



BIG BANK

Febrero 2020

Equipo Super López

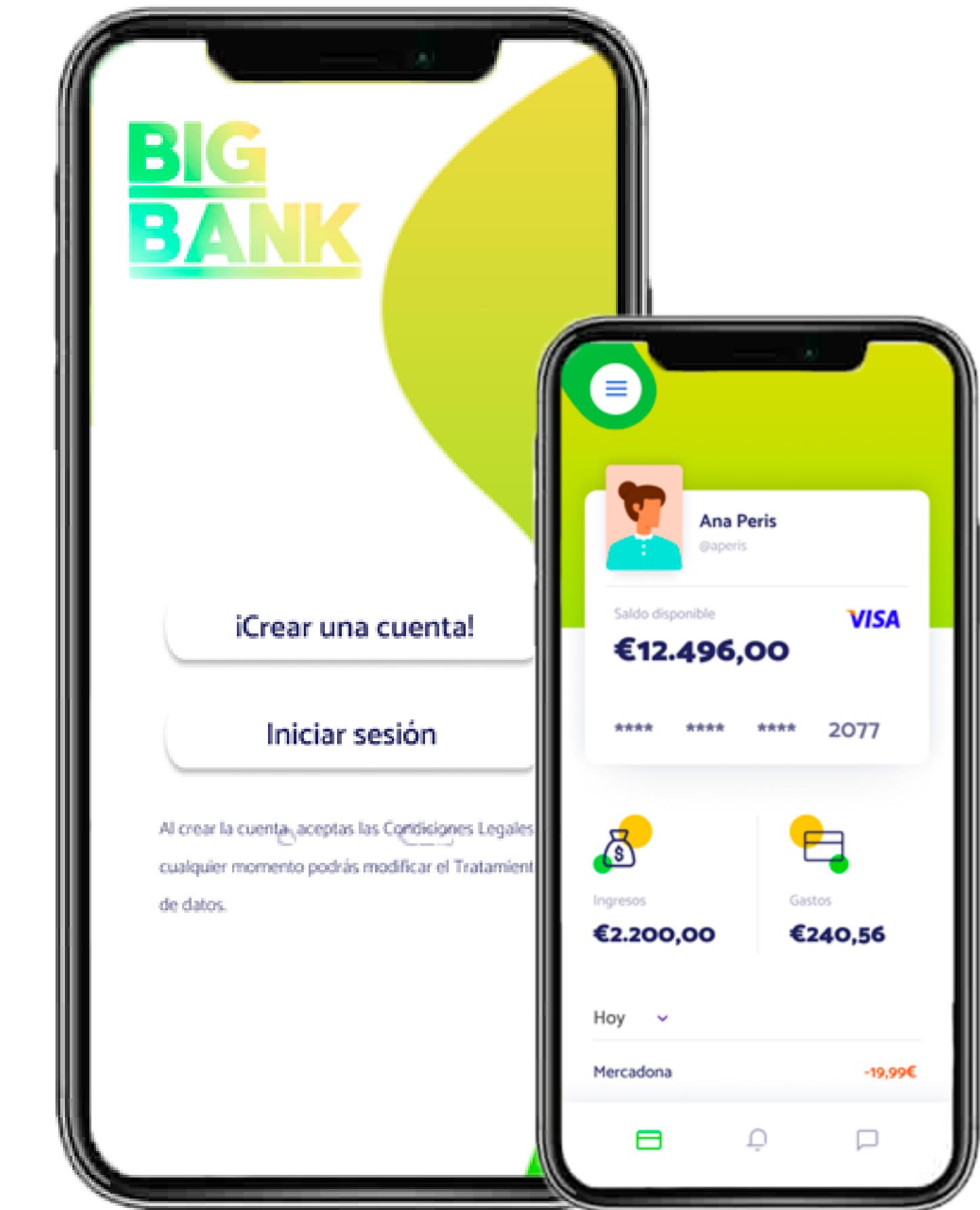


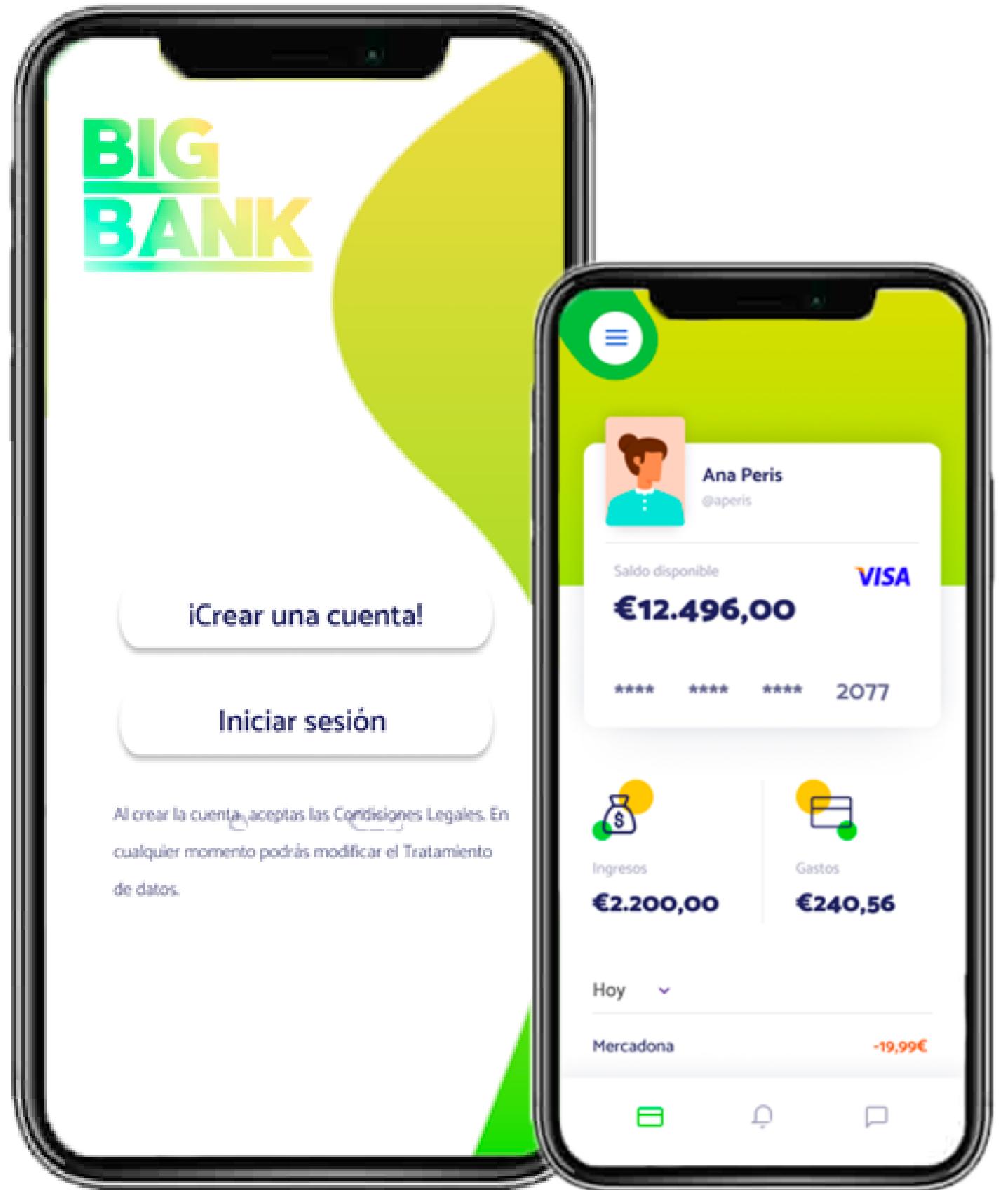
Rocío Cort
Ignacio Costa
Javier de Toro
Sergio Díez
Victor Malvar

Índice

- Porqué
- Meta
- Descripción casos de uso
- Campaña de marketing
contratación a plazo fijo
- Viabilidad para la concesión
de un préstamo personal
- Propuesta de valor







Agilidad

En captación de clientes y análisis de riesgos

Customer Centric

Personalizar y adecuar la oferta

Competitividad

Océano rojo de la mano de la tecnología

Eficiencia

Reducir costes de plantilla e inmovilizado

Casos de uso

1

Campaña de marketing contratación a plazo fijo

Medir la efectividad de una campaña de marketing para conseguir que el cliente realice un depósito a plazo fijo en la entidad.

2

Viabilidad para la concesión de un préstamo personal

Predecir si un cliente es un buen candidato para recibir un préstamo. Si el préstamo será devuelto por el cliente, o se convertirá en impagado.

