



**Proyecto: BD+ML Apply  
Confirming**

**Optimización Inteligente del Confirming  
mediante Big Data y Machine Learning**

**EJE TECNOLÓGICO**

## Presentación del proyecto

Logros alcanzados

Conclusiones



## Presentación del proyecto - Antecedentes

---

- **alvantia** es una compañía tecnológica **especializada en soluciones para financiación** a empresas.
- Dispone de un **equipo multidisciplinar** formado por expertos tecnológicos y especialistas de negocio.
- Ha desarrollado **soluciones completas de financiación** de circulante orientadas a la financiación comercial (Factoring y Confirming).
- **Comercialmente**, entre los clientes potenciales de **alvantia** (entidades financieras no bancarias de España y Latinoamérica) se detecta la **necesidad de ofrecer una solución completa y avanzada** que proporcione una panorámica visual del estado del negocio y una fiel previsión del comportamiento de los proveedores.



## Presentación del proyecto - Objetivos técnicos específicos

---

- Aumentar el *stack* tecnológico en la rama de Big Data, BI e Inteligencia Artificial mediante el uso de lenguaje de programación y herramientas de desarrollo punteras.
- Investigar el uso de las tecnologías que puedan permitir la escalabilidad de la solución en orígenes de datos con gran volumetría.
- Investigar y desarrollar algoritmos de Machine Learning que permitan optimizar los anticipos en Confirming mediante previsión de comportamiento de proveedores.
- Investigar y desarrollar un sistema de propensiones y sensibilidades que sean capaces de ayudar a entrenar los modelos de Machine Learning y a la toma de decisiones en Confirming.
- Implantar un sistema de predicción y visualización de datos que presente una visión global y dinámica de los proveedores y el negocio de Confirming.

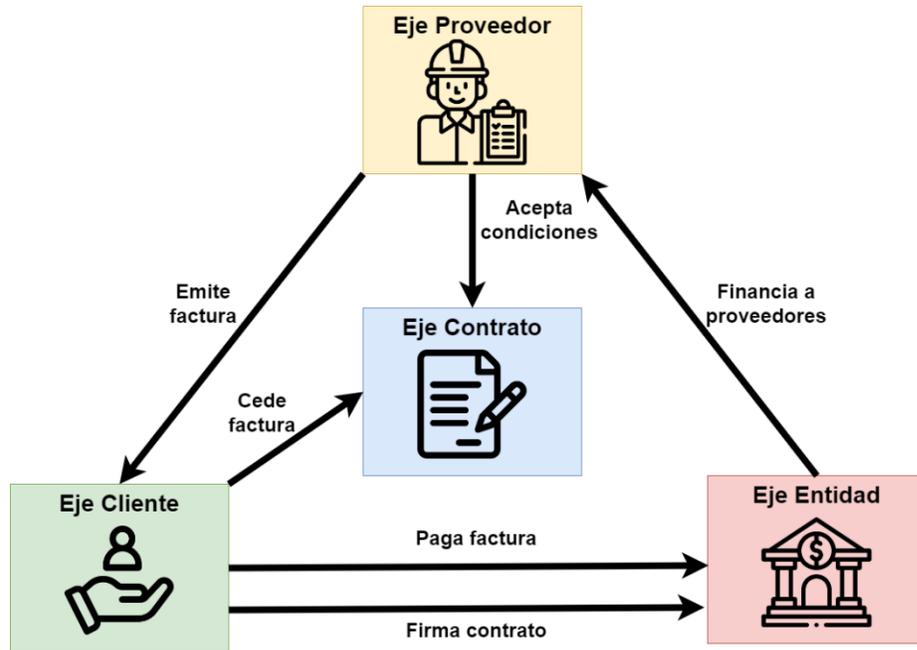
Presentación del proyecto

**Logros alcanzados**

Conclusiones



# Logros alcanzados - Identificación de las principales variables de Confirming



- Se han identificado variables agrupados en **4 ejes distintos**: Cliente, Contrato, Proveedor y Proveedor-Contrato.
- Los ejes reflejan el estado de todos los **actores principales** del negocio de confirming, permitiendo predecir con mayor precisión si un proveedor financia o no y diseñar adecuadamente el DWH.



