



Comisión Calificadora de
Competencias en Recursos
y Reservas Mineras

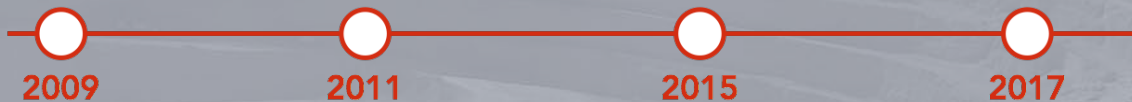
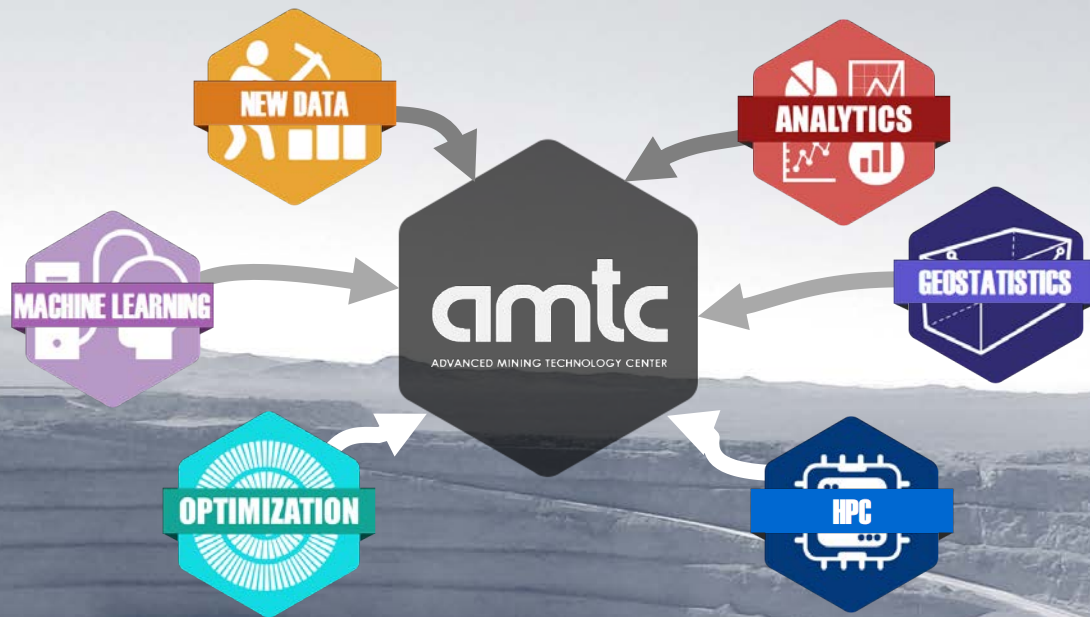


Ingeniería de Minas
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Uso de machine learning en la estimación de recursos minerales como alternativa a kriging

Felipe Navarro

fnavarro@alges.cl



El AMTC fue creado en marzo de **2009** como parte del Programa de CONICYT para Centros Científicos y Tecnológicos de Excelencia



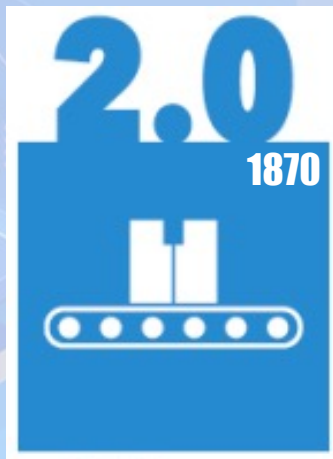
CONICYT
Comisión Nacional de Investigación
Científica y Tecnológica

REVOLUCIÓN INDUSTRIAL



Agua y Vapor

Producción mecánica



Electricidad

Producción en masa
Líneas de producción



Electrónica y TI

Sistemas y producción
automatizada, Internet,
Energía Nuclear



Automatización y datos

Digitalización

Inteligencia artificial

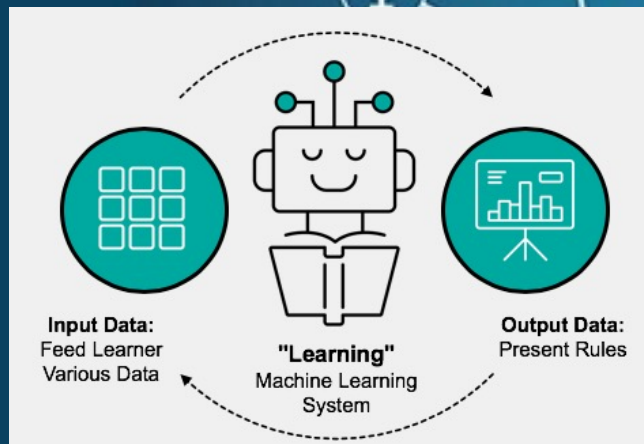
Sistemas que imitan el pensamiento y las acciones humanas

Machine Learning

Sistemas que aprenden basándose en grandes cantidades de datos

Deep Learning

Sistemas que funcionan con redes neuronales sin guía humana



¿QUÉ ES M.L.?