Millennials



1

Campaña de marketing contratación a plazo fijo

Pasos:

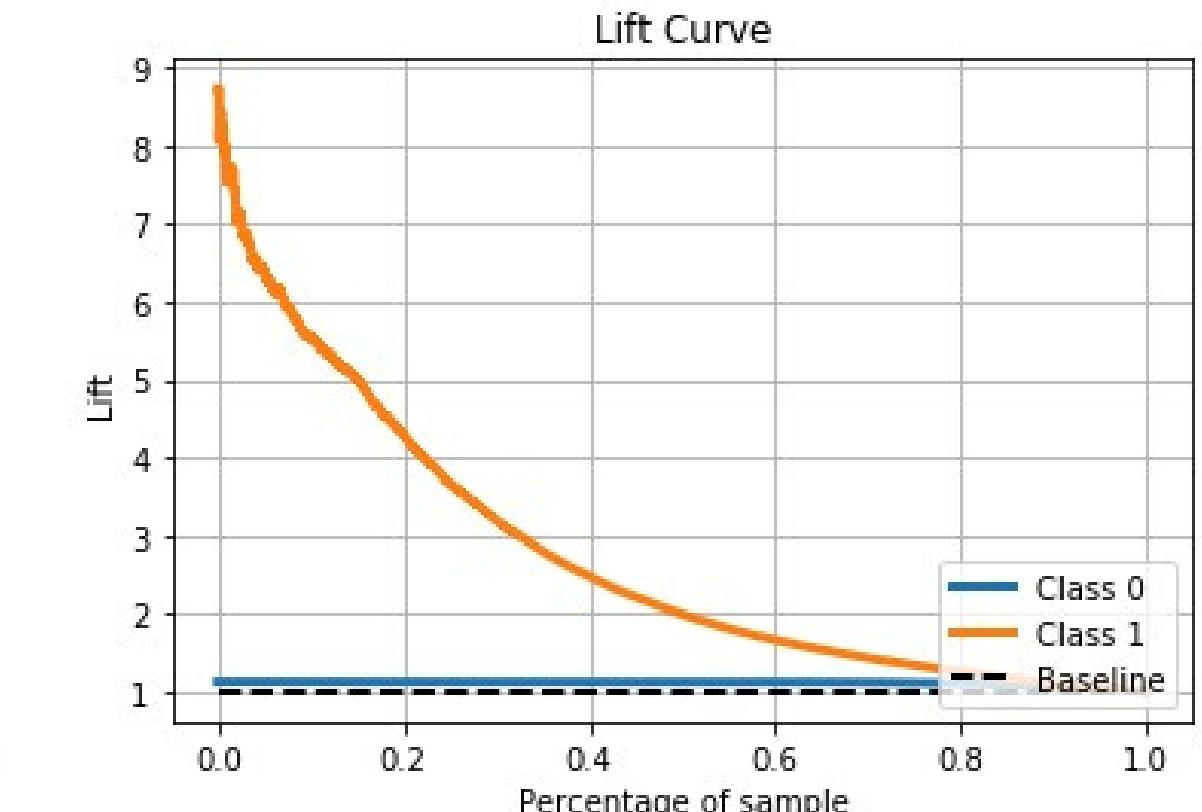
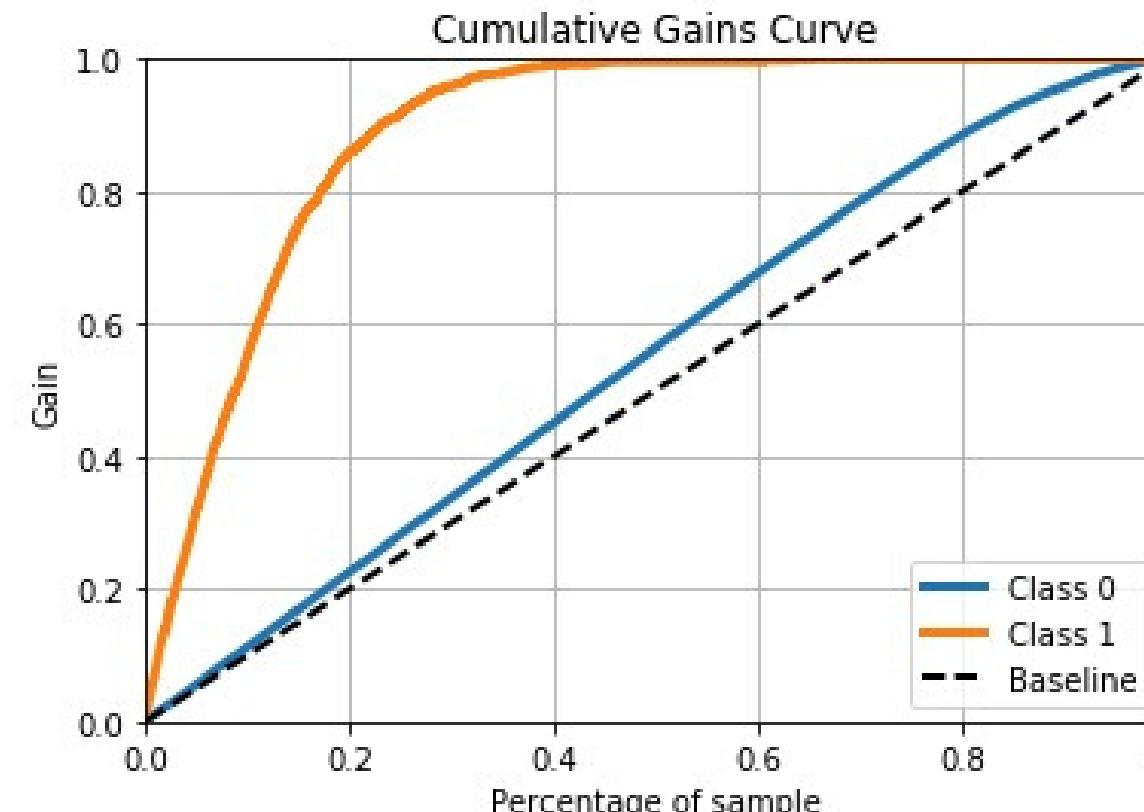
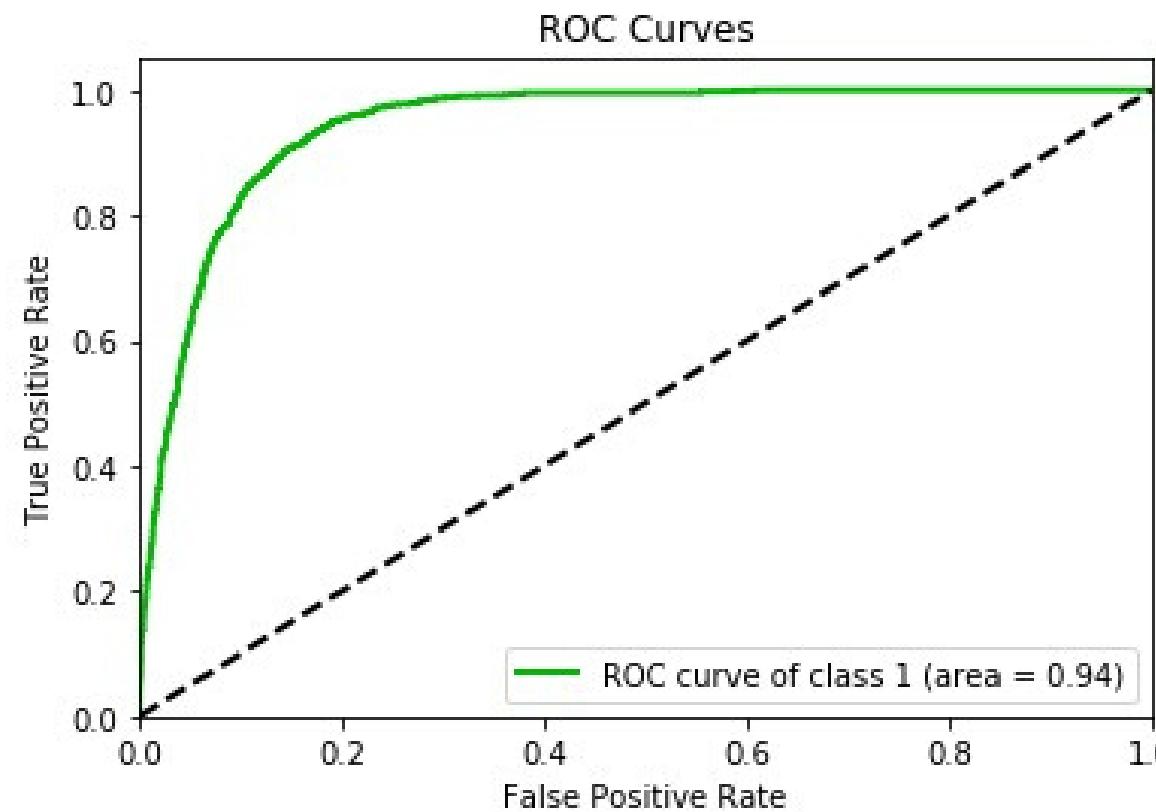
- Optimización del dataset origen:
 - Descripción de variables y categorización
 - Discretización de variables
- Utilización algoritmo Random Forest
- Partición de 80/20 en Training/Evaluación
- 5 iteraciones en validación cruzada
- GridSearch: Iteración de hiperparámetros para elección y afinación del modelo
- Identificación de *millenials* (nacidos entre 1980-1996)
- Clusterización dentro del grupo de *millenials*.



