro	30,22%
Plata	16,98%
Bronce	31,88%
Plomo	16,8%
Inactivo	0,1%
Marginal	0,67%
Fuga	0,75%

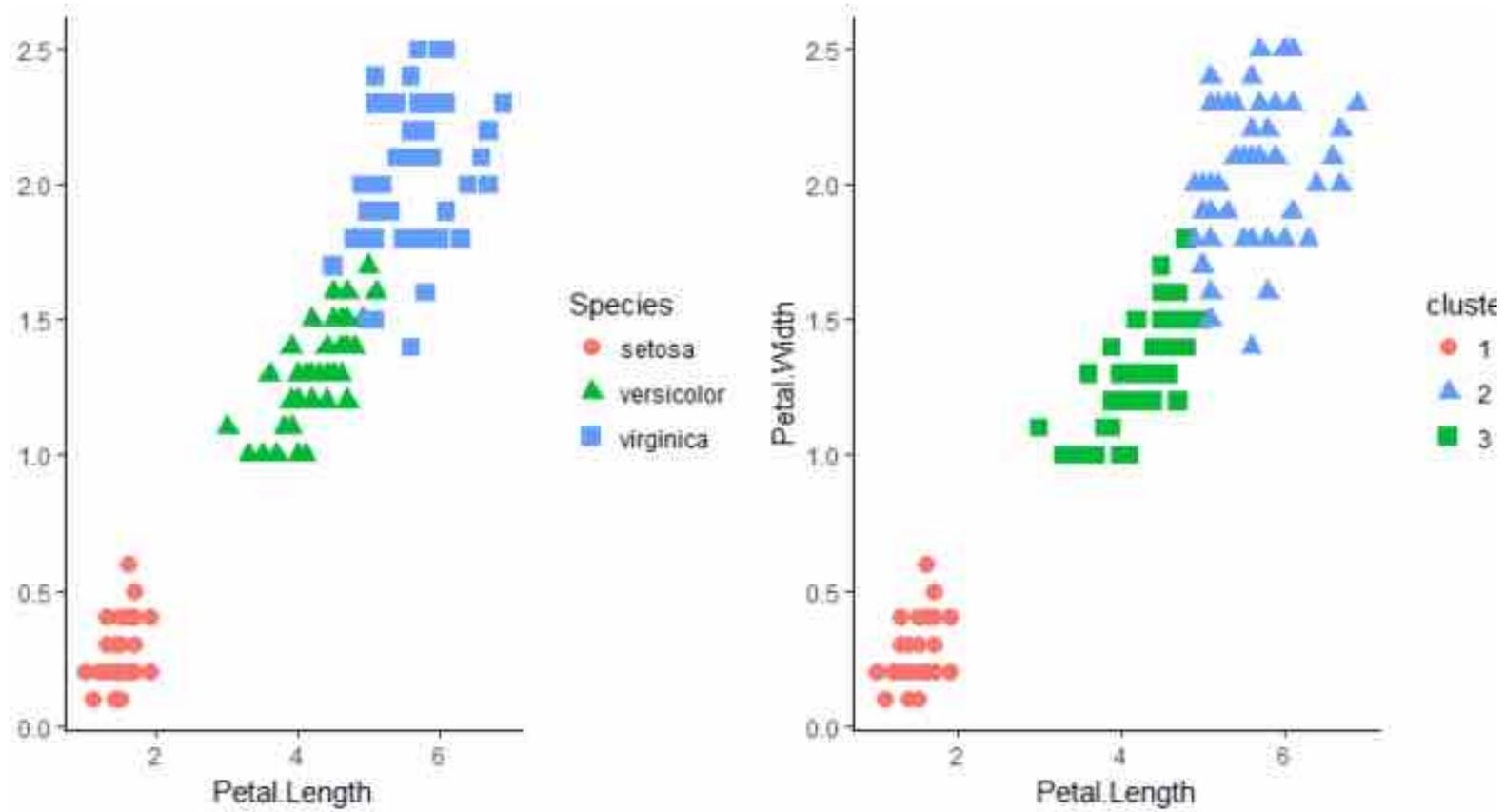
Cerrar