Es el campo de estudio que da la habilidad de aprender a las computadoras sin que hayan sido explícitamente programadas para ello [A. Samuel, 1959]

El estudio de algoritmos de computación que mejoran automáticamente su rendimiento gracias a la experiencia. Se dice que un programa informático aprende sobre un conjunto de tareas, gracias a la experiencia y usando una medida de rendimiento, si es que su desempeño en estas tareas mejora con la experiencia [T. Mitchell, 1997]

En el aprendizaje de máquinas un computador observa datos, **construye un modelo** basado en esos datos y **utiliza ese modelo** a la vez como una hipótesis acerca del mundo y una pieza de software que puede resolver problemas [Russell and Norvig, 2021]

¿QUÉ ES M.L.?

**Abstracción de
la realidad**

**Problemas
específicos**

**Tareas
repetitivas**



Extrapolación

Construir modelos

Datos disponibles

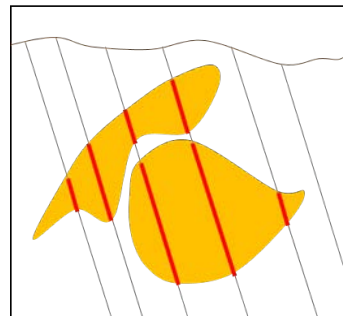
Obtener una **estimación** sin sesgo en volúmenes, leyes, tonelajes y cantidad de mineral

Considerando:

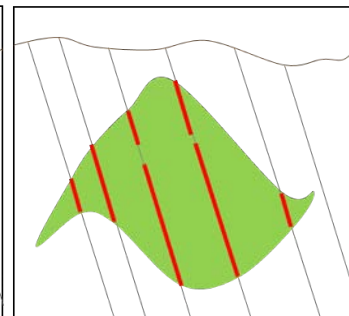
- La calidad de los datos
- La calidad del *modelo geológico*
- Está limitada por el número de muestras disponibles
- Requiere un procedimiento que sea repetible y auditable
- Requiere tiempo

“Si no se entiende la geología de la mineralización, según los lineamientos del código JORC, no se está preparado para estimar recursos minerales”

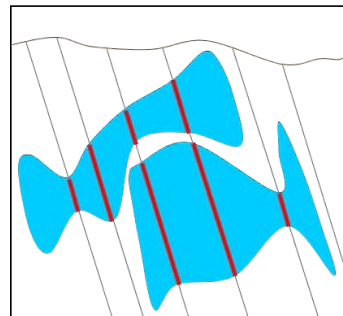
The Pessimistic Geologist



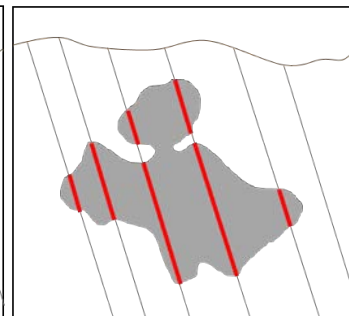
The Optimistic Geologist



The Geophysicist



The Mining Engineer



(*) Fuente: <https://www.micon-international.com/minex/>

Metodología tradicional

- Preparación de datos
- Análisis exploratorio de datos
- Definición de dominios de estimación
- Variografía
- Definición de parámetros de estimación
- Estimación
- Validación del modelo de bloques

Usando Machine Learning

- Preparación de datos
- Análisis exploratorio de datos
- **Definición de dominios de estimación**
- ~~Variografía~~
- ~~Definición de parámetros de estimación~~
- **Estimación**
- Validación del modelo de bloques

Podemos usar ML

no supervisado

Podemos usar ML

supervisado

no supervisado

No recibe ninguna etiqueta, por lo que tiene que encontrar por sí mismo la estructura de los datos de entrada.

