

CRISP-DM

La Metodología CRISP-DM, aunque se desarrolló para llevar a delante grandes proyectos, es suficientemente amplia y flexible para aplicarla a proyectos de cualquier tamaño. En la figura 1 se detalla, nuevamente, el ciclo de vida de un proyecto de Explotación de Datos.

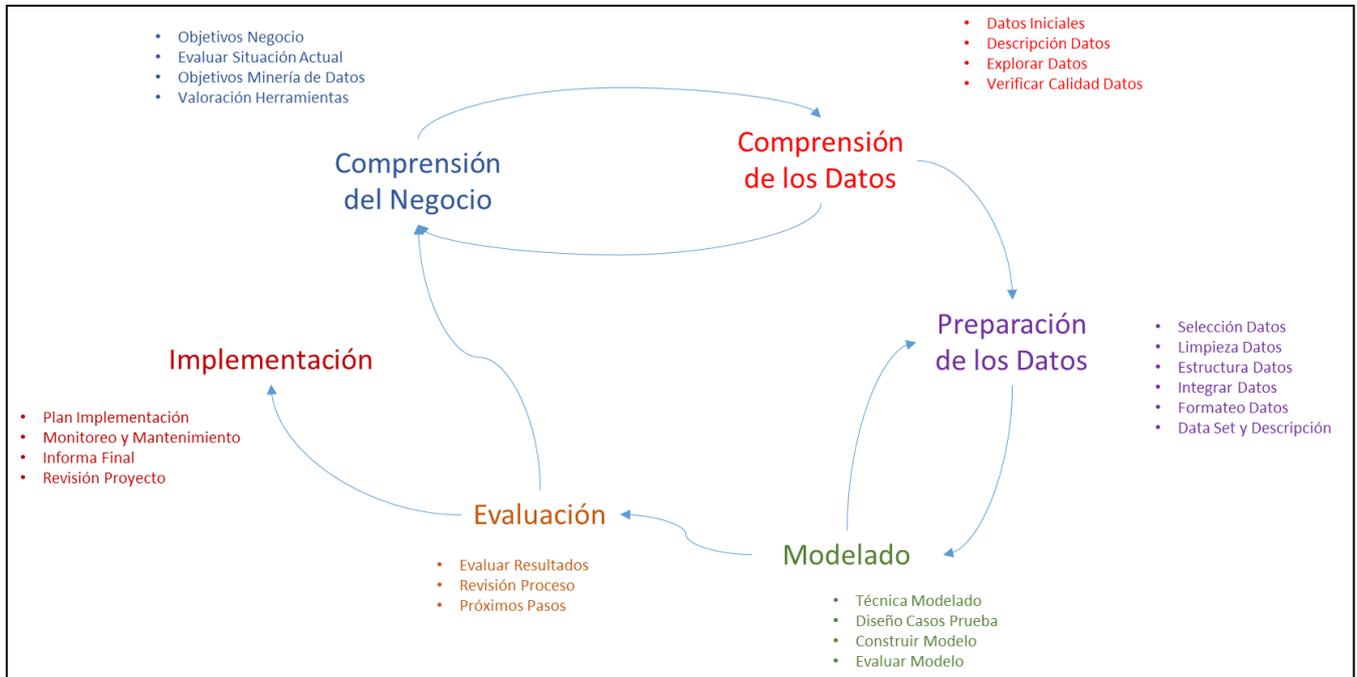


Figura 1: Ciclo de Vida de un Proyecto de Explotación de Datos

El ciclo de vida de un proyecto de Explotación de Datos consiste en seis fases cuya sucesión no es rígida, y se puede mover entre ellas siempre que se requiera. Las flechas indican las dependencias más importantes y frecuentes entre las fases. El círculo exterior simboliza la naturaleza cíclica de los proyectos de Explotación de Datos.

La metodología se presenta en términos de un proceso jerárquico. Consiste en juego de tareas descritas en niveles de abstracción (de lo general a lo específico): la fase, la tarea genérica o subfase, la tarea especializada y el caso del proceso.