# Logros alcanzados - Investigación y elección de herramientas

Se ha investigado las distintas **tecnologías disponibles para el desarrollo del proyecto**, seleccionando las siguientes:



- Plataforma **opensource** (durante el desarrollo)
- Empleada para el desarrollo de las **ETLs**



- PostgreSQL es un sistema de **gestión de BD** relacional **opensource**
- Empleada para **alojar el DWH**



- Plataforma unificada y escalable de **BI**
- Empleada para el desarrollo de los **cuadros de mando**



- **Entrenamiento de modelos** mediante Python.
- La **gestión de entornos** se realiza con Anaconda y con la IDE Visual Studio



- Librería Python para el **desarrollo de APIs**.
- Empleada para la **implementación del modelo** en un microservicio



- Plataforma **opensource** para monitorear **flujos de trabajo**.
- Empleado para **monitorizar la ETL**.



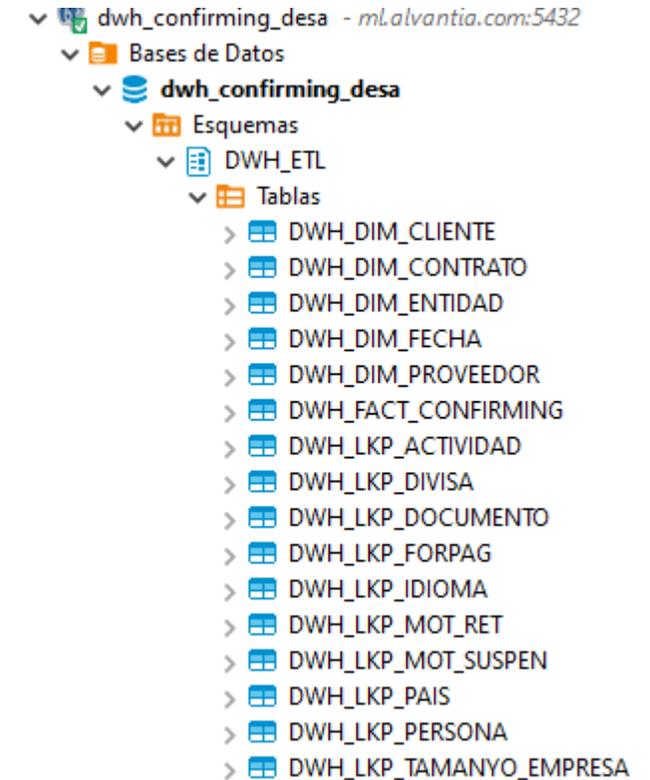
# Logros alcanzados - Desarrollo de procesos ETL para el BI

- Se han desarrollado empleando Talend varios procesos ETL. Estos procesos se dividen en **Lookup, Dimensión y Hechos**.
- Orquestando dichos procesos mediante **AirFlow**, se puede tanto **inicializar como actualizar** el DWH para el Bussines Intelligence (BI) en poco tiempo.
- Como resultado de todos estos procesos se **generan diversas tablas**, también agrupables en Lookup, Dimensión y Hechos.