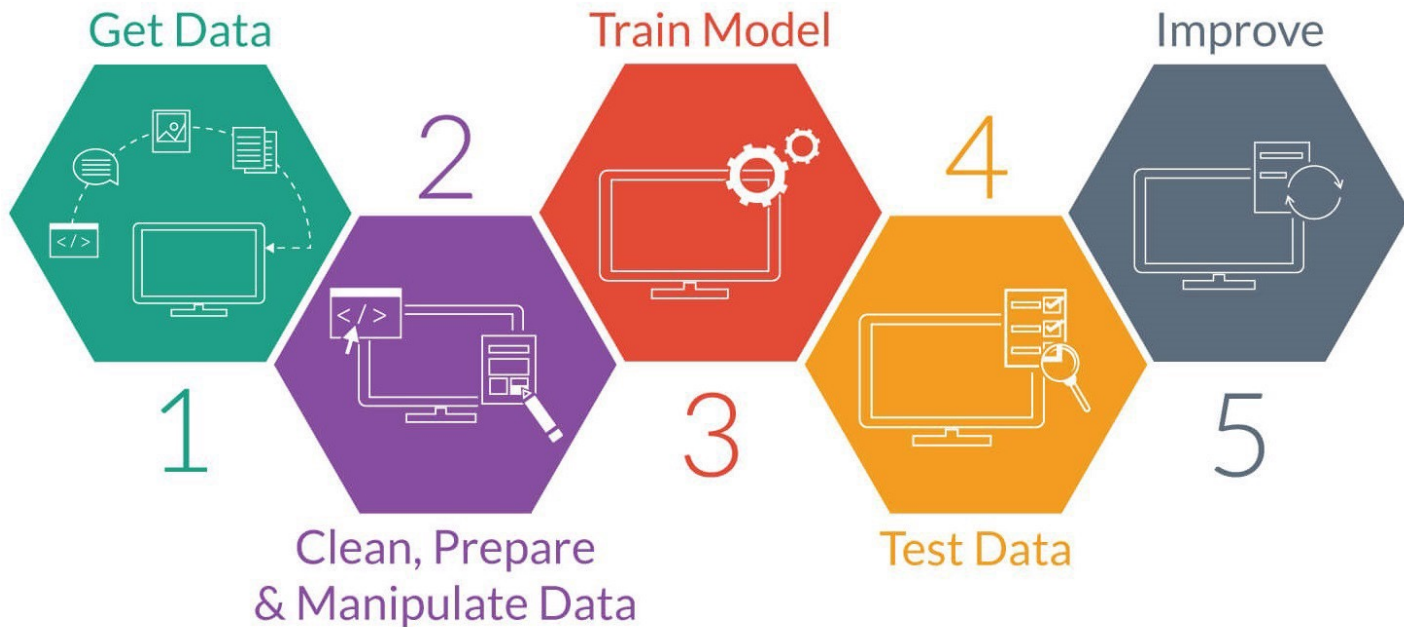
- Clustering (agrupamiento)
- Reducción dimensional
- Asociaciones

supervisado

Al modelo se le presentan ejemplos de entradas y sus salidas deseadas, dadas por un "profesor", y el objetivo es aprender una regla general que asigne las entradas a las salidas