04. ALGORITMOS DE SEGMENTACIÓN

04. Algoritmos de segmentación

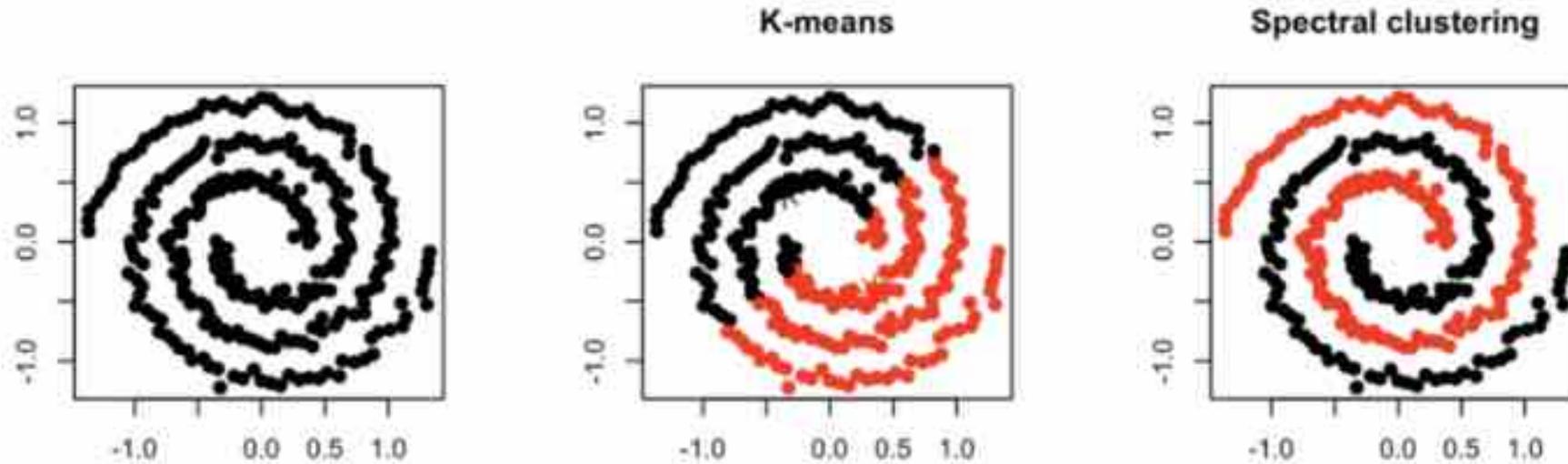


K-means

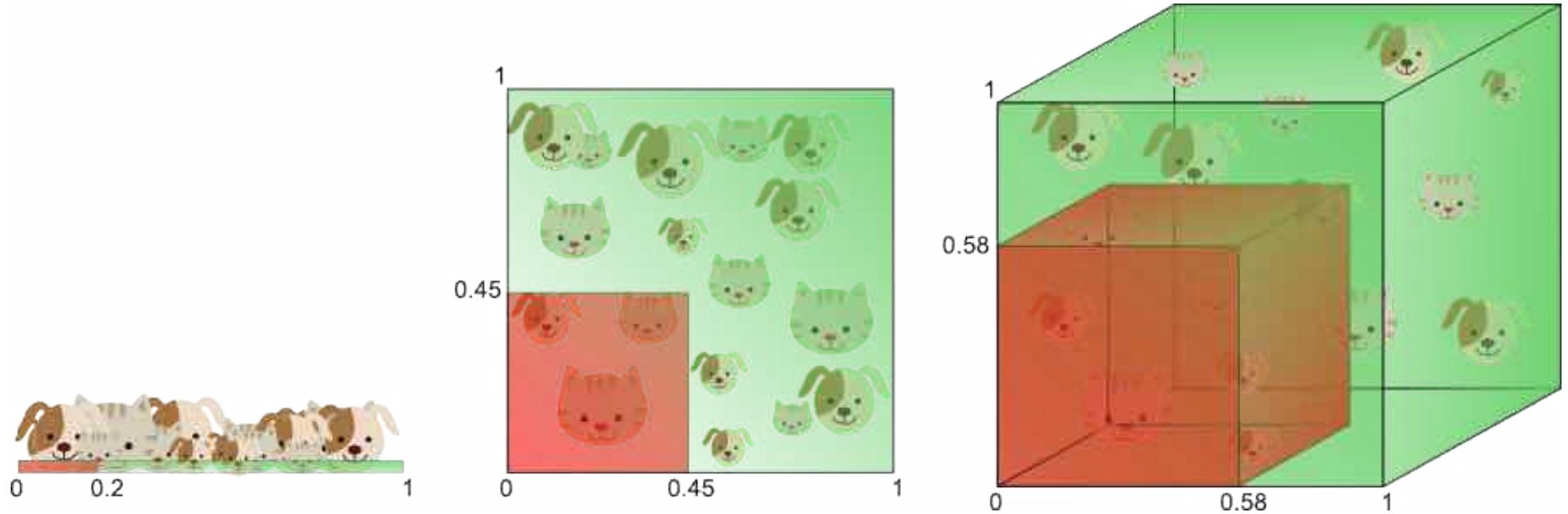