1

Elabora un algoritgmo

- **ROC Curve** - Evalúa la capacidad discriminativa de nuestro modelo
- **Cumulative Gains Curve** - Determina la predicción del modelo vs los resultados obtenidos con una elección aleatoria
- **Lift Curve** - Representa gráficamente la mejora que proporciona el modelo vs. una elección aleatoria



1

Elabora un algoritgmo

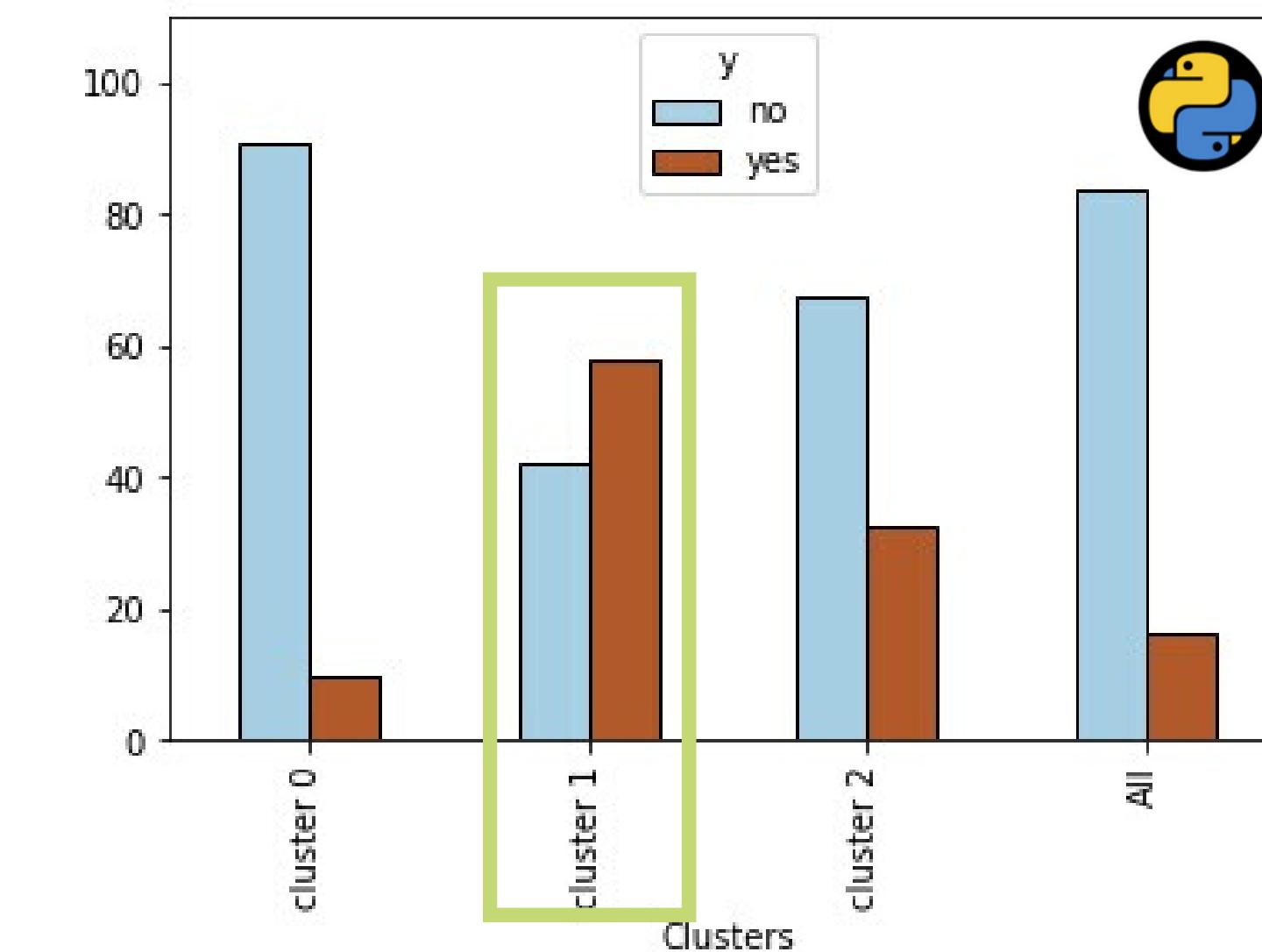
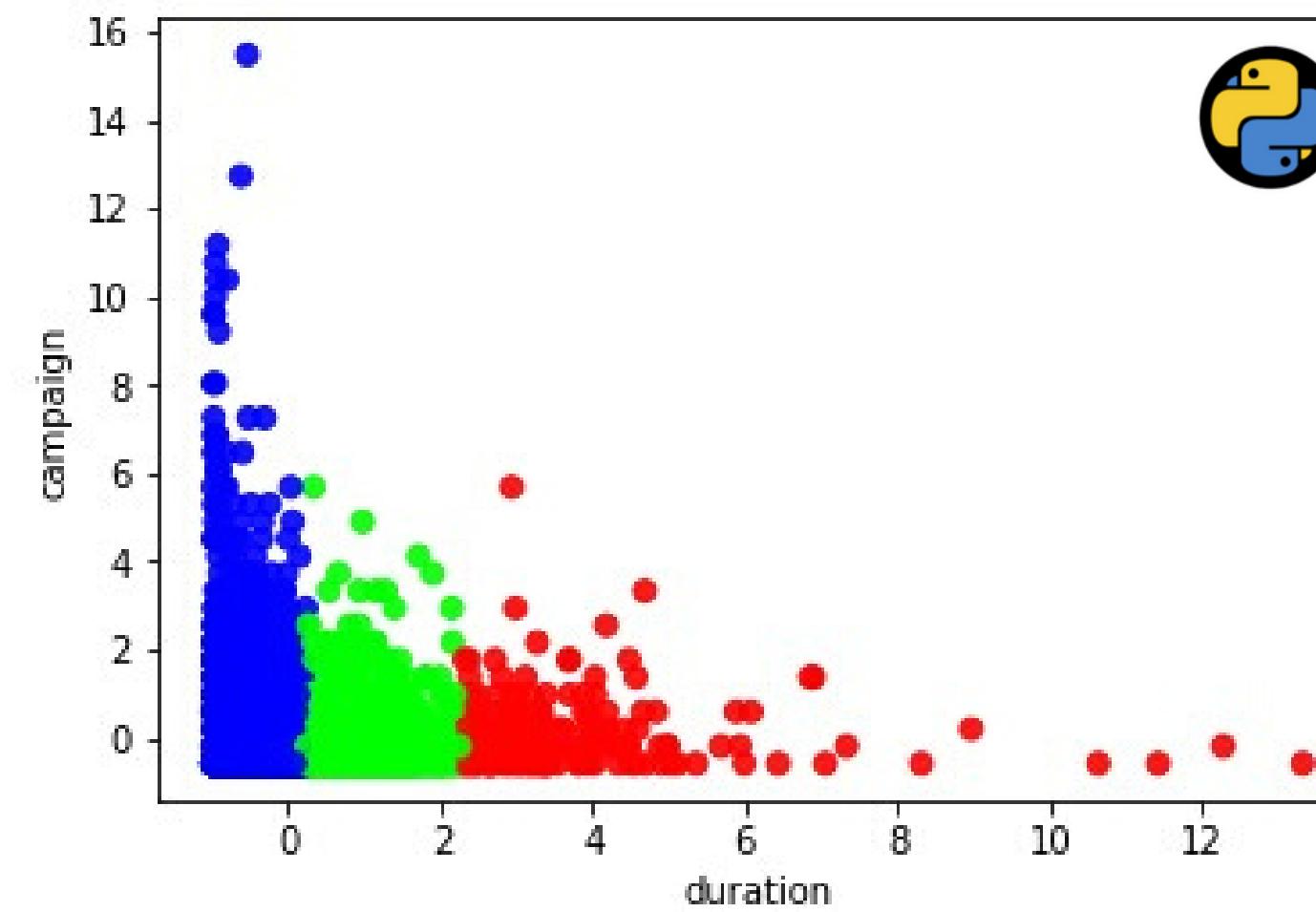
Matriz de confusión

		Actual	
		No (0)	Yes (1)
Accuracy: 91,20%		7142	152
Predicted	0	7142	152
	1	576	368

1

Segmentación de clientes, ¿Para qué nos puede ser útil?

- 3 clusters dentro del **grupo *millenials*** donde las variables más significativas son *Duration*, *Campaign* y *Previous*.
- Conclusión: Los que más tiempo están al telf., que han sido contactados más veces en esta campaña y no han tenido contacto en anteriores son nuestro target



