CRISP-DM | Etapa 6: Deployment

El proyecto no termina cuando llegamos a producción. Allí comienza la prueba de realidad.

Durante las primeras etapas de CRISP-DM hemos recorrido un proceso de investigación.

Partimos de una necesidad del negocio. Entendimos los datos disponibles, los preparamos, construimos alternativas de solución y evaluamos si los resultados obtenidos respondían a la necesidad que dio origen al proyecto.

Durante ese recorrido dejamos algo fundamental: **evidencia**.

Documentamos qué investigamos, qué datos utilizamos, qué transformaciones realizamos, qué supuestos adoptamos, qué resultados obtuvimos y bajo qué condiciones esperamos que esos resultados se produzcan.

Ahora llega Deployment.

Y aquí cambia la pregunta.

Ya no preguntamos solamente:

¿El resultado obtenido es técnicamente correcto?

Ahora debemos preguntarnos:

¿Funcionará en la realidad y cómo vamos a demostrarlo?

---

1. Deployment es llevar nuestro aprendizaje a la operación

Deployment suele asociarse inmediatamente con poner un modelo de Machine Learning en producción.

Pero puede ser mucho más amplio.

Podemos desplegar un modelo, una segmentación de clientes, una estrategia comercial, una recomendación, una regla de negocio, un dashboard, un proceso analítico o simplemente utilizar los conocimientos descubiertos durante nuestra investigación para cambiar la manera en que la organización toma decisiones.

Por eso, Deployment no consiste solamente en conectar técnicamente una solución.

Consiste en **convertir conocimiento en acción y acción en resultados medibles**.

El verdadero ciclo debería ser:

Investigamos → encontramos evidencia → implementamos → medimos → aprendemos → ajustamos → volvemos a medir.

---

2. No debemos llegar a producción preguntándonos cómo vamos a medir

La estrategia de medición debería estar diseñada **antes del lanzamiento**.

Si esperamos hasta después para preguntarnos cómo demostrar si algo funcionó, posiblemente descubriremos que nunca capturamos los datos que necesitábamos.

Antes del Deployment deberíamos poder responder al menos:

- ¿Qué esperamos que ocurra?
- ¿Con qué indicador lo vamos a medir?
- ¿Cuál era nuestra situación antes de implementar?
- ¿A qué población estamos impactando?
- ¿Existe una población comparable que no reciba la intervención?
- ¿Cuándo comienza la medición?
- ¿Cuánto tiempo esperamos que tarde en aparecer el efecto?
- ¿Durante cuánto tiempo seguiremos observándolo?
- ¿Qué consideraremos éxito?
- ¿Qué consideraremos una señal de alerta?
- ¿Cuándo debemos ajustar?
- ¿Cuándo debemos detener?
- ¿Quién será responsable de tomar esas decisiones?

Si estas preguntas no tienen respuesta, realmente no tenemos todavía un plan completo de Deployment.

Tenemos solamente un lanzamiento.

---

3. El tiempo también forma parte del modelo de medición

Consideremos una campaña bancaria relacionada con **Bono 14 o aguinaldo**.

Supongamos que identificamos determinados segmentos de clientes y lanzamos una promoción.

Sería un error establecer una relación automática:

Campaña → compra.

El comportamiento humano raramente funciona de forma tan inmediata.

Un cliente podría:

- comprar inmediatamente;
- anticipar una compra que pensaba realizar posteriormente;
- esperar hasta recibir el ingreso;
- comprar días o semanas después;
- comprar independientemente de nuestra promoción;
- recibir la promoción y no cambiar su comportamiento.

Por eso necesitamos definir una **ventana de observación**.

La pregunta deja de ser solamente:

¿Cuántos clientes compraron?

y evoluciona hacia:

¿Cuándo comenzó a observarse un cambio, cuánto duró y qué evidencia tenemos de que nuestra intervención contribuyó a producirlo?

Incluso puede existir algo muy valioso para el negocio:

un punto de saturación.

Puede llegar un momento en que continuar pagando publicidad, comunicaciones, incentivos o infraestructura deje de producir valor incremental.

Detectarlo también es un resultado del Deployment.

---

4. Una conversión no demuestra necesariamente éxito

Supongamos:

> Enviamos una promoción a un cliente y posteriormente el cliente compró.

Eso demuestra que hubo una compra.

Pero todavía debemos investigar si la compra ocurrió **por efecto de nuestra intervención**.

Quizás ese cliente ya pensaba comprar.

Quizás habría comprado exactamente igual sin recibir la promoción.

Quizás simplemente adelantamos una compra que habría ocurrido posteriormente.

Por eso una organización Data Driven intenta evolucionar de preguntas descriptivas hacia preguntas causales:

¿Qué pasó?

Luego:

¿Nuestra intervención produjo algún cambio?

Y finalmente:

¿Por qué ocurrió y qué deberíamos hacer a continuación?

La madurez analítica no está únicamente en producir mejores algoritmos. También está en formular mejores preguntas.