Spectral Clustering

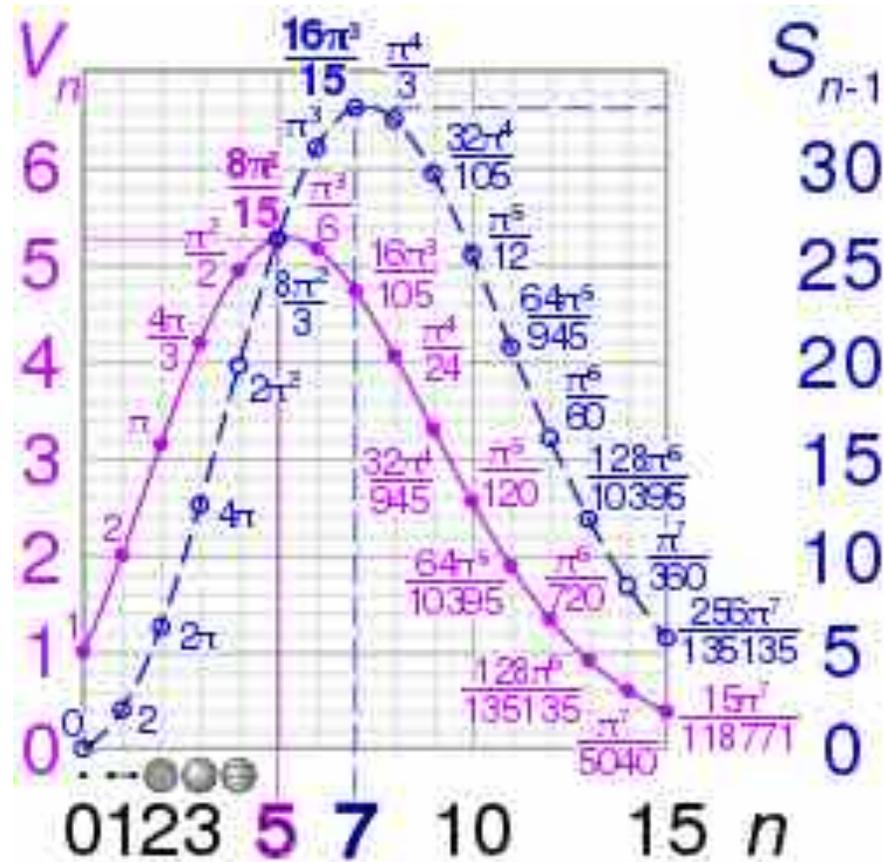
The difference between the 2 can easily be shown by this illustration:



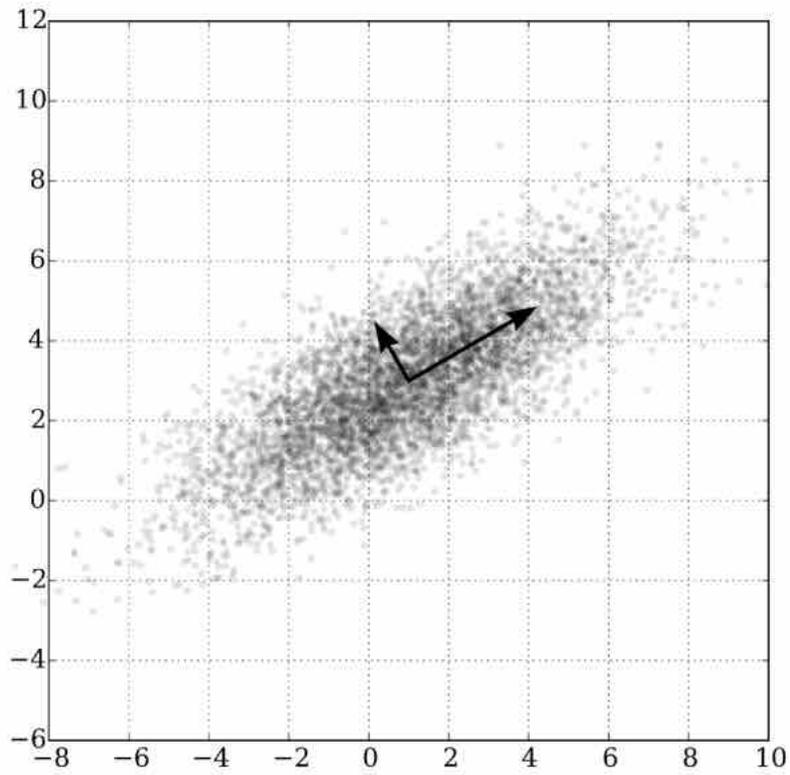
Maldición de la dimensionalidad



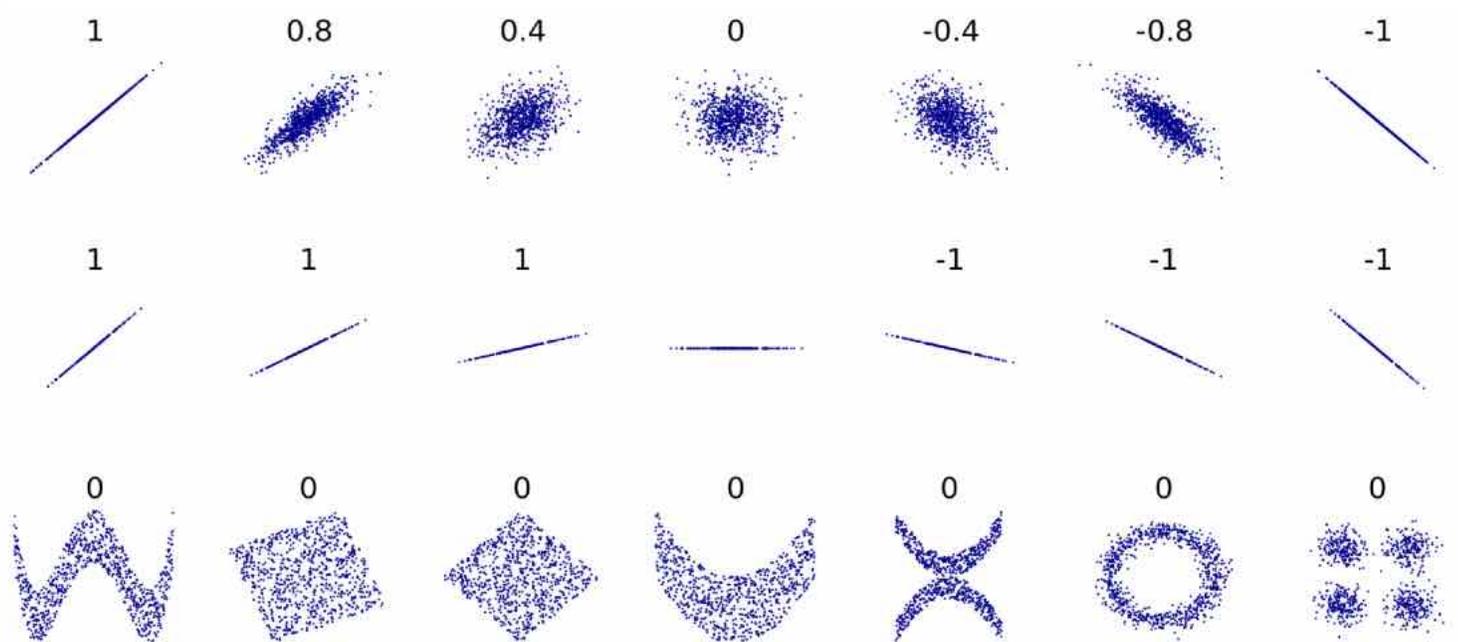
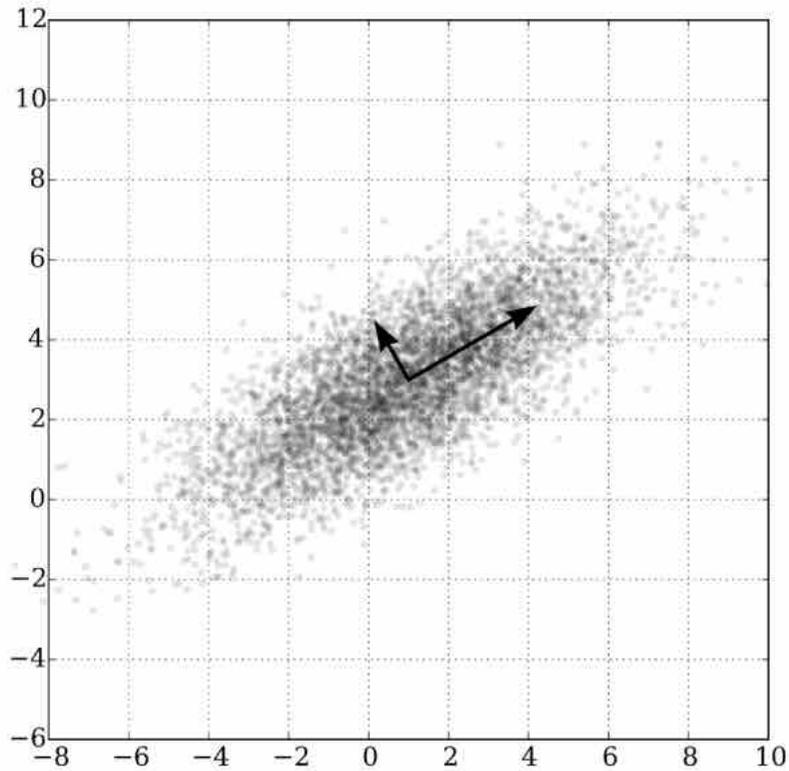
Maldición de la dimensionalidad