## Talend

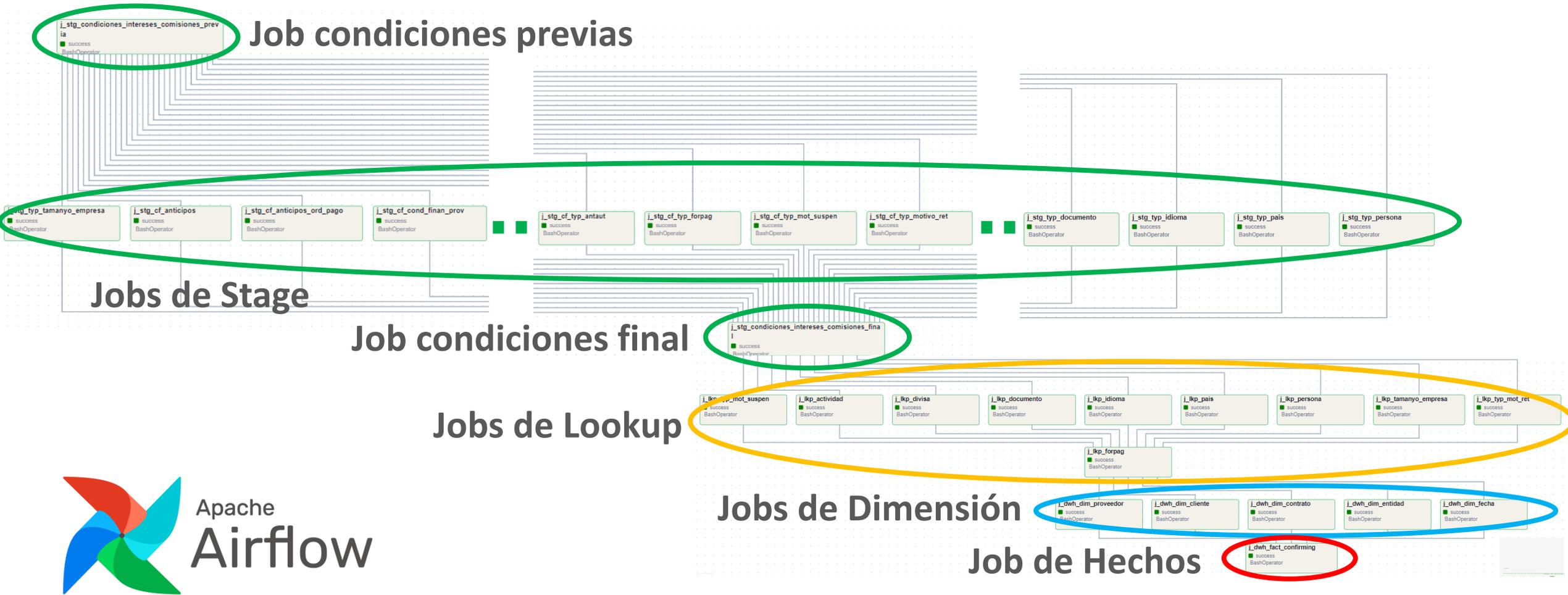


## PostgreSQL





# Logros alcanzados - Desarrollo de procesos ETL para el BI





## Logros alcanzados - Desarrollo de procesos ETL para ML

---



Job de **inicialización de la base de datos** de destino del proceso ETL.

Este job se utiliza para hacer una carga inicial de las experiencias de las facturas existentes en la base de datos, lo cual implica que este job se demorará bastante, por lo que no es lo ideal para actualizar la base de datos cada cierto tiempo.

