



Hackathon-Enaex

Análisis exploratorio

Equipo Avatar:

Bastian Díaz
Felipe Bahamonde



Resumen

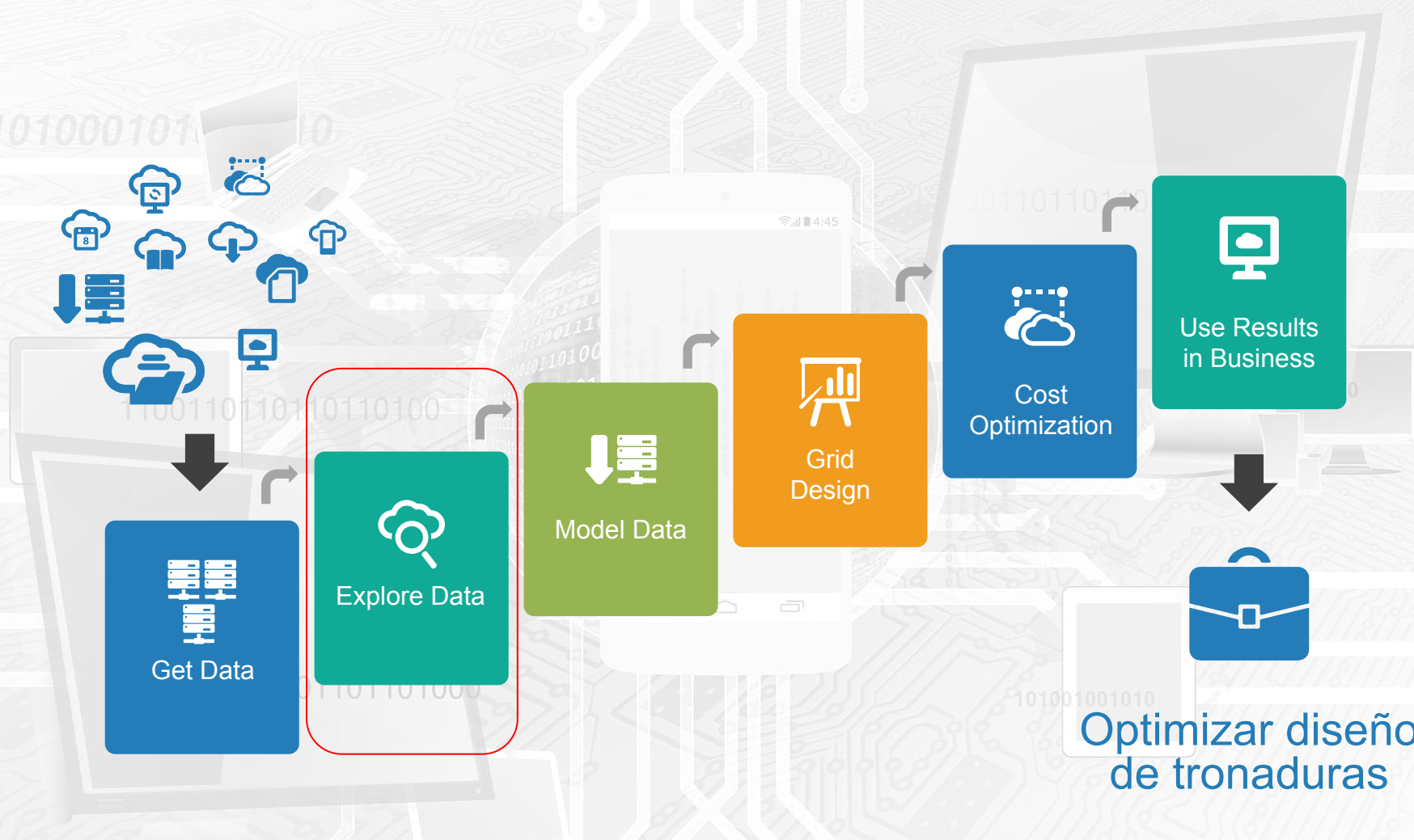
Esquema del desafío.

Limpieza

Curación de datos históricos.

Análisis Exploratorio

Insights relevantes para
modelamiento.