1

Descripción de buyer persona

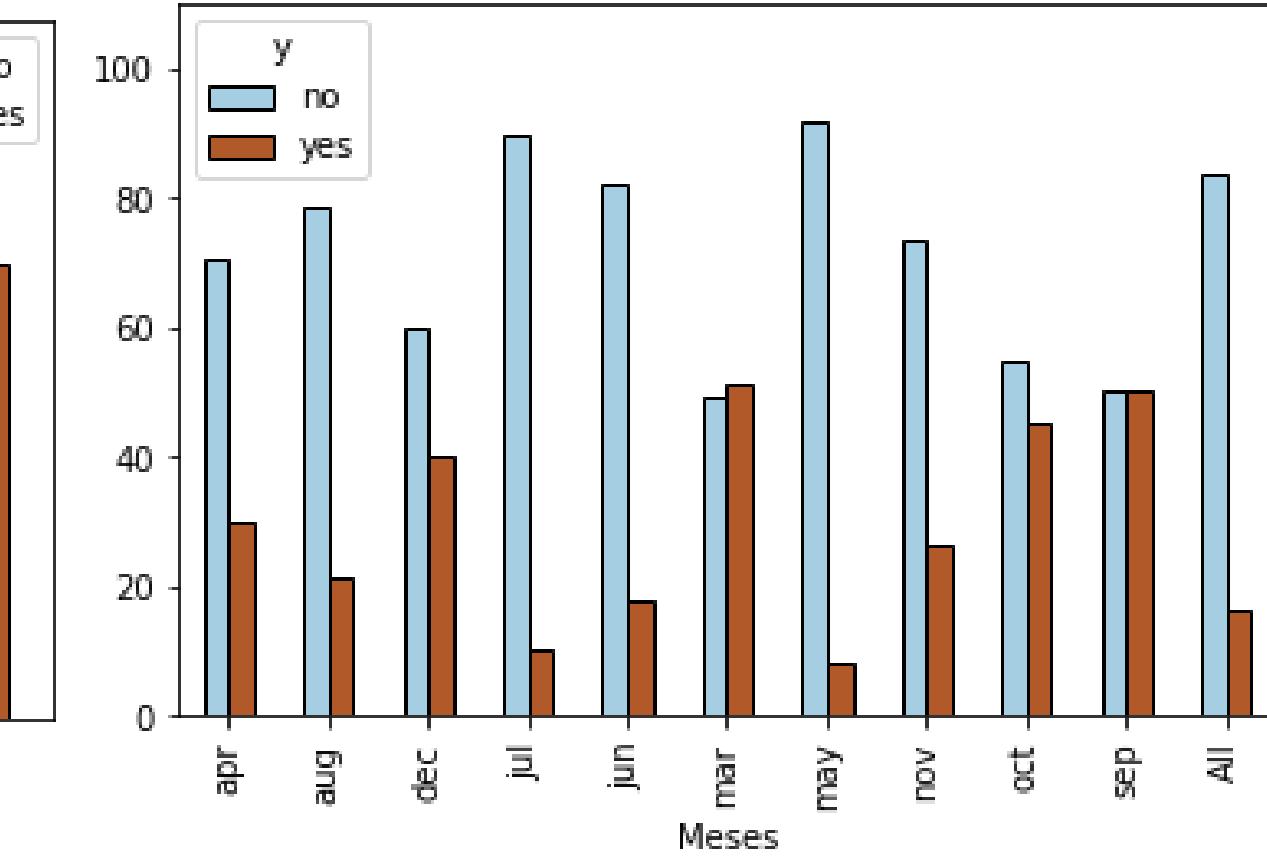
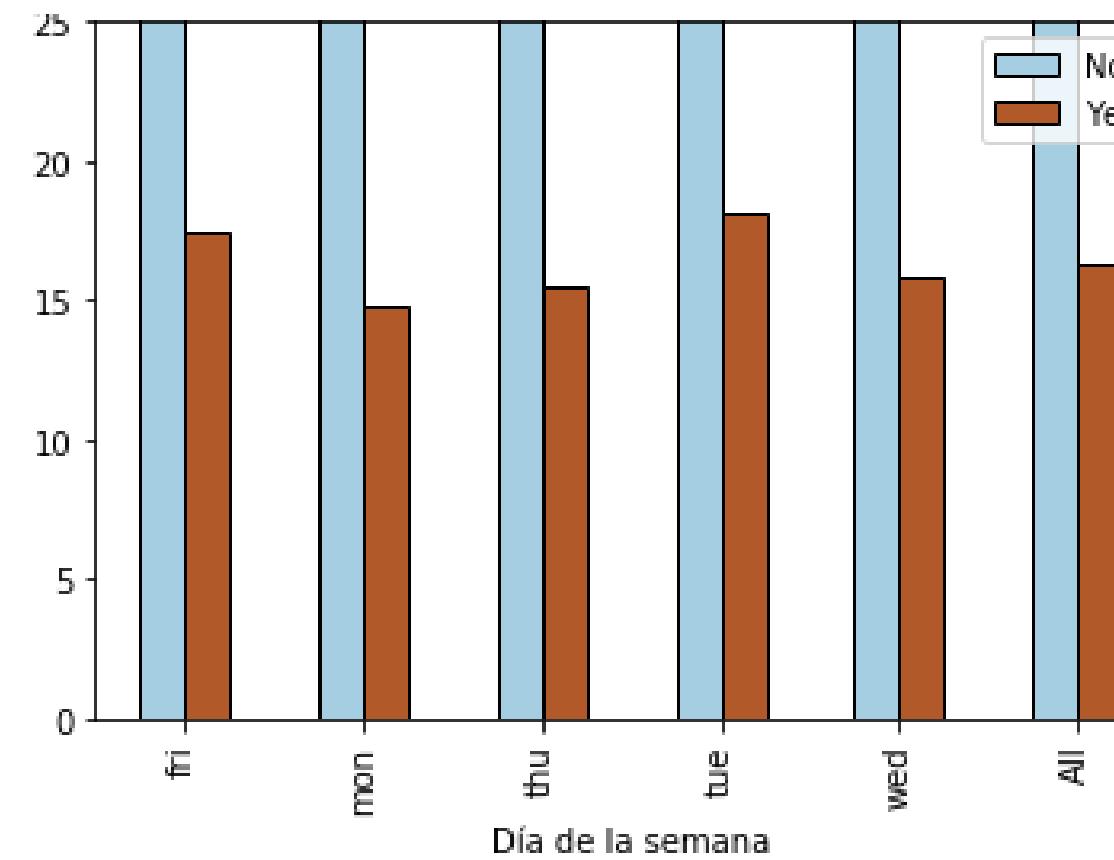
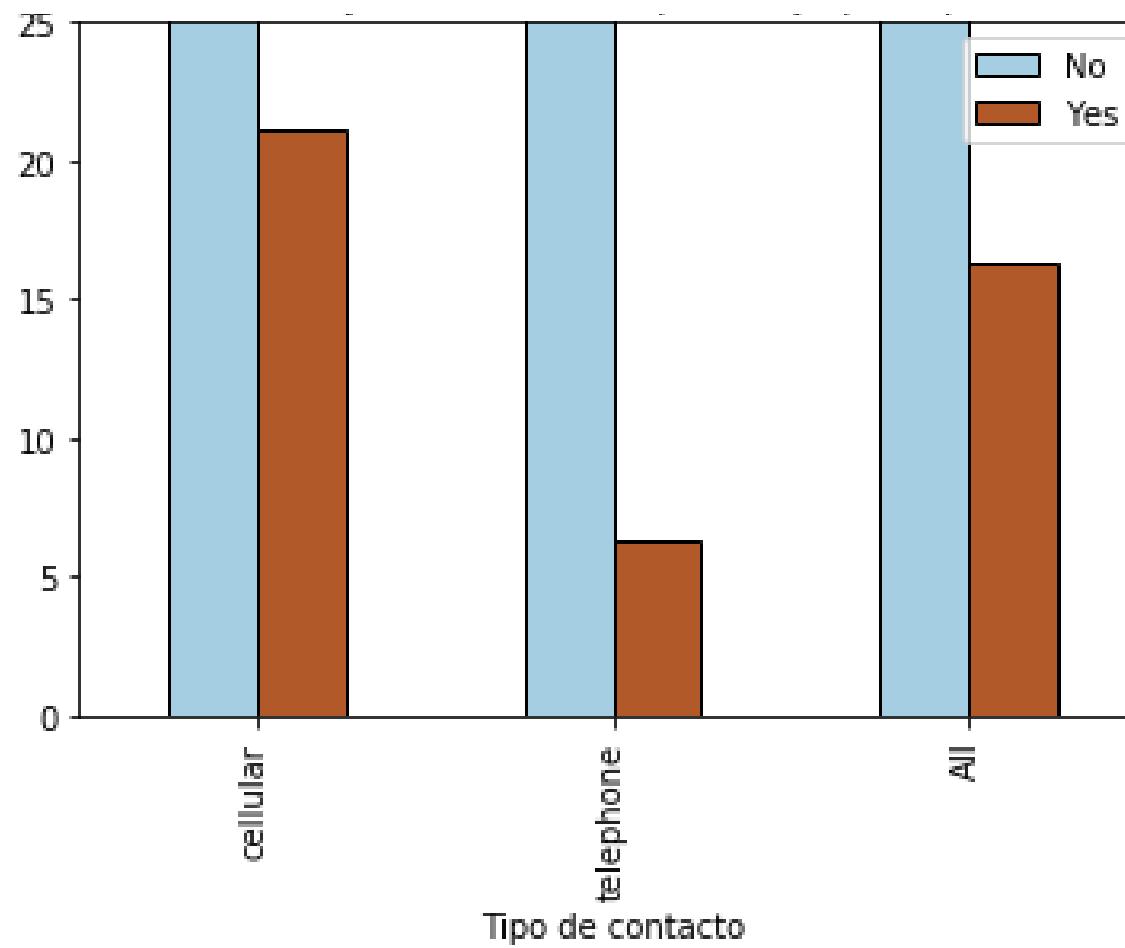


Características

- 26 años
- Trabajan en la administración
- Solteros
- Con educación secundaria
- Tienen casa propia
- No tienen préstamo

Posibles soluciones para la próxima campaña de marketing

- Círculo Clientes del Cluster 1 - **Se recomienda por histórico llamada a móvil, larga, martes y viernes en Marzo, Septiembre y Octubre.**



1

Inversión en Bitcoin

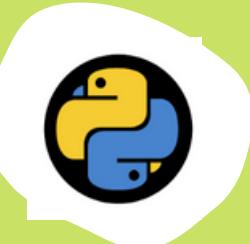
- Los depósitos a plazo fijo son productos “opuestos” a las inversiones en criptomonedas
- El Dataset dado no es adecuado para determinar qué tipos de clientes aceptarían este tipo de inversiones
- Se necesita añadir variables que nos indiquen la aceptación/rechazo al riesgo y determinen el perfil inversor
 - Género
 - Edad
 - Perfil de inversor