El contexto de CRISP-DM se maneja entre lo genérico y el nivel especializado, dentro del cual se distinguen cuatro dimensiones diferentes:

- Dominio de la aplicación. Especifica el área en que el proyecto de explotación de datos tiene lugar.
- Tipo de problema. Describe la clase y objetivos del proyecto.
- Aspecto técnico. Cubre procesos específicos de la explotación de datos, describe diferentes desafíos que normalmente ocurren.
- Herramienta técnica específica que se aplica durante el proceso de explotación de datos.

A continuación, en las figuras 2, se detallan las fases que componen a la metodología CRISP-DM:



Figura 2: Fases componentes de la metodología CRISP-DM

A continuación, se detalla cómo se compone cada una de las fases de la metodología:

Fases de CRISP-DM

Fase 1: Comprensión del negocio

Tareas Componentes	Salidas a generar
Determinar los objetivos del negocio	<ul style="list-style-type: none"> ➤ Escenario actual ➤ Objetivos del negocio ➤ Criterios de éxito del negocio
Evaluación de la situación	<ul style="list-style-type: none"> ➤ Inventarios de recursos ➤ Requisitos, supuestos y requerimientos ➤ Riesgos y contingencias ➤ Terminología ➤ Costos y beneficios
Determinar objetivos de Minería de Datos	<ul style="list-style-type: none"> ➤ Objetivos de minería de datos. ➤ Criterios de éxito de minería de datos.
Realizar el Plan del Proyecto	<ul style="list-style-type: none"> ➤ Plan de proyecto ➤ Valoración inicial de herramientas

Tabla 1: Composición del Negocio

Fase 2: Comprensión de los datos

Tareas Componentes	Salidas a generar
Recolectar los datos Iniciales	<ul style="list-style-type: none"> ➤ Reporte de recolección de datos iniciales
Descubrir datos	<ul style="list-style-type: none"> ➤ Reporte de descripción de los datos
Exploración de los datos	<ul style="list-style-type: none"> ➤ Reporte de exploración de datos
Verificación de calidad de datos	<ul style="list-style-type: none"> ➤ Reporte de calidad de datos

Tabla 2: Comprensión de los datos

Fase 3: Preparación de los datos

Tareas Componentes	Salidas a generar
	<ul style="list-style-type: none"> ➤ Data set ➤ Descripción del data set
Seleccionar los datos	<ul style="list-style-type: none"> ➤ Inclusión/ exclusión de datos
Limpiar los datos	<ul style="list-style-type: none"> ➤ Reporte de calidad de datos limpios
Estructurar los datos	<ul style="list-style-type: none"> ➤ Derivación de atributos ➤ Generación de registros
Integrar los datos	<ul style="list-style-type: none"> ➤ Unificación de datos
Formato de los datos	<ul style="list-style-type: none"> ➤ Reporte de calidad de los datos

Tabla 3: Preparación de los datos

Fase 4: Modelado

Tareas Componentes	Salidas a generar
Seleccionar una técnica de modelado	<ul style="list-style-type: none"> ➤ La técnica modelada ➤ Supuestos del modelo
Generar el diseño de las pruebas	<ul style="list-style-type: none"> ➤ Diseño de las pruebas
Construir el modelo	<ul style="list-style-type: none"> ➤ Configuración de parámetros ➤ Modelo ➤ Descripción del modelo
Evaluar el modelo	<ul style="list-style-type: none"> ➤ Evaluar el modelo ➤ Revisación de la configuración de parámetros

Tabla 4: Modelado

Fase 5: Evaluación

Tareas Componentes	Salidas a generar
Evaluar Resultado	<ul style="list-style-type: none"> ➤ Valoración de resultados mineros con respecto al éxito del negocio ➤ Modelos aprobados
Proceso de revisión	<ul style="list-style-type: none"> ➤ Revisión del proceso
Determinar Próximos pasos	<ul style="list-style-type: none"> ➤ Listado de posibles acciones

Tabla 5: Evaluación

Fase 6: Implementación

Tareas Componentes	Salidas a generar
Plan de Implementación	<ul style="list-style-type: none"> ➤ Plan de Implementación
Plan de monitoreo y mantenimiento	<ul style="list-style-type: none"> ➤ Plan de monitoreo y mantenimiento
Informe Final	<ul style="list-style-type: none"> ➤ Informe final ➤ Presentación Final
Revisión del proyecto	<ul style="list-style-type: none"> ➤ Documentación de la experiencia

Tabla 6: Implementación

Análisis de las fases de CRISP-DM

A continuación, se detalla cómo se compone cada una de las fases metodológicas de CRISP-DM:

COMPRESIÓN DEL NEGOCIO

Esta fase requiere la valoración de varios factores, la comprensión de lo que es el negocio, adonde va, y que preguntas del negocio necesitan ser contestadas para llegar allí.