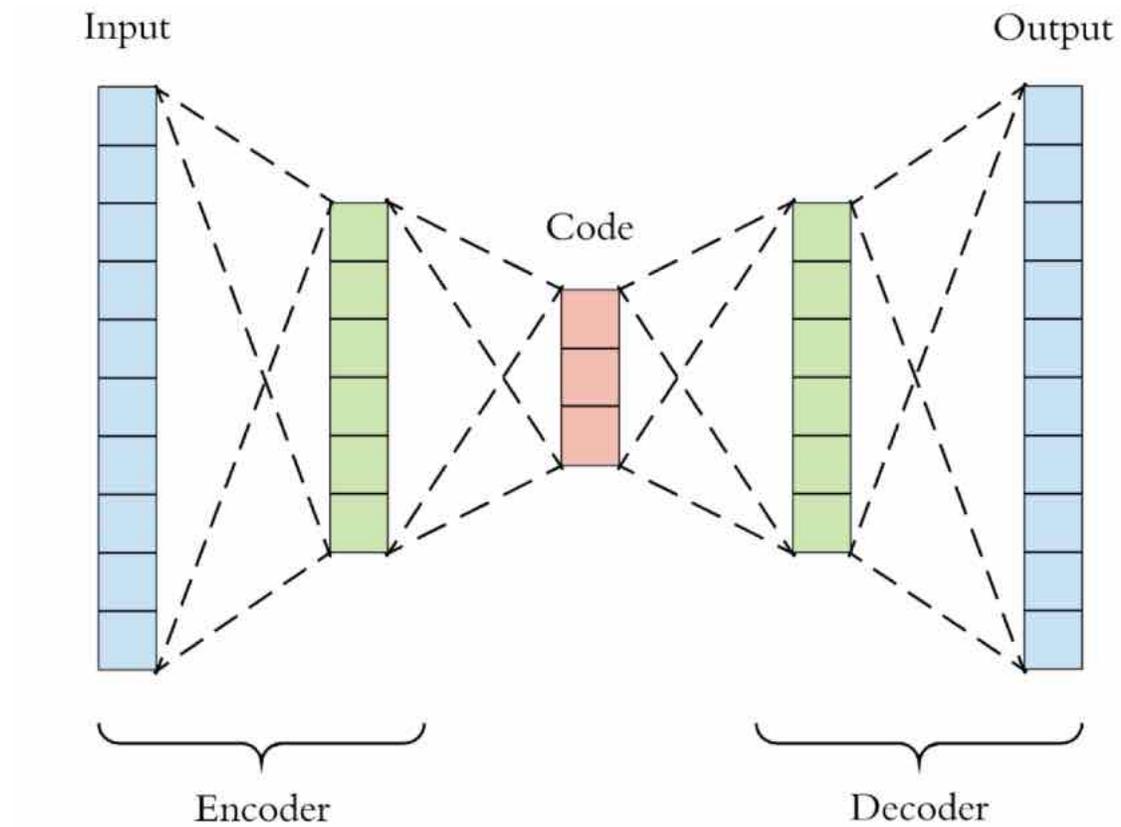
Reducción dimensional: PCA



Reducción dimensional: PCA



Autoencoder



Resultados

Clustering espectral

Resultados

Con todos los valores probados, un gran clúster principal y pequeños clúster con menos de 10 elementos (posibles outliers)

Razones del fracaso

- Mala elección del kernel (otra función, otro parámetro gamma)
- Algunos elementos viven en las fronteras entre segmentos

k-means

Resultados

Segmentaciones con:

- $k = 8$. A partir del método del codo
- $k = 53$. A partir de la estadística de gap

Evaluación de negocio

- Similitud con los grupos definidos por negocio, pero que incorporaba nueva información.
- Faltaría evaluación en un modelo predictivo.

05. CONCLUSIONES

Muchas gracias

[José Luis Huerta Fernández](#)

[José Manuel Roperro Tagua](#)