2

Viabilidad para la concesión de un préstamo personal

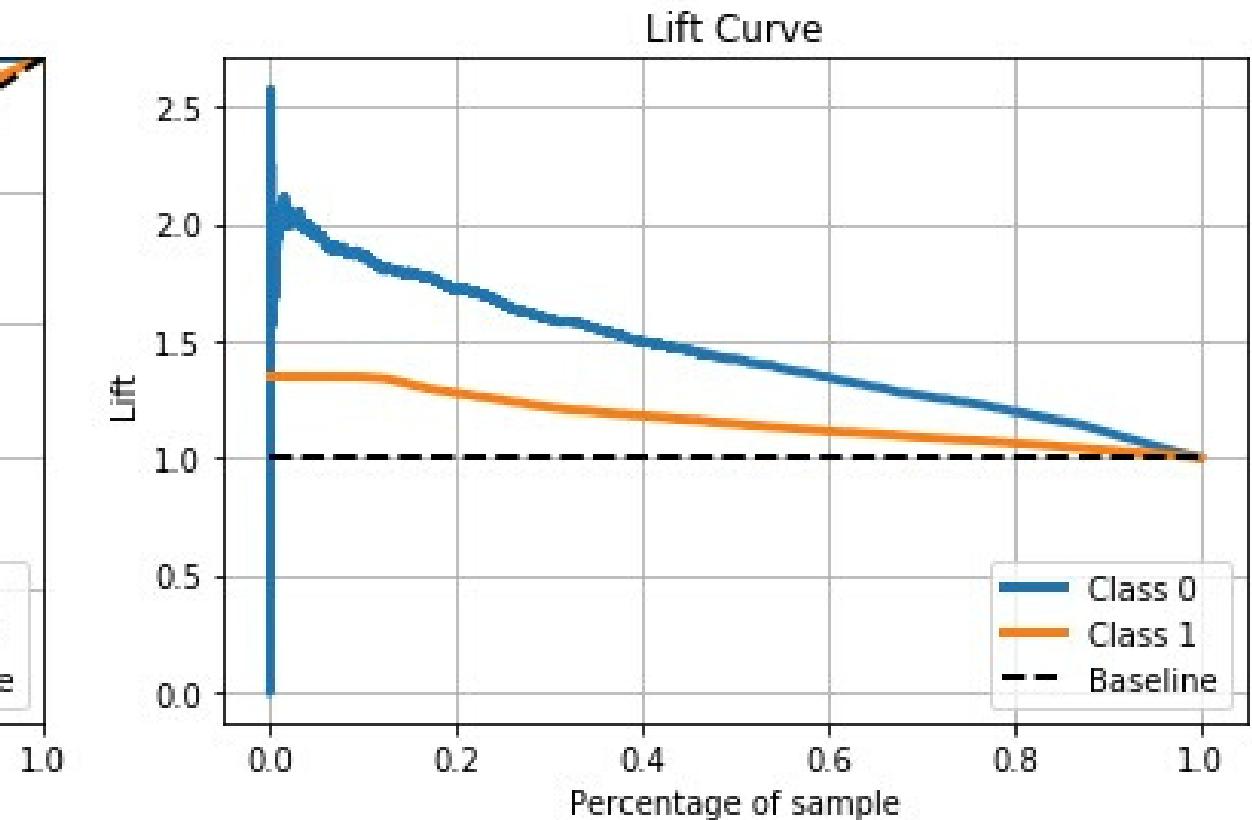
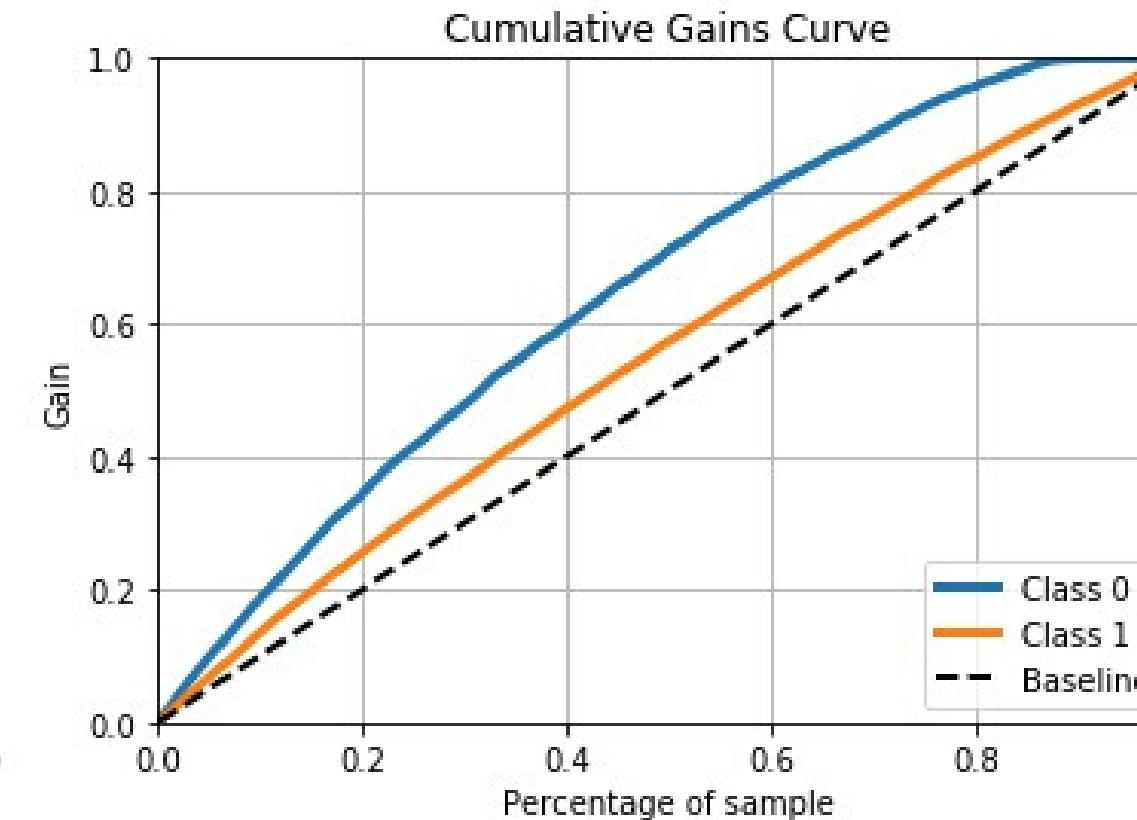
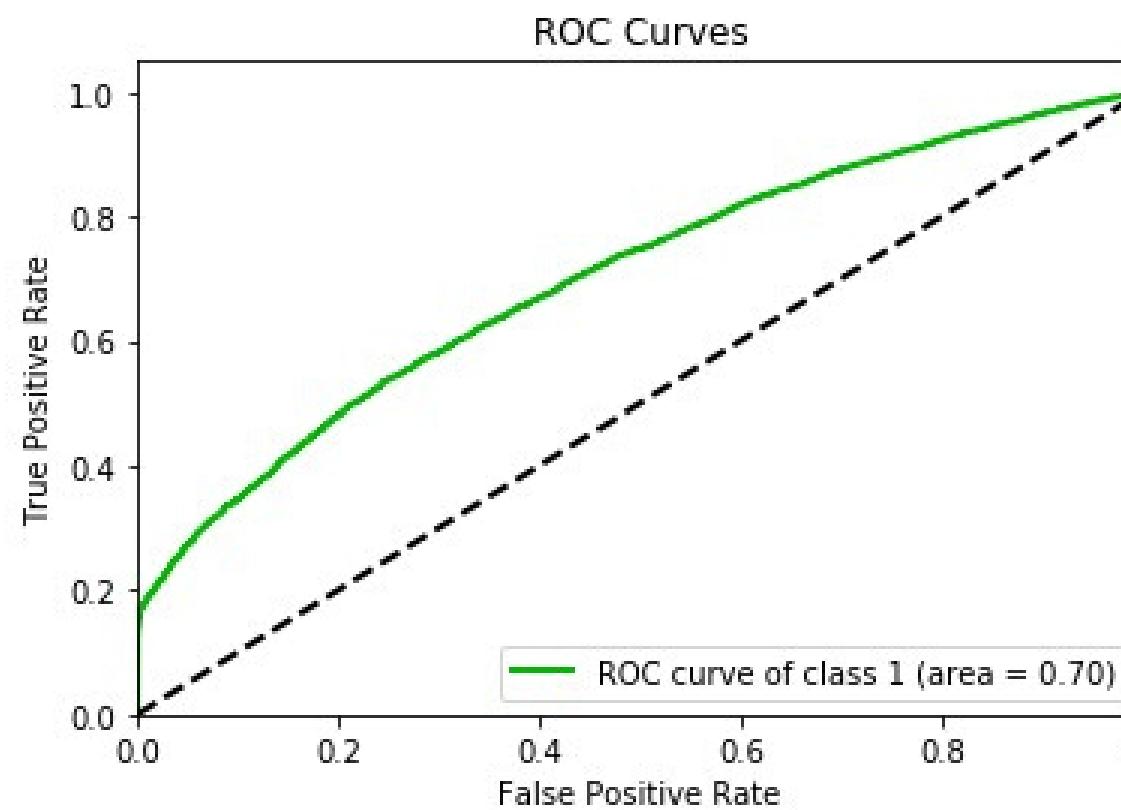
Pasos:



- Optimización del dataset origen:
 - Descripción y eliminación de variables.
 - Completar el Dataset:
 - "Fully paid" para 999999
 - Todos los valores de Credit Score en 3 cifras
 - Reducción de clases (Home Ownership & Purpose)
 - Eliminación de valores MUY anómalos
 - Completar valores perdidos con la media/moda
 - Categorización de variables
- Identificación de variable objetivo: “Loan Status”
- Utilización de Random Forest
- Partición en 70/15/15 en Training/Validación/Evaluación
- GridSearch: Iteración de hiperparámetros para elección y afinación del modelo