---

5. La línea base: necesitamos saber desde dónde partimos

Para medir un cambio debemos conocer nuestro punto de referencia.

Antes del despliegue deberíamos definir una **baseline** que represente la situación previa.

Dependiendo del proyecto podría incluir:

- ventas históricas;
- comportamiento habitual del segmento;
- tasa de conversión;
- frecuencia de compra;
- ticket promedio;
- uso del producto;
- abandono;
- mora;
- rentabilidad;
- costo operativo;
- tiempo de respuesta;
- satisfacción;
- comportamiento estacional.

La métrica dependerá del objetivo definido durante Business Understanding.

Y aquí podemos observar algo importante:

Deployment sigue conectado con la primera etapa de CRISP-DM.

Si inicialmente dijimos qué problema queríamos resolver y cómo reconoceríamos el éxito, ahora tendremos que demostrar si realmente ocurrió.

---

6. Diseñemos la medición desde el principio

Un buen Deployment debería definir, como mínimo:

Objetivo de negocio

¿Qué resultado queremos cambiar?

Hipótesis

¿Qué esperamos que suceda y por qué?

Población objetivo

¿Sobre quién aplicaremos la acción?

Baseline

¿Cuál era el comportamiento previo?

Grupo de comparación o control

Cuando sea metodológicamente y operativamente posible, ¿contra qué compararemos los resultados?

Indicadores de éxito

¿Qué KPI demostrará que estamos avanzando?

Ventana temporal

¿Cuándo esperamos comenzar a observar el efecto y cuánto tiempo debemos seguir midiéndolo?

Criterios de decisión

¿Qué resultado significa continuar, ajustar, escalar o detener?

Captura de datos

¿Qué información debe registrar nuestros sistemas para posteriormente reconstruir lo sucedido?

Responsables

¿Quién monitorea y quién decide?

Evidencia

¿Dónde documentaremos resultados, decisiones, excepciones y aprendizajes?

---

7. No midamos solamente al cliente: midamos la intervención

Para analizar correctamente una campaña no basta con tener información del cliente y de la compra.

También necesitamos capturar adecuadamente la intervención.

Por ejemplo:

- qué cliente fue seleccionado;
- a qué segmento pertenecía;
- qué campaña recibió;
- qué oferta recibió;
- cuándo recibió el impacto;
- por qué canal;
- cuál era la vigencia;
- cuál fue el costo asociado;
- qué comportamiento ocurrió posteriormente;
- cuándo ocurrió;
- qué resultado económico produjo.

Esto permite construir algo fundamental:

trazabilidad.

Debemos poder recorrer el camino desde una decisión analítica hasta el resultado observado.

Sin trazabilidad podemos tener muchas transacciones y muy poca evidencia.

---

8. Antes de hablar de Machine Learning, debemos hablar de medición

Podemos utilizar algoritmos extremadamente sofisticados.

Podemos tener Big Data.

Podemos construir un Cliente 360.

Podemos usar Inteligencia Artificial.

Podemos disponer de dashboards extraordinarios.

Pero ninguna tecnología reemplaza una pregunta básica:

¿Cómo sabemos que funcionó?

Una solución sofisticada sin estrategia de medición puede producir menos conocimiento que un proceso sencillo correctamente diseñado y medido.

No es el tamaño de la tecnología lo que convierte una organización en Data Driven.

Es su capacidad de utilizar evidencia para tomar decisiones.

---

9. ¿Cómo puede ayudarnos la IA durante Deployment?

La IA puede convertirse en un excelente copiloto para esta etapa.

Puede ayudarnos a:

Diseñar

- identificar métricas potenciales;
- proponer hipótesis;
- revisar criterios de éxito;
- identificar variables que quizá no estamos considerando;
- plantear escenarios de medición.

Monitorear

- analizar grandes volúmenes de resultados;
- identificar anomalías;
- comparar comportamientos;
- señalar desviaciones respecto de lo esperado;
- resumir evidencia periódicamente.

Interpretar

- buscar patrones;
- formular nuevas hipótesis;
- identificar segmentos cuyo comportamiento difiere del esperado;
- explicar resultados para diferentes audiencias.

Documentar

- mantener bitácoras;
- generar reportes;
- comparar resultados esperados contra observados;
- registrar aprendizajes;
- preparar conocimiento reutilizable para investigaciones futuras.

La IA puede permitirnos observar mucho más rápidamente.

Pero observar no es lo mismo que decidir.

---

10. Human-in-the-Loop: la decisión continúa siendo humana

Durante Deployment, el Human-in-the-Loop se vuelve todavía más importante.

Un sistema puede detectar:

> La tasa de conversión disminuyó.

Pero un humano debe preguntarse:

¿Por qué?

Puede ser consecuencia del producto.

Puede existir un problema en el canal.

Puede haber estacionalidad.

Puede haber cambiado el contexto económico.

Puede existir un problema de datos.