Actividades que se realizan:

- Identificación de objetivos de negocio y criterios de éxito
- Realización de una valoración circunstancial (recursos, supuestos, riesgos, costos y beneficios)
- Determinación de metas de Explotación de Datos y criterios de éxito
- Producción de un plan de proyecto

Determinar los Objetivos del Negocio:

El primer objetivo del analista de datos es comprender completamente la perspectiva del negocio, es decir, lo que el cliente realmente quiere lograr.

A menudo un cliente tiene muchos objetivos que compiten y deben ser equilibrados apropiadamente. La meta del analista es encontrar factores importantes que pueden influir en el resultado del proyecto.

Una posible consecuencia de descuidar esta fase es malgastar el tiempo y trabajo realizado a responder preguntas que no se corresponden con el objetivo del negocio.

Salidas a generar:

- **Escenario actual:**
Registra la información actual del negocio al principio del proyecto.
- **Objetivos del negocio:**
Describir el objetivo principal del cliente, desde la perspectiva del negocio. Además del objetivo principal, hay otras preguntas del negocio relacionadas con el cliente que sería bueno para el negocio conocer.
- **Criterios de éxito del negocio:**
Describir los criterios para un resultado exitoso o útil al proyecto desde el punto de vista del proyecto. Esto debe ser lo más específico posible para poder medir el mismo objetivamente.

Evaluación de la Situación Actual:

Esta tarea un estudio más detallado sobre todos los recursos, supuesto y otros factores que deben ser considerados determinado la meta de análisis de datos y el plan de proyecto. En la tarea anterior, su objetivo era conseguir rápidamente el factor principal de la situación, aquí se quiere encontrar los detalles.

Salidas a generar:

- **Inventario de Recursos:**
Listar los recursos disponibles para el proyecto, incluye: recursos humanos (expertos del negocio, expertos en datos, asistencia técnica, expertos en explotación de datos, etc.), datos (extractos fijos, acceso a los datawarehouse o a los operacionales), recursos técnicos (fundamentalmente hardware y software).
- **Requisitos, supuestos y requerimientos:**
Listar todos los requisitos del proyecto incluyendo: horarios de realización, alcance, calidad de los resultados y seguridad, además se deben evaluar los posibles problemas legales.
Como parte de esta salida, se debe asegurar que le permitan asegurar los datos.
Se deben listar los supuestos realizados en el proyecto. Estos pueden ser sobre los datos que pueden verificarse durante el proceso de explotación de datos, pero también puede incluir supuestos no controlables sobre el negocio en el resto del proyecto.
Listar los requerimientos del proyecto. Estos pueden ser requerimientos de la disponibilidad de los recursos, pero también puede incluir requerimientos tecnológicos como el tamaño de los datos.
- **Riesgos y Contingencias:**
Listar los riesgos o contingencias que pueden demorar el proyecto o que a causa de ellos el proyecto pueda fallar. Listar los planes de contingencia correspondientes, es decir, que acciones serán tomadas si estas situaciones ocurren.
- **Terminologías:**
Organizar un glosario de términos pertinente al proyecto. El mismo debe estar compuesto de:
 - ❖ Un Glosarios de términos del negocio disponible en el proyecto. Este glosario es útil para la producción de conocimiento.
 - ❖ Un glosario de términos del proceso Explotación de Datos.
- **Costos y beneficios:**
Construir un análisis de costo-beneficio para el proyecto, que compare los costos de llevar a delante el proyecto con los beneficios potencial a obtener gracias a él.