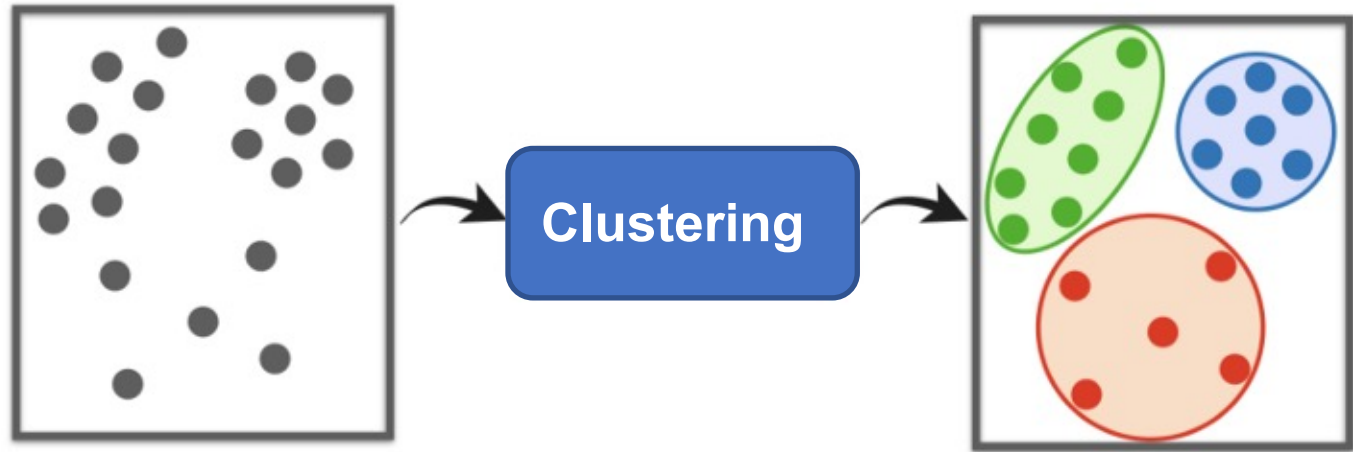
- Clasificación
- Regresión

Datos etiquetados



<https://medium.com/@pytyagi/work-flow-in-machine-learning-project-327eddb946b4>

- Las principales aplicaciones de **aprendizaje no supervisado** son:
 - Segmentación de conjuntos de datos por atributos compartidos
 - Detección de anomalías que no encajan en ningún grupo
 - Simplificación de datasets agregando variables con atributos similares

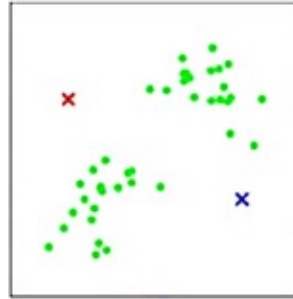


- Se busca **agrupar los datos sin una etiqueta determinada**
 - Encontrar una forma de establecer la similitud entre dos datos es un aspecto muy relevante

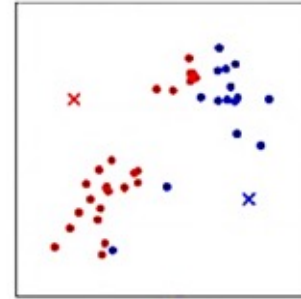
- Método de agrupamiento iterativo
- Asignación de clase según *centroide* más cercano
- Parámetros relevantes: Cantidad de clases a considerar



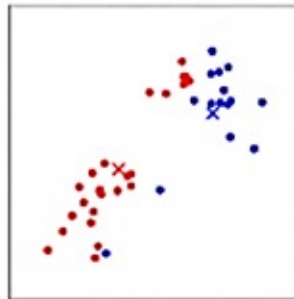
(a)



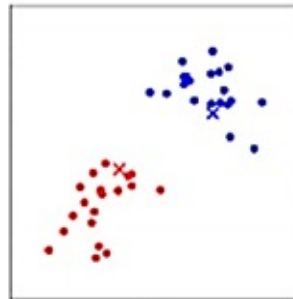
(b)



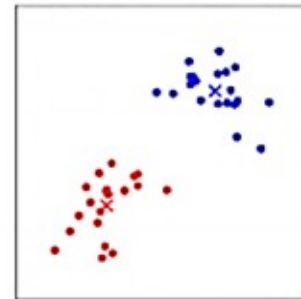
(c)



(d)



(e)



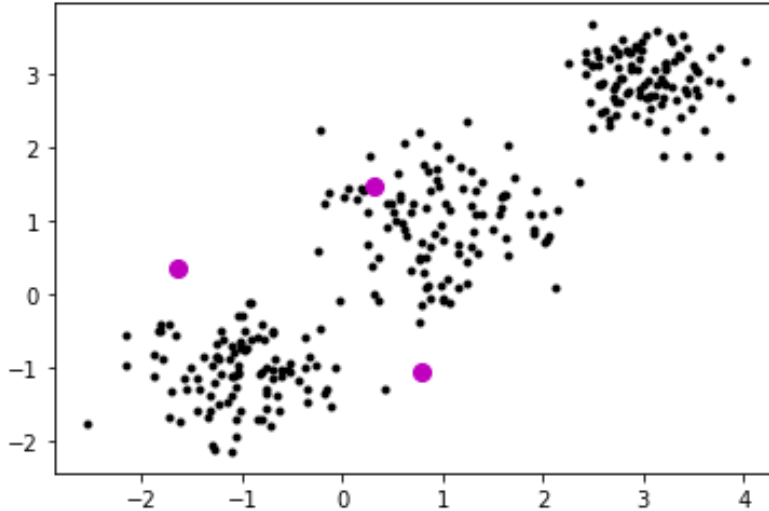
(f)