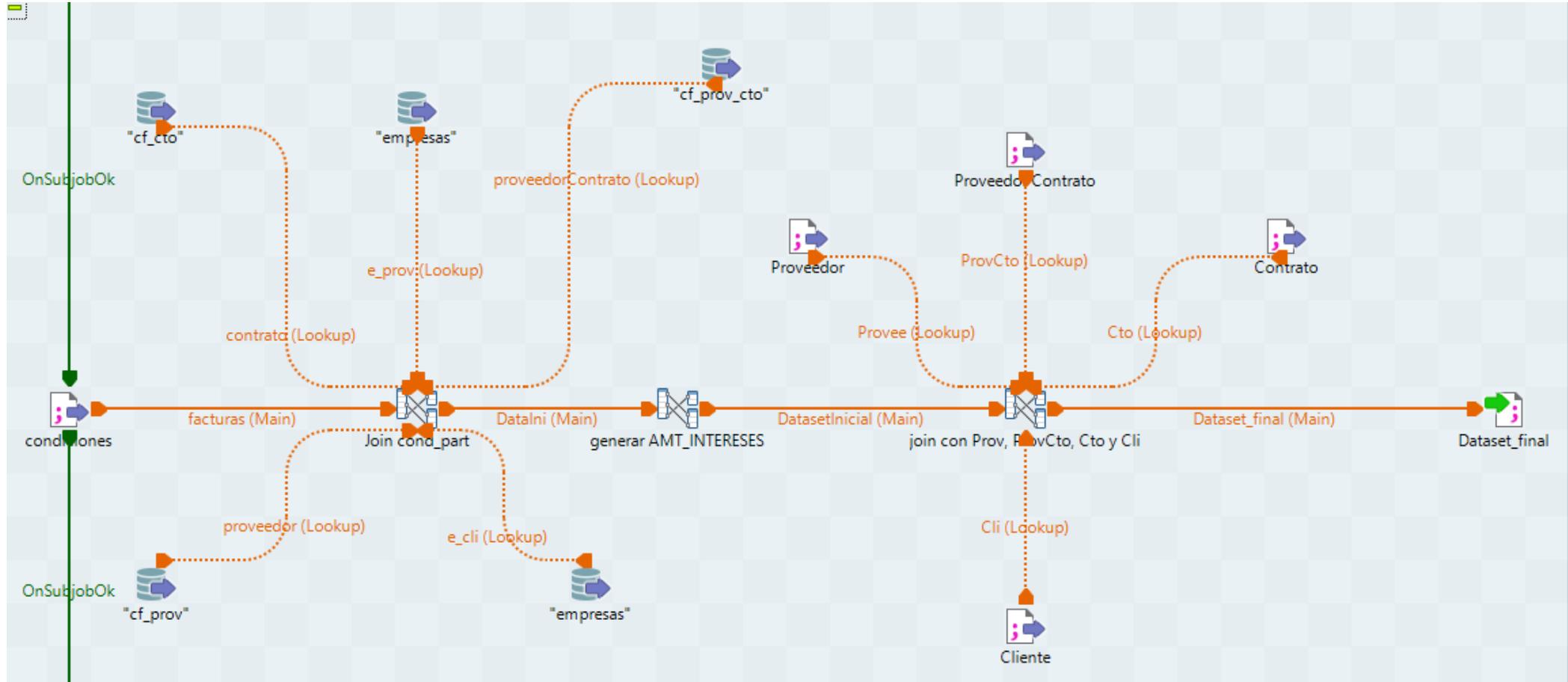


Job de **actualización de la base de datos** de destino del proceso ETL.

Este job se utiliza para cargar las nuevas facturas como experiencias en la base de datos, lo cual implica que se tiene que ejecutar periódicamente, lo cual implica que debe tardar menos que el job de inicialización.



# Logros alcanzados - Desarrollo de procesos ETL para ML

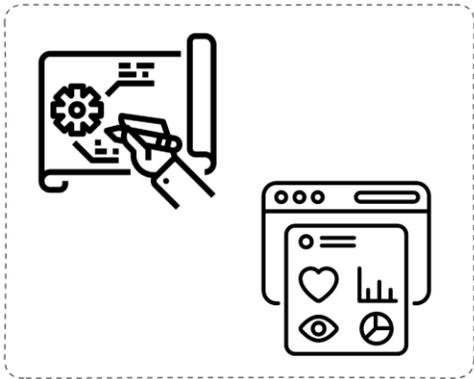




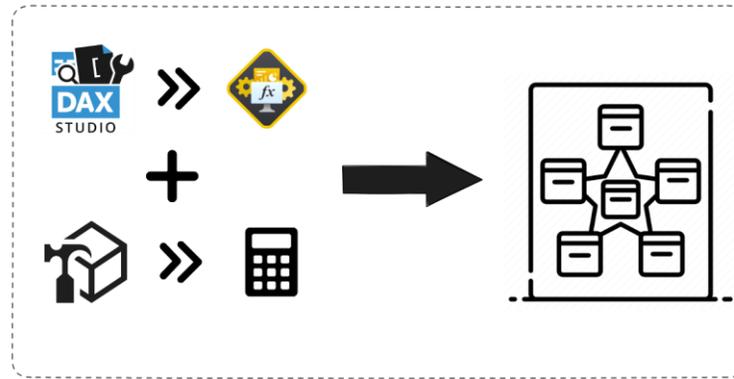
# Logros alcanzados - Desarrollo de Dashboards analíticos

A continuación, se detallan los pasos para el desarrollo de los cuadros de mando o Dashboards.