Determinar los Objetivos del Proceso de Minería de Datos:

Esta fase tiene como objetivo definir las metas a alcanzar con el desarrollo del proyecto de Explotación de Datos, por ejemplo, una meta a alcanzar puede ser “realizar ventas por catálogos a clientes existentes)

Salidas a Generar:

- **Objetivos de minería de datos:**
Describir los rendimientos intencionales del proyecto que habilita el logro de los objetivos del negocio.
- **Criterios de éxito del proceso de minería de datos:**
Definir los criterios para un resultado exitoso en términos técnicos, por ejemplo, el nivel de exactitud de la predicción.

Plan de Proyecto:

Describir el plan intencional para alcanzar las metas. El plan debe especificar los pasos a realizar durante el resto del proyecto, incluyendo una selección inicial de herramientas y técnicas.

Salidas a Generar:

➤ **Plan de Proyecto:**

Listar las actividades a desarrollar en el proyecto, junto con su duración, los recursos requeridos, las entradas, las salidas y las dependencias. Donde es posible hacer explícito las iteraciones de gran potencia en los datos del proceso de explotación de datos. Como parte del plan del proyecto, es también importante analizar las dependencias entre los tiempos y los riesgos. Se debe señalar los resultados de estos análisis explícitamente en el plan de proyecto, con acciones y recomendaciones por si los riesgos aparecen.

El plan de proyecto es un documento dinámico en el sentido que al final de cada fase una revisión de progreso y logros es necesaria y una actualización del mismo es recomendada de acuerdo con ello. Se debe especificar los puntos de revisión ya que estos son parte del plan de proyectos.

➤ **Validación inicial de Herramientas:**

Al final de la primera fase, se realiza también una valoración inicial de herramientas y técnicas. Aquí se selecciona la herramienta de Explotación de Datos que apoya los métodos para las diferentes fases del proceso.

Es importante evaluar herramientas y técnicas en esta fase del proceso, porque estas posiblemente influirán en todo el proyecto.

COMPRESIÓN DE LOS DATOS

Esta fase involucra las tareas que hacen la comprensión de los datos, para lo cual, los datos se deben describir, coleccionar, organizar, verificar y limpiar a priori del análisis de los mismos. Esto puede consumir mucho tiempo y es crítico para el proyecto de Explotación de Datos.

Actividades que se realizan:

- Auditoría de los datos
- Extracción de datos de una base de datos o datawarehouse
- Unificación de tablas desde una base de datos corporativa
- Combinación de archivos de diferentes sistemas
- Reconciliación de valores inconsistentes en campos
- Identificación de valores perdidos, datos incorrectos, o datos externos
- Selección de datos
- Reestructuración de los datos según el análisis lo requiere
- Transformación de campos (tomando proporciones, diferencias, etc.)

SUBFASES

Recolectar los Datos Iniciales:

Adquirir o acceder a los datos definidos en el plan de recursos del proyecto. Esta recolección inicial incluye datos que se cargan, si es necesario, para la comprensión de los mismos.

Es importante destacar que si los datos se adquieren de múltiples fuentes, la integración de los mismos es una tarea adicional a realizar, aquí o en la posterior fase de preparación de los datos.

Salidas a Generar:

➤ **Reporte de Descripción de los Datos:**

Listar el data set o data sets adquiridos, junto con sus localizaciones dentro del proyecto, los métodos usados dentro de la recolección y cualquier problema encontrada. La registración de problemas encontrados y cualquier solución lograda, ayuda en repeticiones futuras de este proyecto o proyectos similares.

Describir los Datos:

Examina la totalidad o la superficie de los datos adquiridos e informa los resultados.

Salidas a Generar:

➤ **Reporte de Recolección de los datos iniciales:**

Describe los datos que han sido adquiridos, indicando: formato de los datos, cantidad de datos. Por ejemplo se debe describir el número de registros y campos en una base de datos, las identidades de los campos y cualquier otro dato de la superficie de los datos que se han descubierto.

Exploración de los Datos:

Esta tarea trae preguntas a la Explotación de Datos, que pueden responderse usando encuestas, visualización e informes. Esto incluye: distribución de atributos importantes, por ejemplo, el atributo designado de una tarea de predicción; relación entre pares o los números pequeños de atributos; resultado de agregaciones y análisis estadístico. Estos análisis pueden dirigir los objetivos de la Explotación de Datos, y también pueden contribuir o refinar la descripción de los datos y la calidad de los reportes.