Get Data

Explore Data

Model Data

Grid Design

Cost Optimization

Use Results in Business

Optimizar diseño de tronaduras

Empty rows

- 2,1 % de las filas se encuentran totalmente vacías.
- Columnas **Cota, Norte y Este** presentan valores nulos (~0.26% de los datos).
- Columna Tiempo entre Pozos Filas ms presenta valores nulos (~0.05% de los datos).

Transformación de características

- Diámetro convertido a numérico.
- BxS debe ser limpiado y separado en 2 columnas (B y S).
- Tiempo entre Pozos Filas debe ser limpiado y separado en 2 columnas (Tiempo 1 y Tiempo 2 en ms.).
- P20 P30 P50 y P70 deben ser corregidos debido a datos mal ingresados.

Cantidad de columnas y tipos de datos

- 26 columnas en donde:
 - 10 corresponden a la target P.
 - 6 son categóricas.
 - 10 son de tipo numéricas.

Datos irrelevantes

- Banco y Cota presentan una correlación $> 90\%$, debido a que ambas hacen referencia a la misma característica.
- Explosivos tipo M se descartan del análisis debido a reglas propias del concurso.

Insights relevantes a partir del análisis*

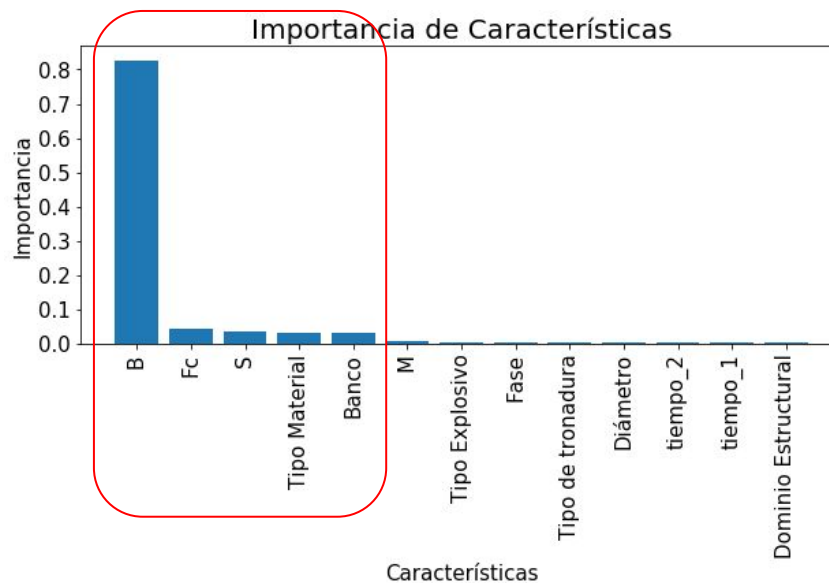
1. Las tronaduras en etapa 2 y 4 son las que **más registros tienen**.
2. La tronadura tipo 2, concentra el **86%** de los datos.
3. La tronadura tipo 1 produce material **más grande que la tipo 2**.
4. Los explosivos **B,C,J y L** se usan más del doble que los demás explosivos.
5. Entre los 2800 y 3000 m de altura el material tiene a **ser más grande** que el generado entre 3400 y 3600. Sobre los 3800 m el resultado es más disperso.
6. A medida que **aumenta FC aumenta** el tamaño del material fragmentado.

Insights relevantes a partir del análisis*

1. La roca tipo 1 se fragmenta en un material de mayor tamaño de la roca tipo 2, mientras que el tamaño del **lastre es más disperso**.
2. En cuanto a burden y espaciamiento se tiene una curva similar a una **normal skewed a la izquierda**.
3. Para la fase se tiene que entre las etapas **1 a 4 el material fragmentado aumenta** de tamaño y en las etapas **5 y 6 disminuye**.

Importancia de características

Análisis preliminar de importancia de características para modelar Ps a través de árboles de decisión simples:



Se destaca la importancia de la característica burden. Seguido por el factor de carga y tipo de material.

Resumen

Esquema del desafío.

Limpieza

Curación de datos históricos.

Análisis Exploratorio

Insights relevantes para
modelamiento.