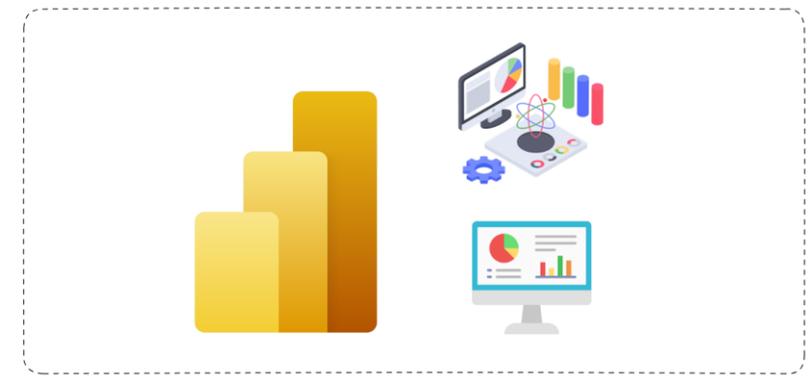
## 1. Diseño del prototipo



## 2. Diseño del modelo de datos de Power BI



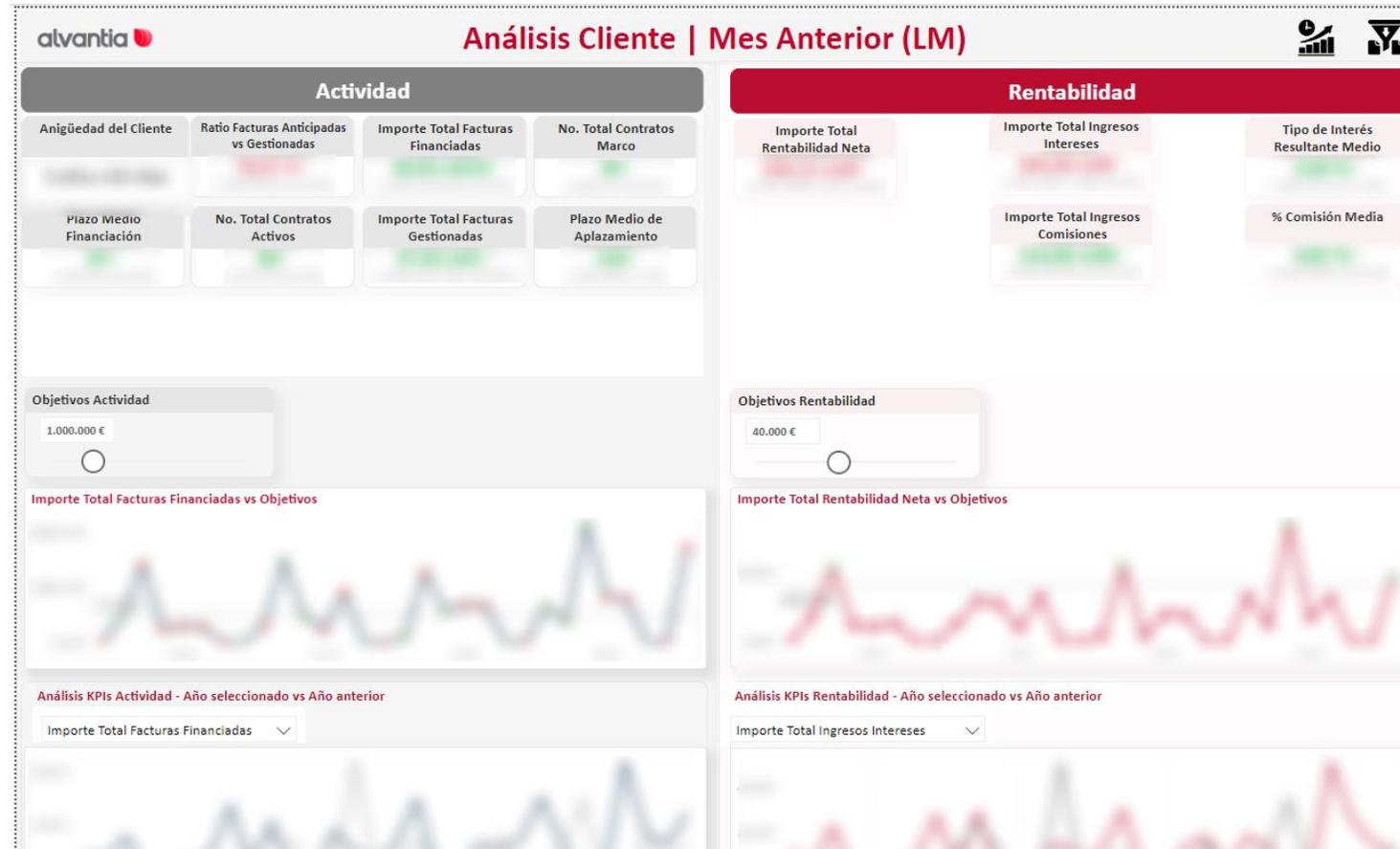
## 3. Diseño de los cuadros de mando con Power BI