Predecir si un cliente es un buen candidato para recibir un préstamo

- **ROC Curve** - Evalúa la capacidad discriminativa de nuestro modelo
- **Cumulative Gains Curve** - Determina la predicción del modelo vs los resultados obtenidos con una elección aleatoria
- **Lift Curve** - Representa gráficamente la mejora que proporciona el modelo vs. una elección aleatoria



2

Predecir si un cliente es un buen candidato para recibir un préstamo

Matriz de confusión

		Actual	
		Charged Off (0)	Fully Paid (1)
Accuracy: 74,20%		376	3051
Predicted	0	376	3051
	1	374	9452

2

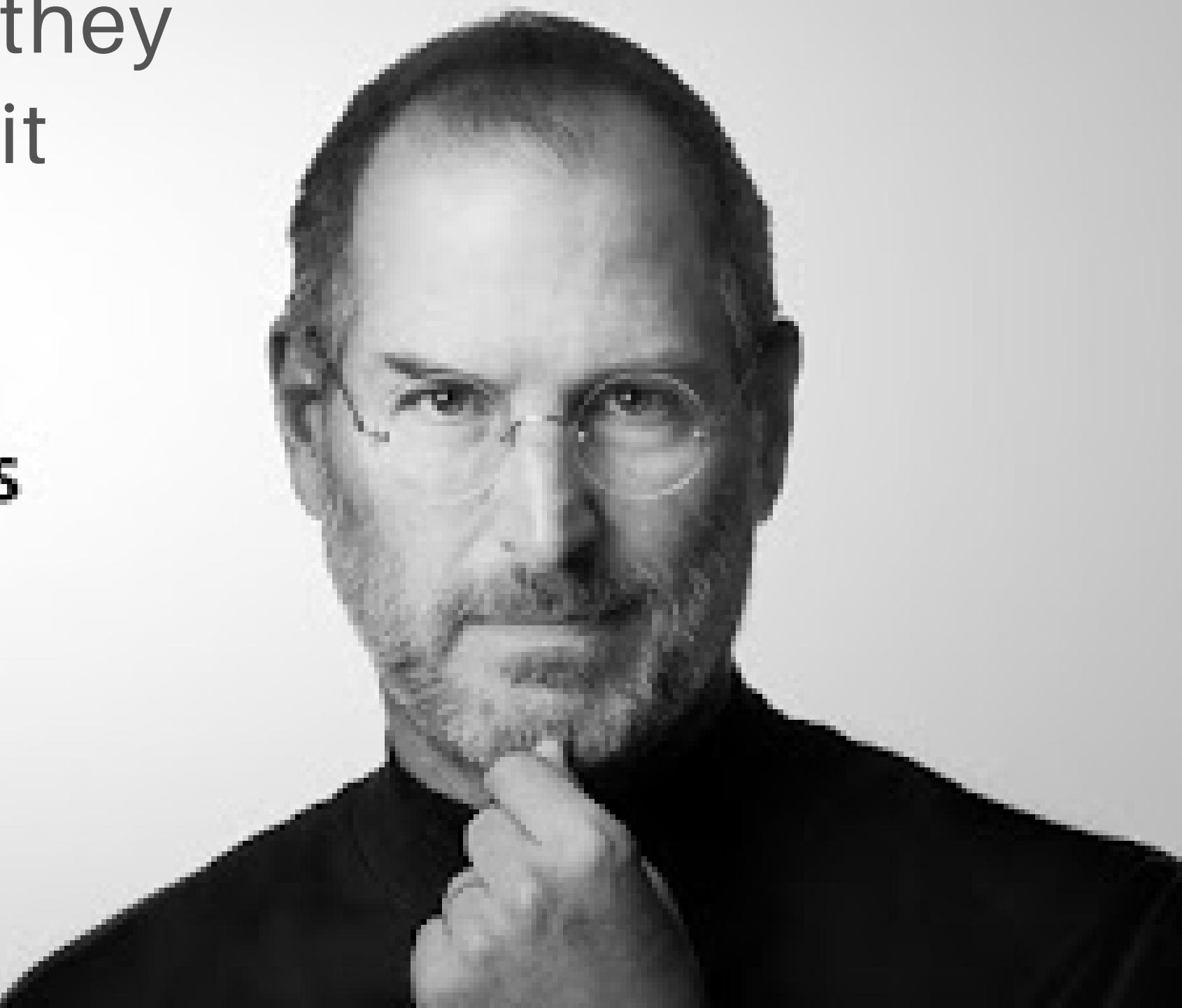
Indicadores que más influyen en un impago

- **Influencia positiva en el pago:** Cantidad deuda actual, Solvencia personal (Score), Deuda a Corto Plazo y Casa Hipotecada
- **Influencia negativa en el pago:** Deuda a Largo Plazo y Casa alquilada



"Get closer than ever to your customer.
So close that you tell them what they
need well before they realize it
themselves"

Steve Jobs
1955-2011





Carmen | 38 años | Profesora de Universidad
Mujer, casada y con 3 hijos

Background

- Paga facturas mensualmente y hace transferencias
- Prefiere el banco online pero acude a oficinas

"Estoy cansada de hacer cola y de que me llegue publicidad que no me interesa"

Motivaciones

- Ascenso y subida de sueldo
- Su hija se casa el año que viene

Frustaciones

- Largas colas
- Le llegan muchos emails de campañas de publicidad



John | 30 años | Diseñador Gráfico

Hombre, soltero

Background

- Gasta el 90% de sus ingresos, no ahorra
- No sabe cómo, ni donde gasta el dinero
- Le han interesado las criptomonedas y ha invertido en ellas

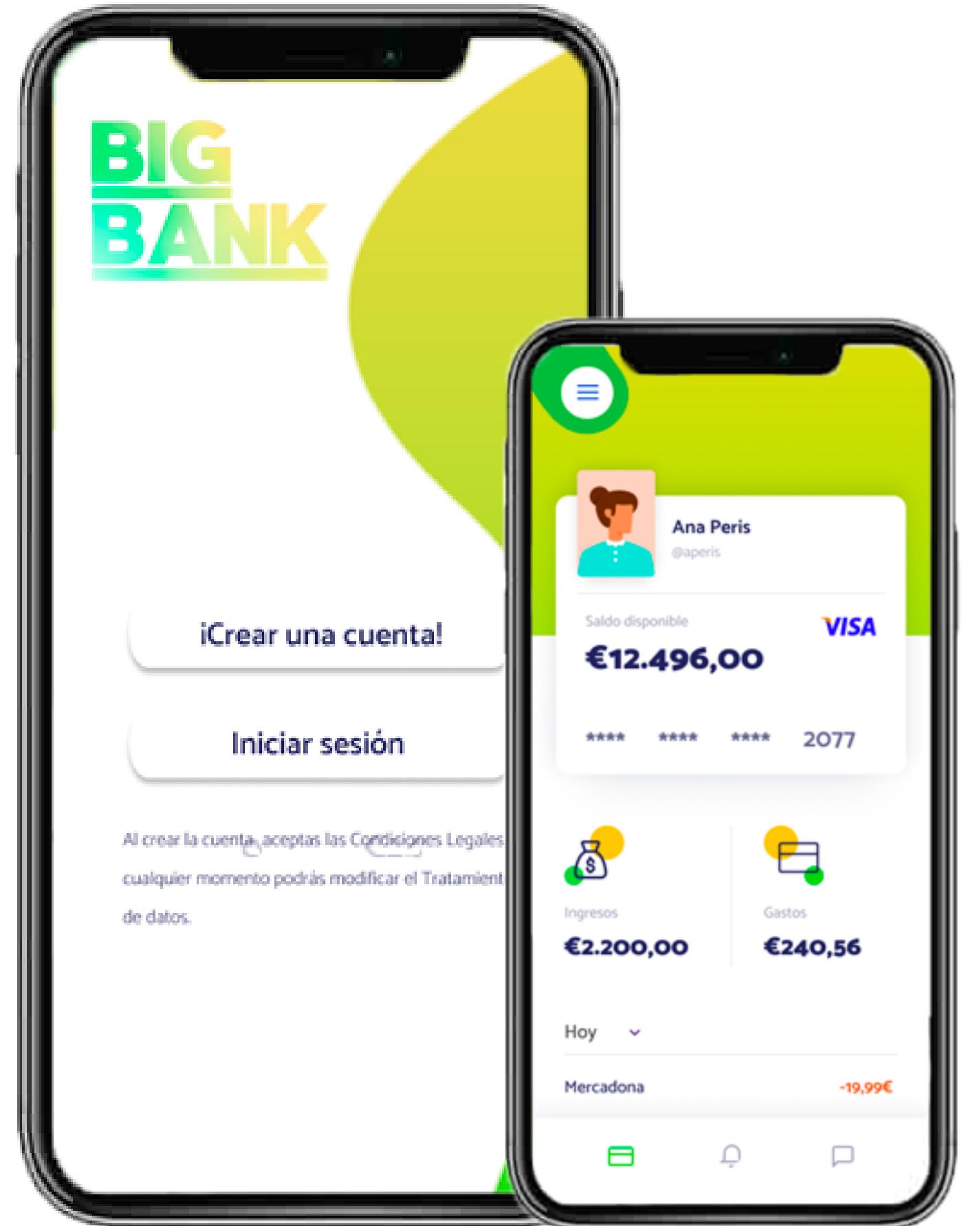
"Tengo que encontrar una manera más óptima de administrar mis finanzas"