Iteración del método se produce entre dos etapas secuenciales:

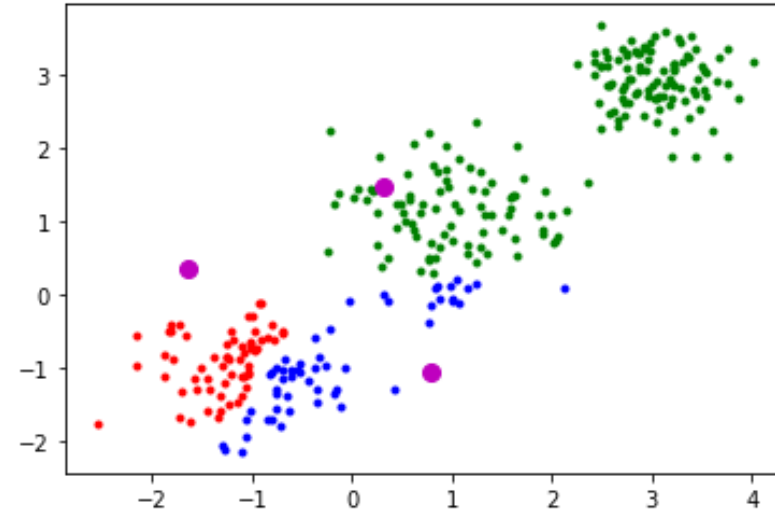
- Actualización de posiciones de centroides
- Reasignación de puntos

Iteración 1

Sorteamos los centroides



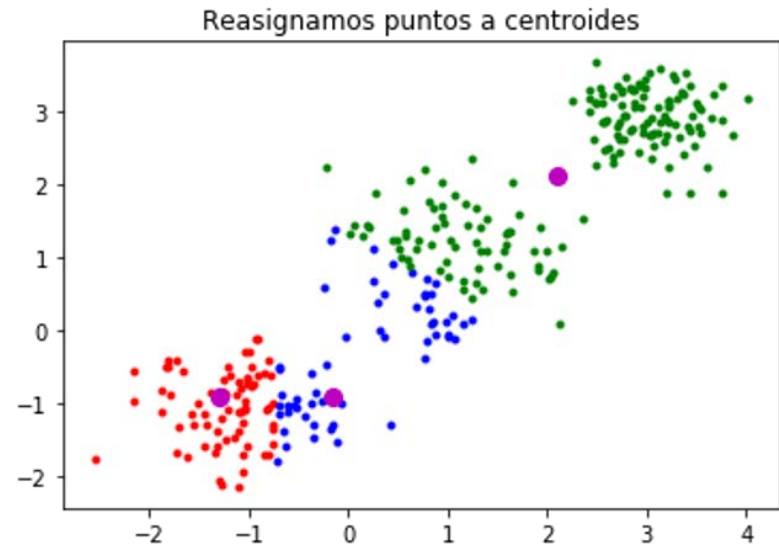
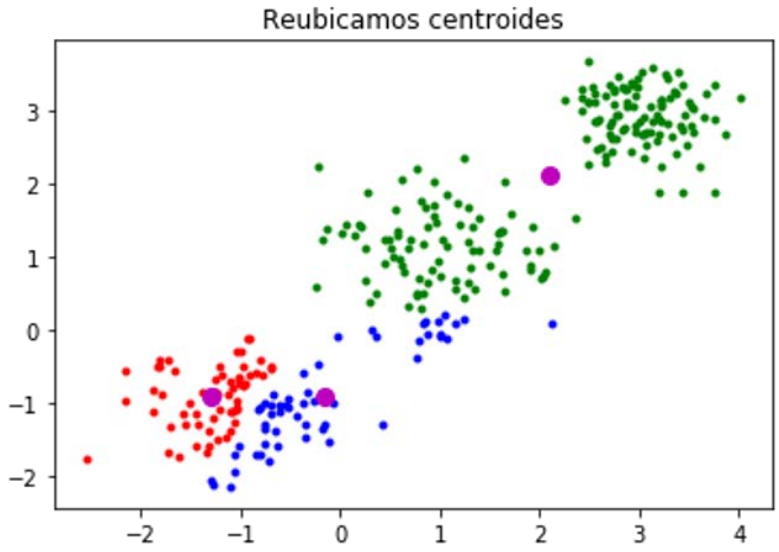
Asignamos puntos a centroides



Iteración del método se produce entre dos etapas secuenciales:

- Actualización de posiciones de centroides
- Reasignación de puntos

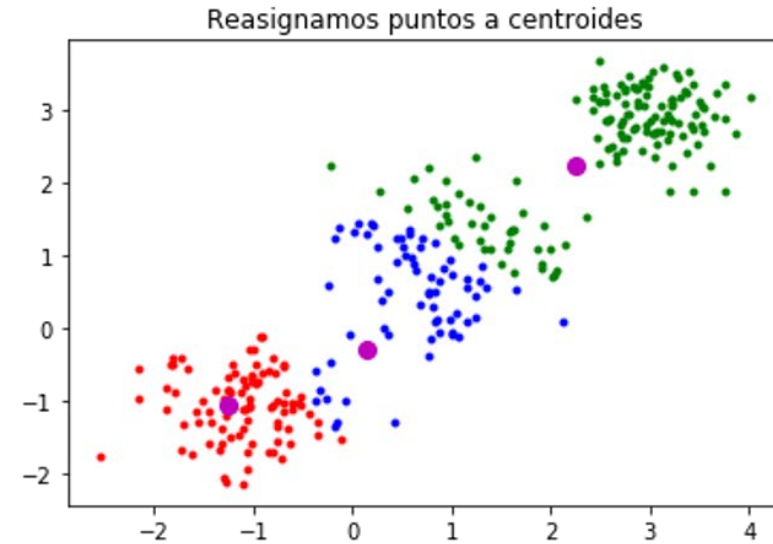
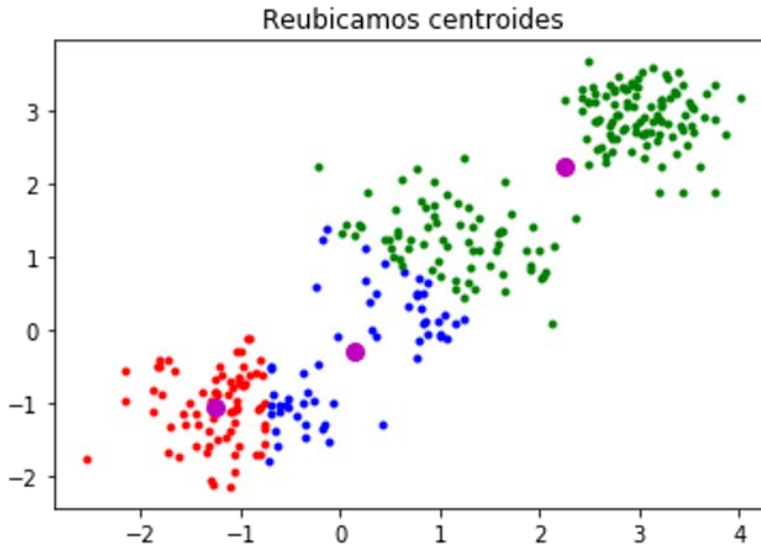
Iteración 2



Iteración del método se produce entre dos etapas secuenciales:

- Actualización de posiciones de centroides
- Reasignación de puntos

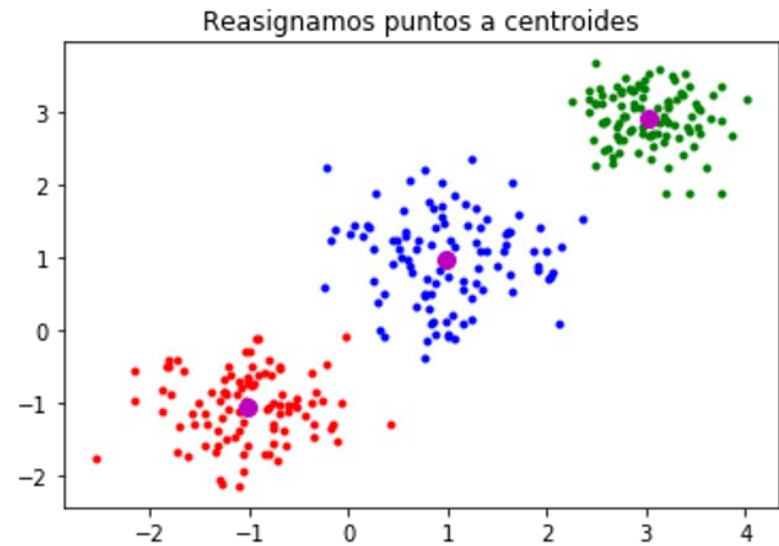
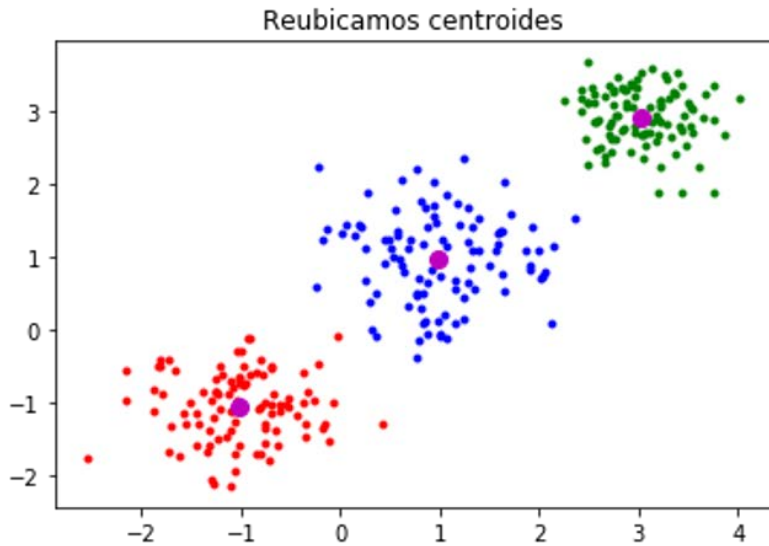
Iteración 3