# Logros alcanzados - Desarrollo de Dashboards analíticos

Como resultado de este proceso se han obtenido 3 Dashboards, que reflejan el estado de los principales actores del negocio de confirming. A continuación, se muestra el Dashboard de clientes:

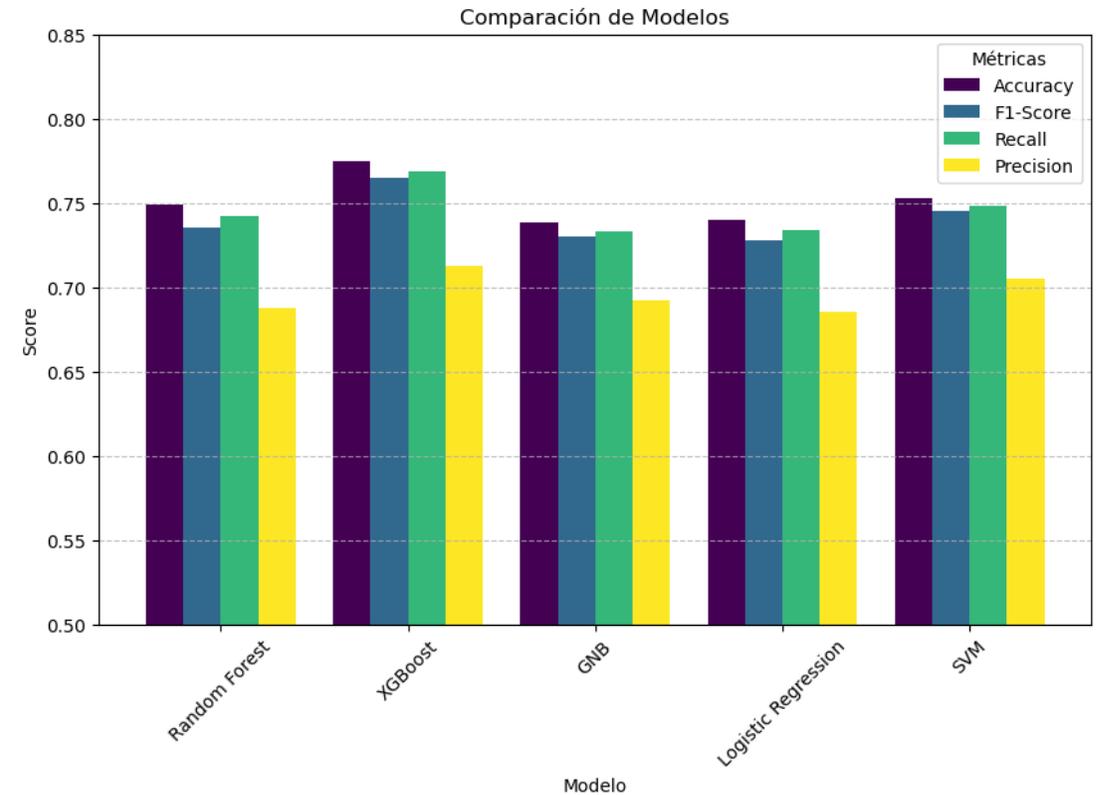




# Logros alcanzados – Entrenamiento del Modelo

Se ha logrado entrenar un modelo capaz de predecir anticipos por facturas, obteniendo los siguientes resultados:

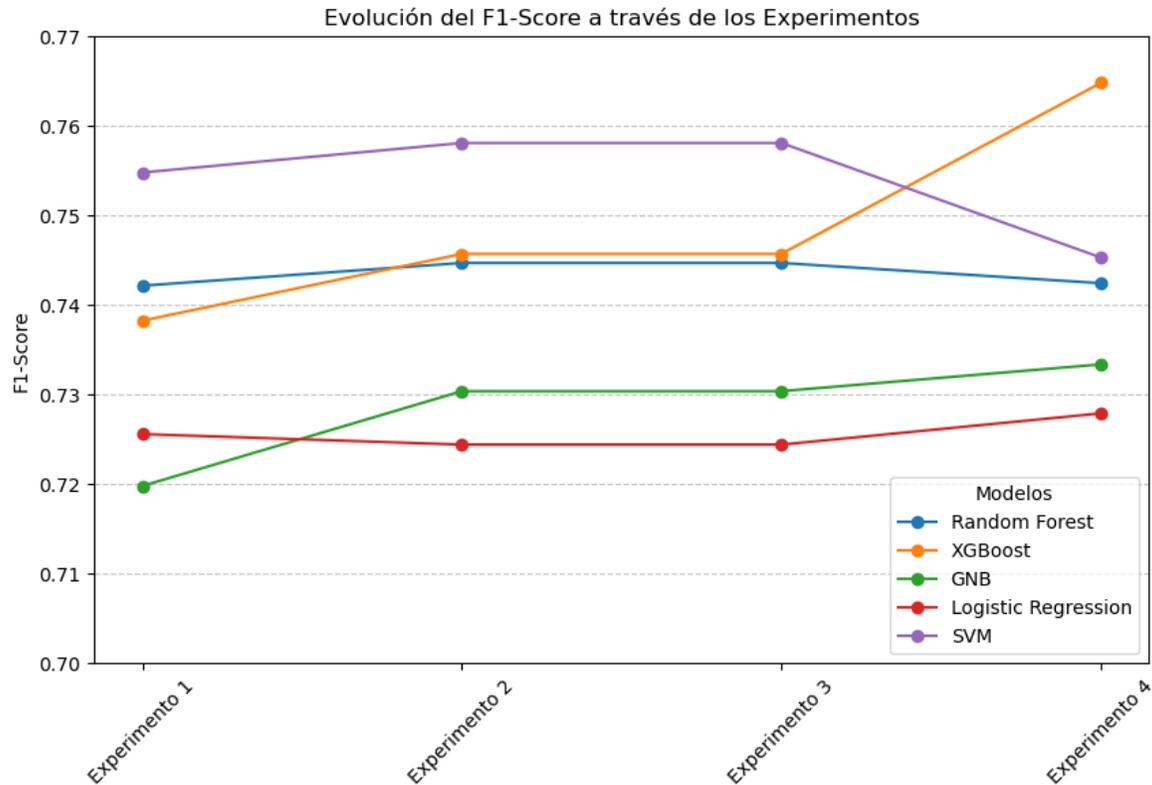
	Accuracy	F1-Score	Recall	Precision
Random Forest	0.74933	0.73572	0.74245	0.68761
<b>XGBoost</b>	<b>0.77458</b>	<b>0.76482</b>	<b>0.76852</b>	<b>0.71235</b>
GNB	0.73877	0.72972	0.73338	0.69208
Logistic Regression	0.73976	0.72791	0.73350	0.68565
SVM	0.75307	0.74531	0.74799	0.70542