Salidas a Generar:

➤ **Reporte de Exploración de datos:**

Describir resultados de esta tarea que incluye hallazgos o la hipótesis inicial y su impacto en el resto del proyecto. Si es apropiado, incluye los gráficos y planos que indican características de los datos interesantes para un examen detallado.

Verificar la Calidad de los Datos:

Se debe verificar la calidad de los datos y realizarse preguntas del tipo: ¿están completos los datos?; ¿cubren todos los casos requeridos?; los datos, ¿son correctos o tienen errores?, si hay errores, ¿cuán común son ellos?; ¿hay valores perdidos en los datos?, en cuyo caso ¿cómo se representan?, ¿dónde ocurren estos errores y cuan común son?

Salida a Generar:

➤ **Reporte de Calidad de Datos:**

Listar los resultados de la comprobación de calidad de los datos; si los problemas de calidad existen, se deben listar las posibles soluciones. Las soluciones a los problemas de calidad de datos generalmente dependen de procesos pesados de datos y del conocimiento del negocio.

PREPARACIÓN DE LOS DATOS

Esta fase involucra las tareas que hacen a la preparación de los datos, es decir seleccionar los datos, limpiarlos, estructurarlos integrarlos y definir el formato final de los mismos.

SUBFASES

Salidas a Generar:

➤ **Data set:**

Este es el data set/s producido/s por la fase de preparación de los datos que es usado con el modelo o en el trabajo de análisis del proyecto.

➤ **Descripción del Data set:**

Describe el data set/s definido en el punto anterior

Selección de los Datos

En esta subfase se decide que datos son usados para el análisis. El criterio de selección aplicado debe ser lo suficientemente amplio para permitir incluir datos de relevancia en función de los objetivos del proyecto de

Explotación de Datos, como así también mantener las normas de calidad y requerimientos técnicos (límites de volumen o tipos de datos).

Salidas a Generar:

- **Inclusión/ Exclusión de datos:**
Lista de datos a Incluir/Excluir y las razones para tomar estas decisiones.

Limpieza de Datos:

Optimizar la calidad de los datos al nivel requerido por las técnicas de análisis seleccionadas. Esto puede implicar selección de subconjuntos limpios de los datos, la inserción de valores por defecto convenientes o aplicación de técnicas de estimación de perdido.

Salidas a Generar:

- **Reporte de limpieza de datos:**
Describir que decisiones se tomaron y que acciones se tomaran para solucionar los problemas de calidad de los datos que se informaron durante la tarea de calidad de datos. Aquí se deben considerar, entre otras cosas, las transformaciones de los datos por limpiezas, los propósitos de los datos y el posible impacto de los mismos en los resultados del análisis.

Estructura de los Datos:

Esta incluye las operaciones de preparación y construcción de datos, como así también la producción de atributos derivados, completando con los nuevos registros o los valores transformados con los atributos existentes.

Salida a Generar:

- **Derivación de atributos:**
Los atributos derivados, son nuevos atributos que se construyen a partir de uno o más atributos existentes en el mismo registro.
- **Generación de registros:**
Describir la creación de nuevos registros. Por ejemplo, crear registros nuevos para los clientes que no realizaron compras el último año.

Integración de los Datos:

Se aplican métodos que combinan información de múltiples tablas o archivos para crear nuevos registros o valores.

Salida a Generar:

- **Unificación de los datos:**
El unificar datos se refiere a unir dos o más tablas o archivos que contienen información diferente sobre el mismo objeto.
Los datos unidos también cubren agregaciones. La agregación se refiere a los funcionamientos donde los nuevos valores son computados resumiendo información.

Formato de los Datos:

Las transformaciones de estructuras se refieren a las modificaciones principalmente sintácticas realizadas a los datos que no cambian su significado, pero podría requerirse por la herramienta del modelo.

Salida a Generar:

- **Reporte de formato de los Datos:**

Algunas herramientas tienen requisitos en el orden de los atributos, como el primer campo es un único identificador para cada registro o el último campo siendo el campo del resultado total del modelo a predecir.

En función de lo expuesto en el párrafo anterior, también puede ser necesario cambiar el orden de los registros en el data set. Hay herramientas que requieren que los estén ordenados conforme al valor del atributo resultado.

