

## EXECUTIVE SUMMARY<sup>1</sup>

### -Lineamientos para el reporte del proyecto (Parte I)-

---

1 **Nota:** El proyecto tiene una ponderación del 30% sobre el total de la calificación de la asignatura (Evaluación continuada) y este 30% se evalúa con un 50% (competencia Kaggle) y 50%(reportes ejecutivos parte I y II).  
Idiomas para el reporte: Los idiomas vehiculares son el catalán y/o castellano. Pero dada la presencia de estudiantes Erasmus también puede entregarse en inglés, italiano, francés o portugués.

# PORTADA

- Debe contener título
- Identificación del Grupo (Número y ID en Kaggle, ejemplo: G10 – ID Kaggle: XXXX)
- Identificación de los Integrantes: Nombre, Universidad de procedencia, Grado de Procedencia, Email.
- Link al GITHUB con acceso.

# SECCIÓN 1 – ENTENDIMIENTO DEL NEGOCIO y ENTENDIMIENTO DE DATOS

- ENTENDIMIENTO DE NEGOCIO (1 slide) → Resumen del problema, contexto de aplicación, objetivo principal y objetivos secundarios. Listado de KPIs de control y opcional listado de KPIs de negocio que se deseen monitorear. - Fuentes externas de apoyo (Links).
- ENTENDIMIENTO DE DATOS (1 slide) → Metadata (Tabla o figura) con detalle de los datos: Nombre, Descripción, Tipología (Justificación de la misma), Rango. Reporte básico de calidad de datos. En caso de fuentes externas se deben listar con sus links.

## SECCIÓN 2 – FRAMEWORK METODOLÓGICO (1 slide)

Figura con la metodología escogida para desarrollar el proyecto (Debe seleccionar alguna de las vistas en clase). En dicha figura debe involucrar en cada paso a los responsables de cada tarea o actividad.

## SECCIÓN 3

### ADVANCED PREPROCESSING

- Exploratory Data Analysis (EDA) –Resultados relevantes (listado de los mismos) con apoyo de figuras, tests y/o tablas. Implicaciones de los resultados para la preselección de variables explicativas. (máximo 2 slides).
- Tratamiento de Missings y Outliers (Máximo 2 slides)
  - a) Estrategias implementadas.
  - b) Resultados Relevantes.
  - c) Decisiones
- Ingeniería de Variables (Feature Engineering) (1 o 2 slides)
  - a) Estrategias implementadas
  - b) Decisiones
  - c) Base de datos preprocesada (Ver ejemplo en la discusión en clase).

## SECCIÓN 4

### ASSOCIATION RULES

(1 o 2 slides)

- Workflow de trabajo para analizar el proyecto bajo la idea de un Market Basket Analysis
- Mención sobre la selección del min\_support y min\_confidence, tamaño de reglas, filtrado, etc.
- Resultados relevantes y posibles cambios, confirmación o updates en la base de datos obtenido de la sección 3.

## SECCIÓN 5

### PREPARACIÓN PARA EL MODELADO

(1 slide)

- Diagrama explicativo para las políticas de training, test y validación que serán empleadas en los modelos.
- Tratamiento de Balanceo (si aplica)
- Aquí también debe listar la estrategia de modelado por etapas (si es disponible) → atender discusión en clase.

## SECCIÓN 6

### PRIMEROS MODELOS – KNN y Naive Bayes

(1 o 2 slides)

- Descripción de la estrategia de modelado.
- Definición de predictores (cambios, transformaciones, adaptaciones, etc)
- Resultados con los KPIs exigidos en la competencia y en adicional (si aplica) de los KPIs listados en la sección 1.



SECCIÓN 7  
CONCLUSIONES y/o PRÓXIMOS PASOS  
(1 slide)

## TIPS DE AYUDA y otros señalamientos

- Use frases cortas y párrafos cortos de máx 3 o 4 líneas.
- Haga uso de recursos de visualización (resaltar resultados, colores de alerta o seguridad, tablas resumen, stickers, etc)
- **El reporte no debe superar las 15 slides (portada NO INCLUIDA).**
- Cualquier detalle que se quiera explicar debe citarse con un link a un banco de datos (nube, github, etc) del cual cada equipo me debe dar acceso. Esos bancos de resultados (con código, documentos, etc deben estar documentos mínimamente).
- Recuerde que la audiencia de su reporte en este caso será de conocedores de Data Mining, Estadística, Matemática y el dueño de la competencia (que asumimos no tiene un perfil determinado) por tanto el nivel de abstracción debe ser mixto: estilo técnico + entendimiento puntual y conciso que permita soportar procesos de toma de decisiones.
- La nota del proyecto dependerá de la calidad técnica de las secciones, enfoque de trabajo y resultados.
- No deberá presentar nada de manera oral, la asistencia el día 04 de noviembre es ALTAMENTE ACONSEJABLE ya que discutiré con cada equipo el reporte y podré de alguna manera daros un feedback sobre el ranking de los equipos para medir un poco el futuro/termómetro de la competencia.