Motivaciones

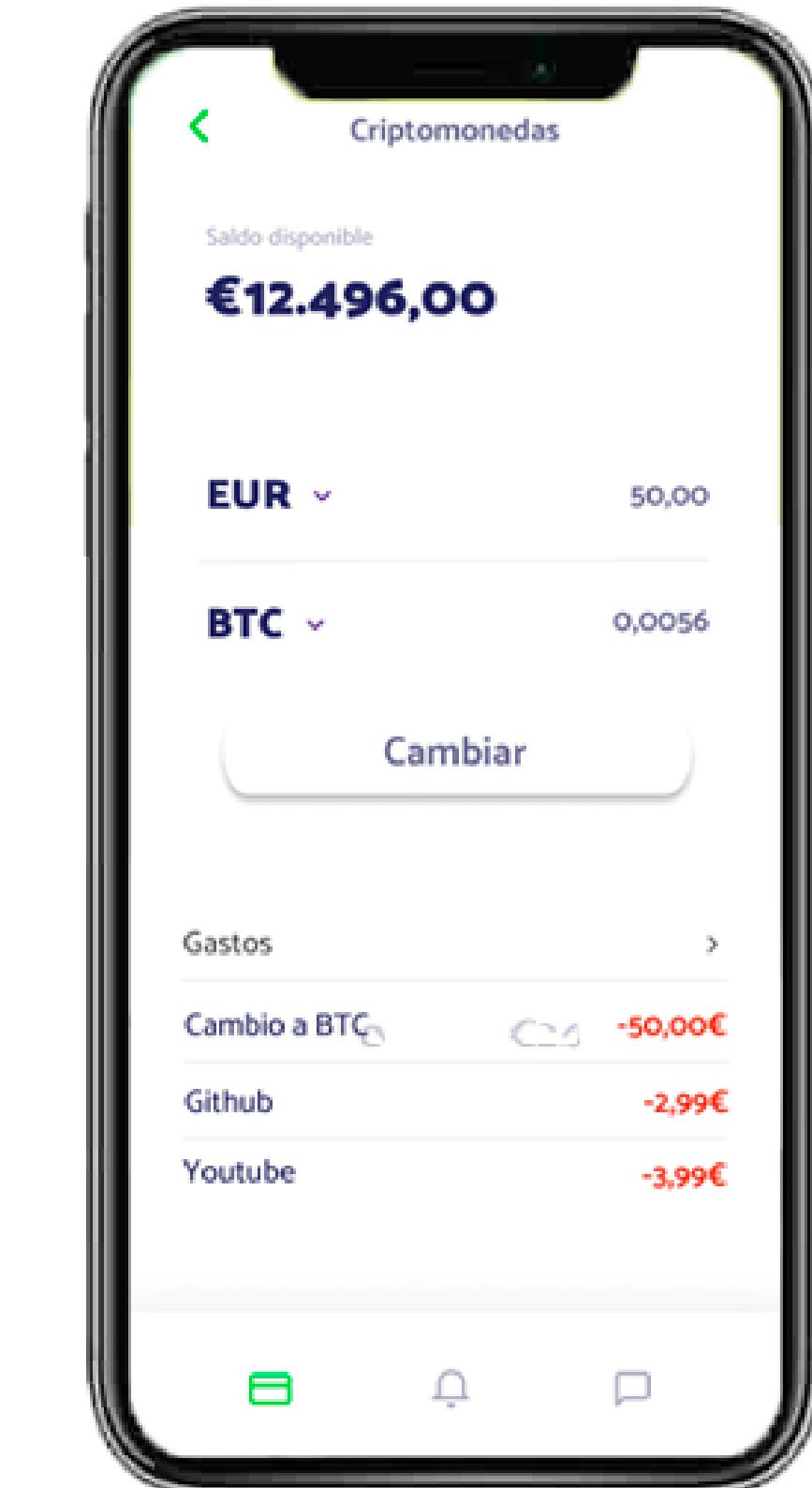
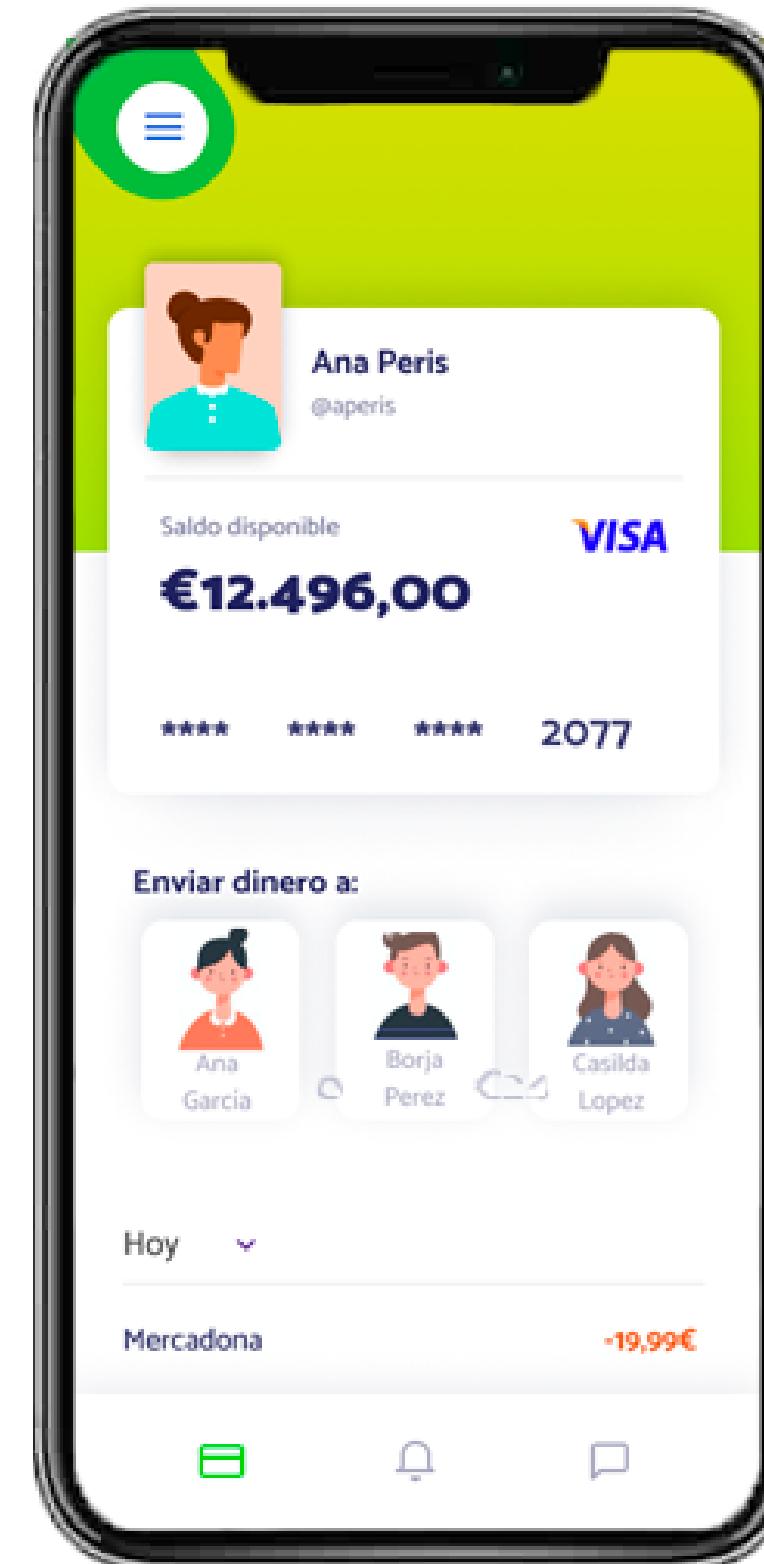
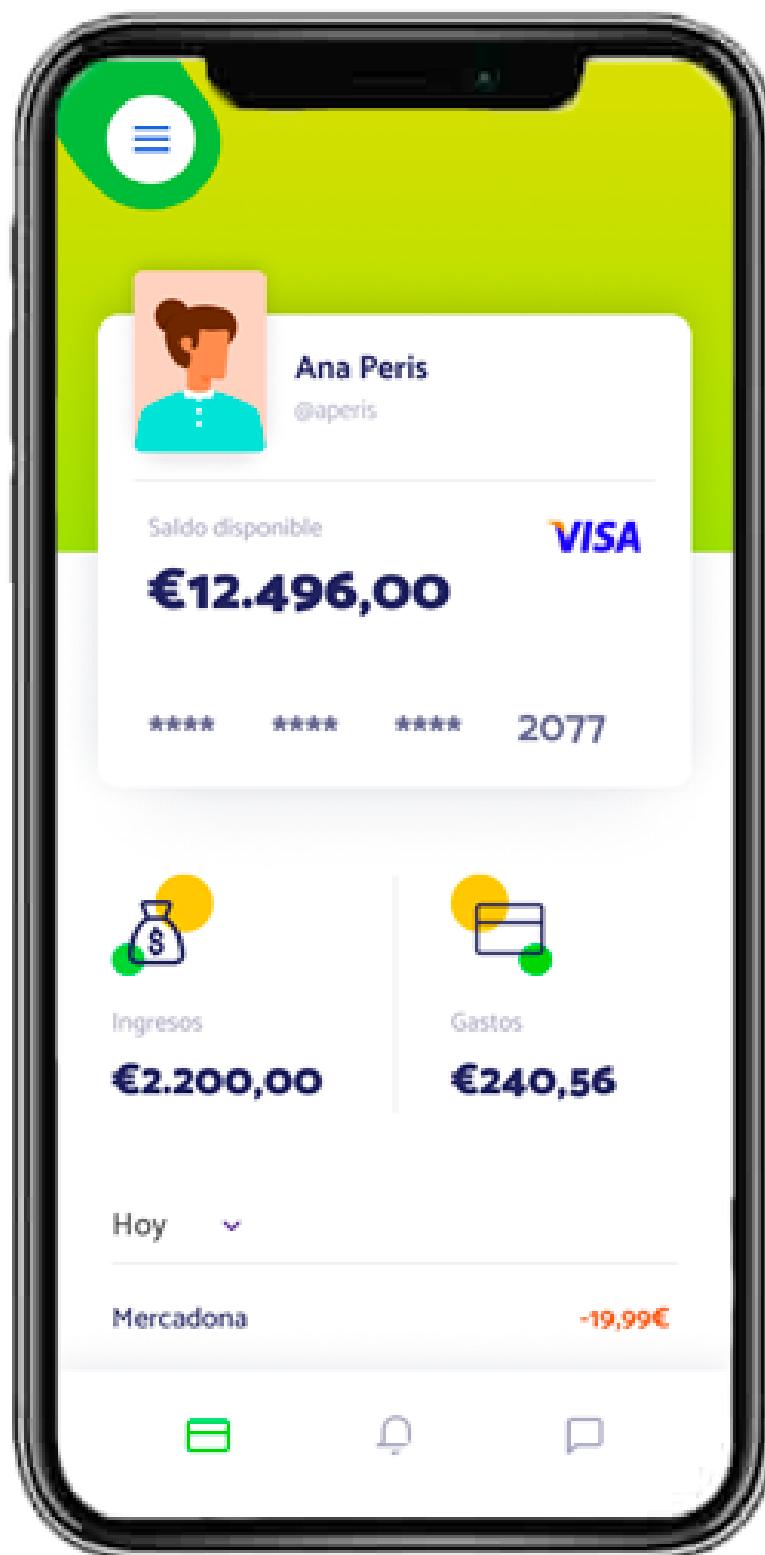
- Comprar un PC el año que viene
- Muy bueno en su trabajo
- Sus clientes le han recomendado

