**Planeamiento para el desarrollo de Soluciones Analíticas**

1. **Datos Generales**

|  |  |
| --- | --- |
| **Nombre del desarrollo** |  |
| **Versión (Major.Minor.Patch)** |  |
| **Riesgo asociado** |  |
| **Fecha (DD/MM/AAAA)** |  |

1. **Responsabilidades**

|  |  |
| --- | --- |
| **Desarrollo** | Área:  Colaborador:  Puesto:  Mail: |
| **Validación** | Área:  Colaborador:  Puesto:  Mail: |
| **Aprobación** | Área:  Colaborador:  Puesto:  Mail: |
| **Propietario** | Área:  Colaborador:  Puesto:  Mail: |
| **Usuario** | Área:  Colaborador:  Puesto:  Mail: |
| **Seguimiento** | Área:  Colaborador:  Puesto:  Mail: |

1. **Machine Learning Canvas[[1]](#footnote-1)**

| **Etapa** | **Comentario** |
| --- | --- |
| 1. **Propuesta de valor**   ¿Cuál es el objetivo del desarrollo? ¿Qué es lo que estamos intentando hacer?  ¿Por qué es importante? ¿Cómo se beneficiarán del modelo desarrollado, qué beneficios se obtendrán?  ¿Quién lo va a usar o se verá impactado (usuario final)?  ¿Existe alguna implementación que busque los mismos objetivos del desarrollo propuesto? |  |
| 1. **Fuentes de datos**   ¿Qué fuentes de datos podemos utilizar (internas, externas)?  ¿Se contará con datos alternativos o con un proceso de enriquecimiento de datos?  ¿Las fuentes de datos se mantendrán disponibles durante el ciclo de vida del proyecto?  Menciona: nombre del servidor, base de datos, tablas; otros archivos como xls, csv, txt, etc. |  |
| 1. **Recolección de datos**   ¿Detalle la estrategia para obtener los datos para el entrenamiento del modelo?  ¿Cuál será el costo de adquisición de los datos para producción?  ¿Las fuentes de datos se actualizan constantemente? |  |
| 1. **Features** (variables o características)   Detalla las variables de entrada que se utilizarán en el proceso de desarrollo del modelo (diccionario de variables). |  |
| 1. **Construcción de modelos**   ¿Cuántos modelos en producción necesitamos?  ¿Cuándo podríamos actualizarlos?  ¿De cuánto tiempo disponemos? |  |
| 1. **Tareas Machine Learning**   ¿Qué tipo de tarea predictiva se desarrollará (clasificación, regresión, serie de tiempo, otro)?  ¿Cuál es el horizonte de predicción?  ¿Cuáles son las definiciones para los posibles valores del target (objetivo)?  ¿En cuánto tiempo**[[2]](#footnote-2)** se realizará el desarrollo? |  |
| 1. **Decisiones**   ¿Cómo se usan las predicciones para tomar las decisiones que aportan valor?  ¿Cuáles son las acciones que se pretenden realizar? |  |
| 1. **Predicciones**   ¿El proceso de predicción se realizará en batch o tiempo real?  ¿Cuándo calculamos las predicciones de cuánto tiempo disponemos? ¿Cuál es el tiempo adecuado de entrega de información sin que afecte la continuidad del negocio? |  |
| 1. **Evaluación offline**   ¿Qué data de prueba se usará para evaluar el modelo?  Criterios de implementación: ¿De qué métricas disponemos para evaluar el modelo antes del pase a producción (considerar nombre del indicador, valores mínimos o intervalos)?  ¿Cada cuánto se realizarán las validaciones del modelo? De existir algún impacto, ¿se realizará algún ajuste en las predicciones y/o los segmentos propuestos? ¿Qué acciones se realizarán para mitigar este impacto? |  |
| 1. **Monitoreo y alertas tempranas**   De qué métricas y herramientas disponemos para evaluar el modelo en producción.  ¿Cuáles son las métricas para cuantificar el impacto de las predicciones y/o segmentaciones una vez se encuentre en producción lo desarrollado? (tanto para el usuario final como para el negocio). |  |

1. Referencia: https://thecanvasrevolution.com/product/the-machine-learning-canvas [↑](#footnote-ref-1)
2. Se consideran días hábiles laborables. [↑](#footnote-ref-2)