MODELADO

El arte del trabajo especializado del proceso de Explotación de Datos toma lugar en esta fase. Si se prueban hipótesis específicas o se corren métodos del descubrimiento automatizados, se interpretan los resultados de análisis realizados en esta fase en el contexto de las preguntas del negocio originales.

Actividades que se realizan:

- Construcción del Modelo
- Interpretación del Modelo
- Validación / Evaluación del Modelo
- Refinamiento y Repetición

SUBFASES

Seleccionar una Técnica de Modelado:

Como primer paso del plan, se selecciona la técnica de modelado a utilizar. Considerando que ya, posiblemente, se seleccionó una herramienta de negocio, esta tarea se refiere a la técnica de modelado específica, por ejemplo árboles de decisión, reglas de decisión, redes neuronales, etc. Si se considera necesario aplicar múltiples técnicas, se debe realizar esta tarea, para cada una de las técnicas, separadamente.

Salida a Generar:

- **Técnica de modelado:**
Descripción de las técnicas de modelado que se utilizaran.
- **Supuestos de modelado:**
Muchas técnicas específicas generan supuestos específicos en los datos, por ejemplo, todos los atributos tienen una distribución uniforme, o no existen valores perdidos. Todos estos supuestos deben ser registrados

Generar el Diseño de las Pruebas:

Antes de generar el modelo se debe generar un procedimiento o mecanismo para probar la calidad y validez del modelo.

Salidas a Generar:

- **Diseño de las pruebas:**
Se debe describir el plan de pruebas y los modelos. Un componente principal del plan de pruebas es como dividir el data set disponible en datos de prueba y datos de aprobación.

Construir el Modelo:

Se debe ejecutar la herramienta de modelado en el data set preparado para crear uno o más modelos.

Salidas a Generar:

- **Configuración de parámetros:**

En general, la mayoría de las herramientas de modelado, proveen un conjunto de parámetros de ajuste a configurar. Se deben listar el conjunto de parámetros y los valores escogidos para los mismos.

- **Modelos:**
Describir los modelos reales generados por la herramienta.
- **Descripción del modelo:**
Se describe el modelo resultante, mediante un informe que detalle la interpretación de los modelos y documente cualquier dificultad encontrada con su significado.

Evaluar el Modelo:

El ingeniero en Explotación de Datos debe interpretar los modelos según su dominio de conocimiento, los datos, el criterio de éxito y el plan de pruebas definido.

Esta tarea interfiere con la fase subsiguiente. Considerando que los datos que se “Explotan” a juicio del ingeniero define el éxito de la aplicación de modelado y las técnicas de descubrimiento. Él comunica, más tarde, a los analistas del negocio y expertos en el dominio de la aplicación los resultados obtenidos, para discutir con estos los resultados de la explotación de datos en el dominio del negocio.

El ingeniero en Explotación de Datos intenta alinear los datos a los modelos. Él analiza los modelos según los criterios de evaluación. En la mayoría de los proyectos, el ingeniero aplica la misma técnica más de una vez o intenta generar los resultados con técnicas alternativas.

Salidas a Generar:

- **Evaluación del modelo:**
Se deben resumir los resultados de la tarea, detallando la calidad de los documentos generados.
- **Revisión de la configuración de parámetros:**
Según las valoraciones, se revisan las configuraciones de los parámetros para las próximas corridas del modelo. Se debe iterar el modelo construido y la configuración de los parámetros hasta encontrar el mejor modelo, documentando todas las revisiones y valoraciones.

EVALUACIÓN

En esta fase se evalúan los resultados del proceso de Explotación de Datos, se formulan recomendaciones y se prepara el material de apoyo.

Actividades que se realizan:

- Evaluación de los resultados del proceso de Explotación de Datos en términos de criterios de éxito del negocio
- Determinación de las recomendaciones y las posibles acciones
- Toma de decisiones

SUBFASES

Evaluar Resultado:

Las tareas de evaluación vistas con anterioridad trataban de la exactitud y generalidad del modelo. Esta fase evalúa el grado en que el modelo reúne los objetivos del negocio y busca determinar si existe algún criterio de negocio por el cual este modelo es deficiente. Otra opción de evaluación es probar el modelo en aplicaciones con datos reales si hay tiempo y presupuesto.