Frustaciones

- Diferentes cuentas bancarias
- No tiene conocimientos financieros
- No sabe la cantidad de dinero que gasta ni donde



- Envío y solicitud de pagos
- Vaults
- Criptomonedas
- Visualización de diferentes cuentas
- Visualización de gastos
- Transferencias de dinero





BIG BANK

Febrero 2020

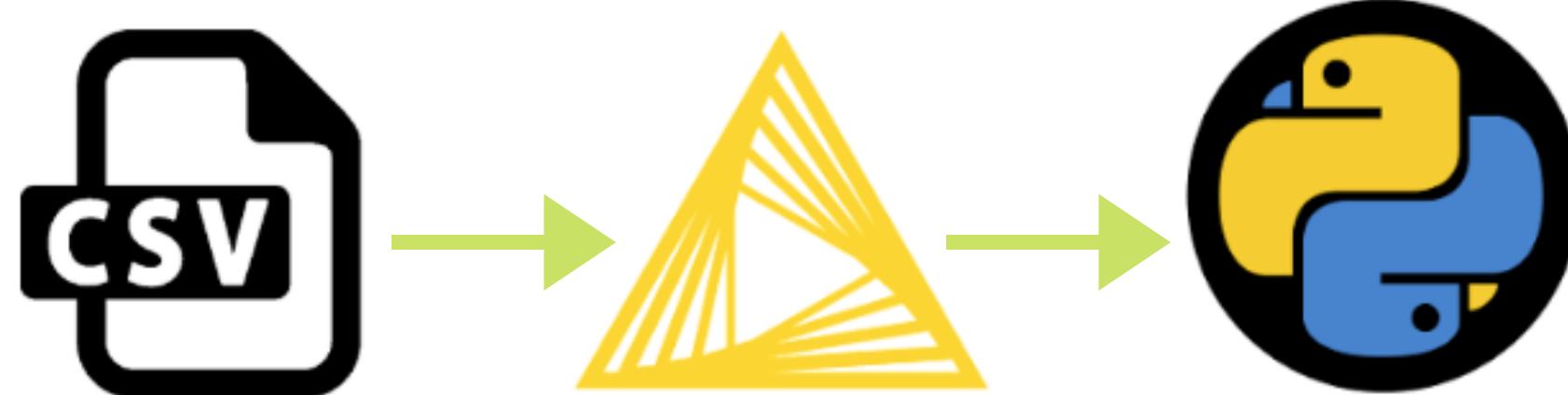
Equipo Super López



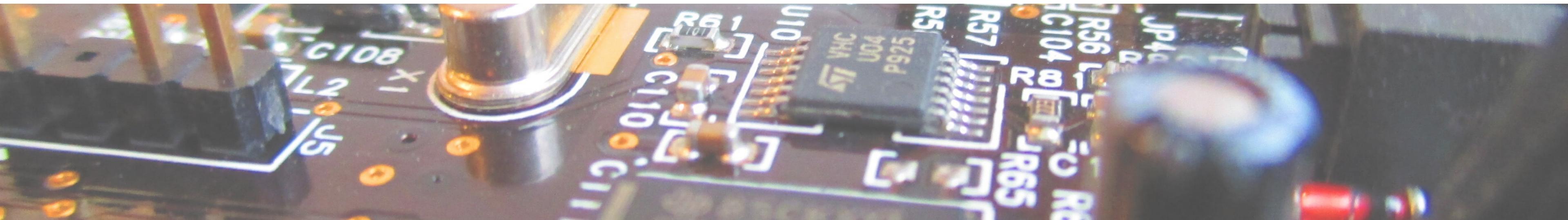
Rocío Cort
Ignacio Costa
Javier de Toro
Sergio Díez
Victor Malvar

Anexo (I)

Arquitectura utilizada



1. Prototipar de forma mínima en la aplicación Knime, que nos permite agilizar y tener una idea clara y rápida de los datasets dados.
2. Utilizar Python como lenguaje de programación donde, con las correspondientes librerías, somos capaces de implementar el código que nos permite construir los modelos de Machine Learning que dan solución a los casos de uso propuestos.



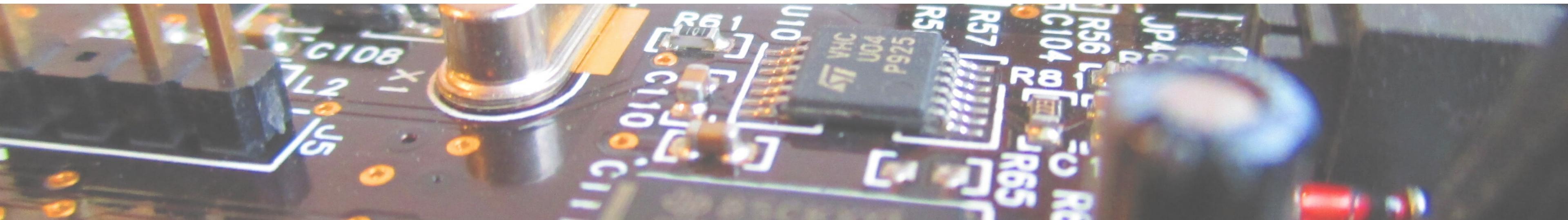
Anexo (II)

Elección algoritmo

Bibliografía:

Kirasich, Kaitlin; Smith, Trace; and Sadler, Bivin (2018) "Random Forest vs Logistic Regression: Binary Classification for Heterogeneous Datasets," SMU Data Science Review: Vol. 1 : No. 3 , Article 9. Available at: <https://scholar.smu.edu/datasciencereview/vol1/iss3/9>

Houtao, Deng; (2018) "Why Random Forests outperform decision trees. Two intuitive reasons illustrated" Towards Data Science. Available at: <https://towardsdatascience.com/why-random-forests-outperform-decision-trees-1b0f175a0b5>



Anexo (III)

Definición de variables

Row ID	ID fila	Number of Open Accounts	Número de cuentas abiertas
Loan ID	ID prestamos	Number of Credit Problems	Número de problemas de crédito
Customer ID	ID cliente	Current Credit Balance	Dinero disponible actual
Loan Status	Estado préstamo	Maximum Open Credit	Cantidad de crédito máximo abierto
Current Loan Amount	Cantidad préstamo actual	Bankruptcies	Quiebras
Term	Plazo préstamo	Tax Liens	Gravamen fiscal (Impuesto para garantizar el pago)
Credit Score	Solvencia		
Annual Income	Ingresos anuales		
Years in current job	Años en puesto trabajo actual		
Home Ownership	Propiedad de la vivienda		
Purpose	Propósito préstamo		
Monthly Debt	Deuda mensual		
Years of Credit History*	Años de historial crediticio*		
Months since last delinquent	Meses desde último delito		

* Years of Credit History - Un historial crediticio es el registro de cómo una persona ha administrado su crédito en el pasado, incluida la cantidad total de la deuda, el número de líneas de crédito y la puntualidad del pago. Los prestamistas analizan el historial crediticio de un cliente potencial para decidir si ofrecen o no una nueva línea de crédito y para ayudar a establecer los términos del préstamo.