Puede simplemente haber pasado demasiado poco tiempo.

El profesional debe evaluar contexto, negocio, riesgo, regulación, impacto sobre los clientes y consecuencias de la decisión.

Por eso un esquema sano es:

IA detecta → IA propone → humano interpreta → humano valida → negocio decide → sistema ejecuta → todos vuelven a medir.

Human-in-the-Loop no es un obstáculo para la automatización.

Es parte de una automatización responsable.

---

11. Las decisiones también forman parte del Deployment

Un Deployment maduro debería permitir periódicamente tomar alguna de estas decisiones:

Continuar

Los resultados evolucionan dentro de lo esperado.

Escalar

Existe evidencia suficiente para ampliar la intervención.

Ajustar

Existe valor, pero algún componente necesita modificarse.

Pausar

Aparecieron señales que requieren investigación.

Detener

La intervención dejó de generar el valor esperado o su costo ya no se justifica.

Reinvestigar

La realidad contradijo nuestras hipótesis y necesitamos regresar a etapas anteriores.

Y regresar **no significa haber fracasado**.

Significa que CRISP-DM está funcionando como metodología iterativa.

---

12. ¿Qué caracteriza a una organización Data Driven?

Una organización Data Driven no afirma:

> “Todas nuestras campañas fueron exitosas.”

Pregunta:

> “¿Cuál era nuestro criterio de éxito?”

Luego pregunta:

> “¿Con qué evidencia podemos demostrarlo?”

Y posteriormente:

> “¿Qué aprendimos y qué decisión tomaremos?”

Una organización madura convierte cada Deployment en una nueva fuente de conocimiento.

El resultado no queda únicamente en un dashboard.

Se documenta:

Objetivo → hipótesis → intervención → medición → resultado → decisión → aprendizaje.

Ese conocimiento puede alimentar próximos análisis, nuevos modelos, futuras campañas y mejores decisiones.

Allí aparece un verdadero círculo virtuoso:

Dato → conocimiento → decisión → acción → nuevo dato → nuevo conocimiento.

---

13. El Deployment debe producir conocimiento institucional

Al finalizar un ciclo de medición deberíamos conservar:

- resultados obtenidos;
- diferencias frente a lo esperado;
- hipótesis confirmadas o rechazadas;
- segmentos que respondieron y segmentos que no;
- comportamiento temporal del impacto;
- anomalías observadas;
- decisiones tomadas;
- cambios realizados;
- lecciones aprendidas;
- recomendaciones para futuros proyectos.

Esta documentación convierte un proyecto individual en **conocimiento organizacional reutilizable**.

El siguiente equipo no debería comenzar desde cero.

Debería comenzar desde lo que la organización ya aprendió.

---

14. El verdadero producto del Deployment

El producto final de Deployment no debería ser solamente:

“El modelo está en producción.”

Tampoco:

“La campaña fue lanzada.”

Ni siquiera:

“El dashboard ya está publicado.”

El verdadero resultado debería ser un sistema capaz de responder:

> **Qué implementamos, por qué lo hicimos, qué esperábamos que sucediera, qué sucedió realmente, cómo lo medimos, qué aprendimos y qué decisión tomamos con esa evidencia.**

Eso convierte una implementación tecnológica en una capacidad de negocio.

---

15. Deployment no es el final

La última etapa de CRISP-DM contiene una paradoja:

Es la última etapa del diagrama, pero puede convertirse en la primera etapa del siguiente aprendizaje.

Antes de producción trabajábamos principalmente con nuestras hipótesis, datos históricos, modelos y supuestos.

Después de producción aparece algo que ningún laboratorio puede reproducir completamente:

la realidad.

Clientes reales.

Decisiones reales.

Cambios de comportamiento.

Estacionalidad.

Competencia.

Errores.

Eventos inesperados.

Resultados mejores de los previstos.

Resultados peores.

Y nuevos datos.

Por eso Deployment no debería cerrar el ciclo.

Debería alimentarlo.

Business Understanding → Data Understanding → Data Preparation → Modeling → Evaluation → Deployment → Medición → Aprendizaje → nueva necesidad del negocio.

Y CRISP-DM comienza nuevamente.

Reflexión final

Quizás la pregunta más peligrosa después de poner algo en producción sea:

> **“¿Funcionó?”**

si nadie definió previamente qué significaba **funcionar**.

Y quizás una pregunta todavía más importante sea:

> **“¿Qué evidencia necesitaríamos para aceptar que estábamos equivocados?”**

Porque una organización verdaderamente Data Driven no utiliza los datos solamente para demostrar que tenía razón.

Los utiliza también para descubrir cuándo estaba equivocada, aprender rápidamente y corregir el rumbo.

Deployment no es llevar un modelo a producción.

Es llevar una hipótesis a la realidad y construir los mecanismos necesarios para medir lo que ocurre después.

Por eso:

Deployment no es el final del proyecto.

Deployment es el momento en que comienza el aprendizaje real.