

Adaptaciones CRISP-DM

Adaptaciones de CRISP - DM por Fase

Estas adaptaciones no alteran los objetivos de CRISP-DM, sino que optimizan su aplicación para que el proceso sea más práctico, limpio y eficiente.

- **Business Understanding**

Esta fase identifica los objetivos del análisis de datos a través de su entendimiento, desde una perspectiva de negocio, contestando qué es lo que el cliente quiere resolver realmente. Es una fase crítica del proyecto porque evita que se produzcan las respuestas correctas a las preguntas equivocadas.

Business Understanding:

- **Determine Business Objectives**
- **Asses Situation**
- **Determine Data Mining Goals**
- **Produce Project Plan**

Cambio 1: Unificación de “Determine Business Objectives”, “Asses Situation”, “Determine Data Mining Goals”

Antes:

- Determine Business Objectives
- Asses Situation
- Determine Data Mining Goals
- Produce Project Plan

Después:

- **Nuevo entregable:** Documento de Alineación del Proyecto (DAP)

- **Data Understanding y Preparation**

En la práctica, las etapas de Understanding y Preparation ocurren en un ciclo continuo.

El objetivo de data preparation es producir un dataset que será utilizado para modelar o para un mayor análisis del trabajo del proyecto, lo cual va muy relacionado con el entendimiento de los datos.

Data Understanding se presenta originalmente con entregables separados:

- **Data Collection Report**
- **Data Description Report**

- **Data Exploration Report**
- **Data Quality Report**

Data Preparation se presenta originalmente con entregables separados:

- **Dataset**
- **Select Data**
- **Clean Data**
- **Construct Data**
- **Integrate Data**
- **Format Data**

Cambio 1: Unificación de “Data Understanding”

Antes:

- Data Collection Report
- Data Description Report
- Data Exploration Report
- Data Quality Report

Después:

- **Nuevo entregable: Reporte de Comprensión de Datos (RCD)**

Cambio 2: Unificación de “Data Preparation”

Antes:

- Dataset
- Select Data
- Clean Data
- Construct Data
- Integrate Data
- Format Data

Después:

- **Nuevo entregable: Reporte de Preparación de Datos (RPD)**

Justificación:

La preparación de datos es un pipeline cohesivo, donde todas las transformaciones dependen unas de otras, y documentarlo en piezas aisladas va en contra de las prácticas modernas de MLOps, reproducibilidad y mantenimiento.

Data Preparation incluye:

- Limpieza
- Transformaciones
- Feature engineering
- Selección de atributos
- Creación del dataset final

En muchos proyectos estas actividades se realizan juntas dentro de un script o pipeline.

- **Modeling**

Modeling se presenta originalmente con entregables separados:

- **Select Modeling Technique**
- **Generate Test Design**
- **Build Model**
- **Assess Model**

Cambio 1: Unificación de “Modeling”

Antes:

- Modeling + Fases

Después:

- **Nuevo entregable: Reporte de Modelado (RM)**

- **Evaluation**

Evaluation se presenta originalmente con entregables separados:

- **Evaluate Results**
- **Review Process**
- **Determine Next Steps**

Cambio 1: Unificación de “Evaluation”

Antes

- Evaluation + Fases

Después:

- **Nuevo entregable: Reporte de Evaluación (RE)**

- **Deployment**

La fase de Deployment, según el enfoque tradicional de CRISP-DM, contempla una implementación completa del modelo en un entorno de producción posterior a una evaluación exhaustiva. Sin embargo, este enfoque resulta poco flexible frente a los ciclos iterativos y de entrega continua que caracterizan a metodologías ágiles como SCRUM. En un contexto real, las necesidades del negocio, los requisitos funcionales y la madurez del modelo evolucionan constantemente; por ello, la implementación debería adaptarse a un enfoque incremental que permita validar el valor del modelo

antes de realizar despliegues definitivos o inversiones técnicas extensas.

La adopción de un MVP (Producto Mínimo Viable) en la fase de Deployment permite entregar una primera versión funcional del modelo en un entorno controlado, donde los usuarios o stakeholders pueden probar, retroalimentar y validar su utilidad real desde etapas tempranas. Este enfoque minimiza riesgos, reduce costos asociados a implementaciones completas prematuras y habilita ajustes rápidos basados en evidencia. Además, alinearse a SCRUM permite incorporar el despliegue dentro de sprints iterativos, donde cada incremento puede incluir mejoras al modelo, ampliación de funcionalidades, integración de nuevas métricas o incorporación de pipelines de monitoreo más robustos.

En este sentido, adaptar Deployment a un esquema SCRUM con un MVP no solo agiliza el proceso, sino que garantiza que la solución desplegada esté orientada a resultados y validada continuamente por usuarios reales. Esto fortalece la alineación entre ciencia de datos y desarrollo de producto, y permite transitar de un modelo experimental a una solución operativa con mayor control, retroalimentación y eficiencia.

De igual Deployment sabemos que se presenta originalmente con entregables separados:

- **Plan Deployment**
- **Plan Monitoring and Maintenance**
- **Produce Final Report**
- **Review Project**

Cambio 1: Unificación de “Deployment”

Antes

- Deployment + Fases

Después:

- **Nuevo entregable: Reporte de Deployment (RD)**