# Logros alcanzados – Entrenamiento del Modelo

Se han desarrollado diversos experimentos probando distintas combinaciones de criterios y de variables, obteniendo los siguientes resultados:



F1-Score	Random Forest	XGBoost	GNB	Logistic Regression	SVM
Experimento 1	0.74217	0.73826	0.71980	0.72560	0.75481
Experimento 2	0.74471	0.74572	0.73039	0.72442	0.75811
Experimento 3	0.74471	0.74572	0.73039	0.72442	0.75811
Experimento 4	0.74245	0.76482	0.73338	0.72791	0.74531

Presentación del proyecto

Logros alcanzados

**Conclusiones**



# Conclusiones – Transformación digital del Confirming

La integración de Big Data y Machine Learning ha permitido evolucionar los procesos tradicionales de Confirming hacia modelos más eficientes, predictivos y automatizados.

Antes, las condiciones de anticipo eran estándar, pero con el desarrollo del modelo de ML se consigue adaptar las condiciones de anticipo a las necesidades específicas de cada proveedor.



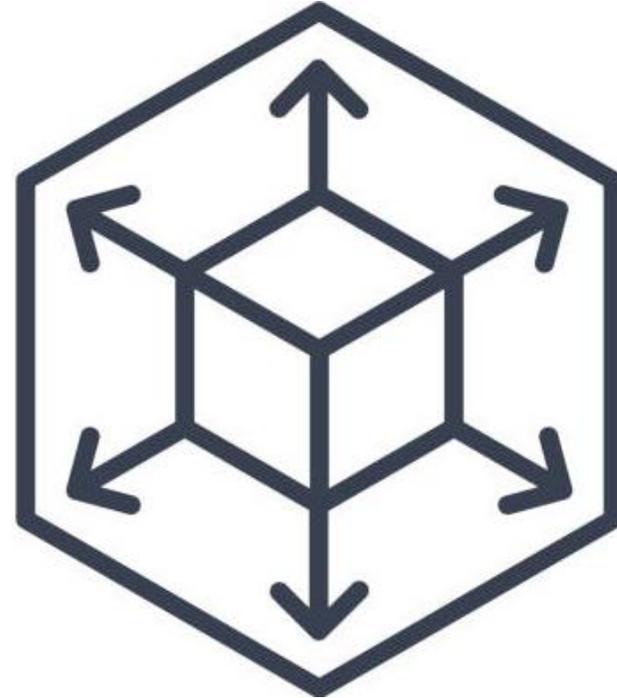


## Conclusiones – Arquitectura escalable y modular

---

El uso de tecnologías open-source y la implementación de microservicios permiten el diseño de una arquitectura flexible preparada para el crecimiento y la integración con diversas fuentes de datos.

Esto permite que la solución pueda ser fácilmente implementada en todo tipo de clientes, requiriendo tan solo unos pequeños ajustes.





## Conclusiones – Automatización de procesos ETL

La creación de procesos ETL robustos y optimizados en la carga, transformación y actualización de los datos ha permitido garantizar unos datos de calidad, consistentes y que estén actualizados siempre.





## Conclusiones – Visualización analítica avanzada

---

Los cuadros de mando desarrollados con Power BI permiten obtener una visión clara y dinámica del estado de clientes, contratos y proveedores, mejorando la toma de decisiones y la capacidad de análisis y seguimiento.





# Conclusiones – Optimización de la toma de decisiones

Los modelos predictivos y los cuadros de mando implementados permiten mejorar sensiblemente la toma de decisiones, pues aportan más información sobre el estado del proveedor y las condiciones más rentables que está dispuesto a aceptar.

Esto permite reducir el riesgo y maximizar los anticipos.





## Conclusiones – Líneas futuras de desarrollo

A futuro, Alvantia continuará evolucionando la solución, integrando:

- Modelos de refuerzo de aprendizaje para optimizar aún más las decisiones en tiempo real.
- Explicabilidad de los modelos para mejorar la transparencia en la toma de decisiones, utilizando herramientas como Shap Values, Lime y feature importance de Ensemble Models.
- Aplicación en mercados internacionales, expandiendo la solución a nuevas geografías.