Los resultados del proceso de Explotación de Datos necesariamente se relacionan con los objetivos del negocio originales y todos los otros hallazgos que necesariamente no están relacionados con el objetivo original del negocio, pero que podrían escalarse como desafíos adicionales.

Salidas a Generar:

- **Valoración de los resultados de minería de datos:**

Se deben resumir los resultados de valoración en términos de criterios de éxito del negocio, incluyendo una declaración final si el proyecto ya se encuentra cumpliendo los objetivos iniciales del negocio.

➤ **Modelo aprobado:**

Después de generar la valoración respecto al criterio de éxito del negocio, se aprueban los modelos que se encuentran seleccionados.

Revisión del Proceso:

En este punto el modelo resultante parece ser óptimo para la satisfacción de las necesidades del negocio. Por tanto se debe realizar una completa revisión de los datos para determinar si existe algún factor importante o tarea que se ha pasado por alto en algún momento.

Salidas a Generar:

➤ **Revisión del proceso:**

Generar un documento que resuma la revisión del proceso y resaltar las actividades que faltan y/o deben repetirse.

Determinar Próximos Pasos:

Según los resultados de la valoración del proceso se define como proceder en esta subfase. Se debe definir si el proyecto se puede dar por terminado, se debe comenzar con las iteraciones o si se debe preparar un nuevo proceso de Explotación de Datos.

En esta tarea se deben tener en cuenta los recursos y presupuesto restantes.

Salidas a Generar:

➤ **Listado de posibles acciones:**

Listar las potenciales acciones junto con las razones a favor y en contra de cada acción.

IMPLEMENTACIÓN

La fase de Implementación incluye tareas de puesta en producción y entrega del informe final. El plan de Implementación se desarrolla para supervisar los cambios que pueden producirse si el modelo de análisis no puede sostenerse. Esto puede comprender el análisis automatizado que produce advertencias cuando ciertos eventos ocurren (por ejemplo, si la brecha entre el predicho y observado excede una cantidad específica).

Actividades que se realizan:

- Producción y entrega del informe final
- Plan de monitoreo y mantenimiento de los resultados del proceso de Explotación de Datos
- Revisión del Proyecto
- Distribución de los resultados.

SUBFASES

Plan de Implementación:

Para implementar los resultados del Proceso de Explotación de Datos en el Negocio, se debe analizar el resultado de la evaluación y generar una estrategia de implementación.

Salida a Generar:

➤ **Plan de implementación:**

Resumir la estrategia del desarrollo incluso los pasos necesarios y como realizarlos.

Plan de Monitoreo y Mantenimiento:

El monitoreo y el mantenimiento son problemas importantes en los procesos de Explotación de Datos, para que el resultado sea parte del negocio diario. La preparación cuidadosa de una estrategia de mantenimiento ayuda a evitar innecesariamente largos periodos de uso incorrecto de los datos.

Salidas a Generar:

➤ **Plan de monitoreo y mantenimiento:**

Resumir el monitoreo y estrategia de mantenimiento, incluyendo los pasos a realizar y como realizarlo.

Informe Final:

Al final de proyecto, el líder de proyecto y su equipo escriben el informe final. Dependiendo del plan de implantación, este informe puede ser solo un resumen del proyecto y sus experiencias o puede ser una presentación comprensiva de los resultados del proceso de Explotación de Datos.

Salidas a Generar:

➤ **Informe final:**

Se genera un informe final que recoja la experiencia del desarrollo del proyecto.

➤ **Presentación final:**

Generalmente, al final del proyecto, se realiza una reunión donde se presentan los resultados del proyecto al cliente.

Revisión del Proyecto:

Una vez finalizado el proyecto, se debe evaluar lo que fue correcto y lo que salió mal, lo que se hizo bien y lo que necesita ser mejorado.

Salida a Generar:

➤ **Documentación de la experiencia realizada:**

Se debe resumir las experiencias importantes obtenidas durante el proyecto. Por ejemplo, las trampas en las que se cayó, los acercamientos engañosos o indirectas para seleccionar mejor las técnicas de Explotación de Datos.