Iteración del método se produce entre dos etapas secuenciales:

- Actualización de posiciones de centroides
- Reasignación de puntos

Iteración n

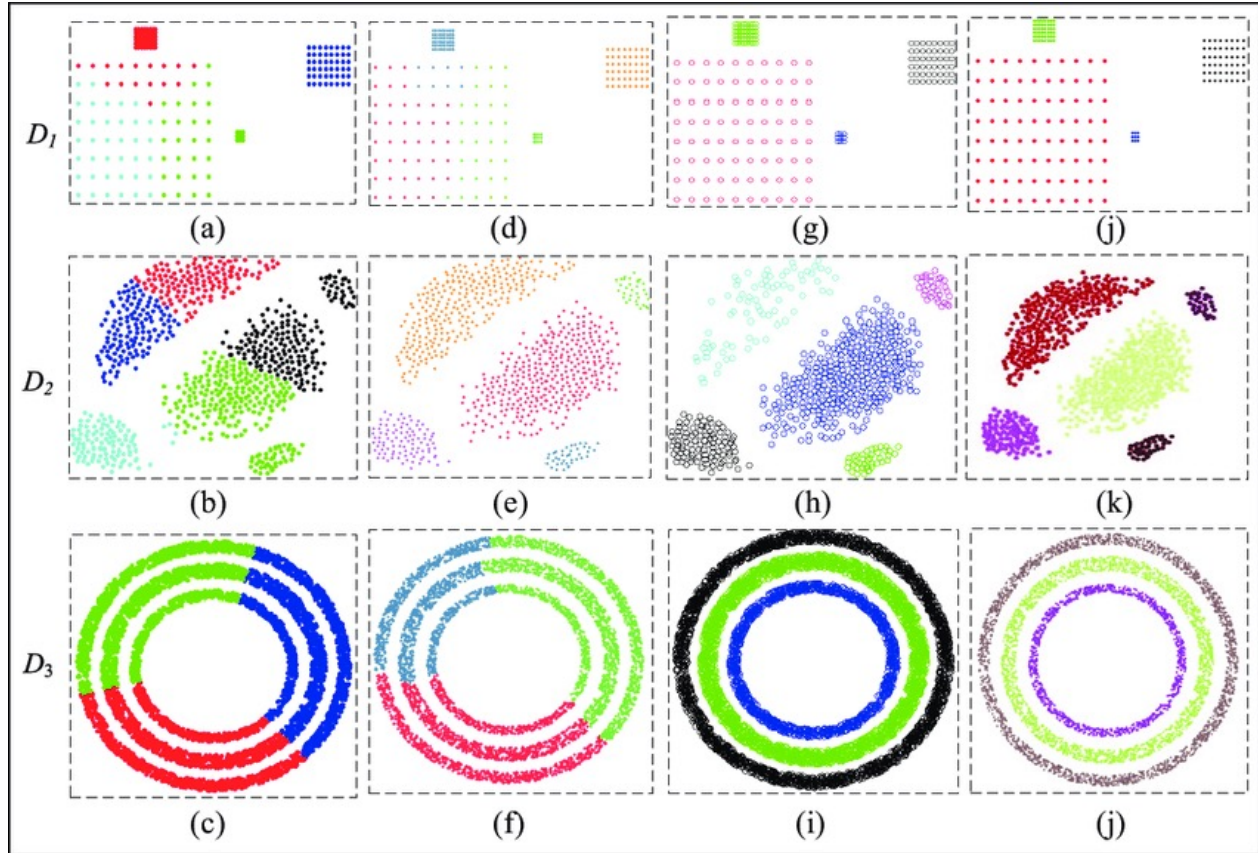


K-Means

Jerárquico

DBSCAN

Region-based



Aprendizaje supervisado:

- Establece la relación existente entre una o más variables (entrada) y una variable objetivo (respuesta)
- Cuando la variable objetivo es continua: regresión

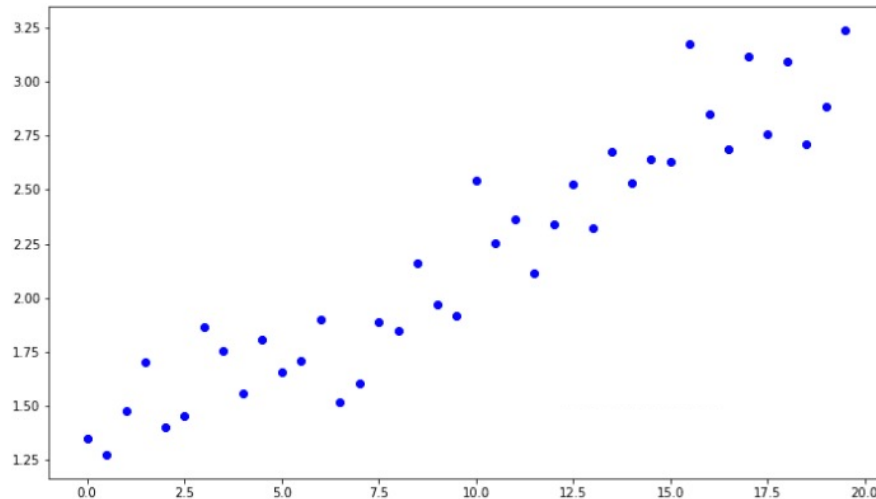
	Cd	Co	Cr	Cu	Land Use	Ni	Pb	Rock	Zn	Este (m)	Norte (m)
0	1.740	9.32	38.32	25.72	3	21.32	77.36	3	92.56	2386	3077
1	1.335	10.00	40.20	24.76	2	29.72	77.88	2	73.56	2544	1972
2	1.610	10.60	47.00	8.88	2	21.40	30.80	3	64.80	2807	3347
3	2.150	11.92	43.52	22.70	3	29.72	56.40	2	90.00	4308	1933
4	1.565	16.32	38.52	34.32	3	26.20	66.40	5	88.40	4383	1081

Ejemplo de variable objetivo



Caso preliminar: Regresión lineal univariada

- Modelo propuesto: $h(x) = a \cdot x + b$
 - Asume relación lineal entre x e y
- a, b -> parámetros del modelo
- x -> variable independiente o explicativa
- y -> variable dependiente o endógena
- Se cuenta con m pares de datos (x_i, y_i)



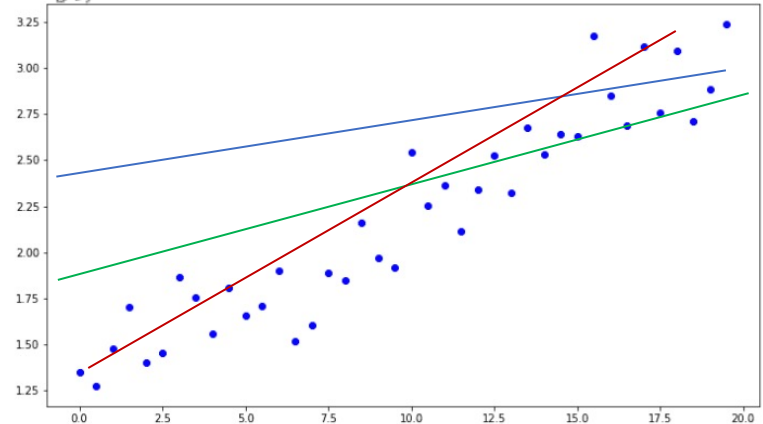
¿Cómo encontrar los "mejores parámetros"?

- Se prueban diferentes parámetros y se escogen los mejores según algún criterio

Función de pérdida

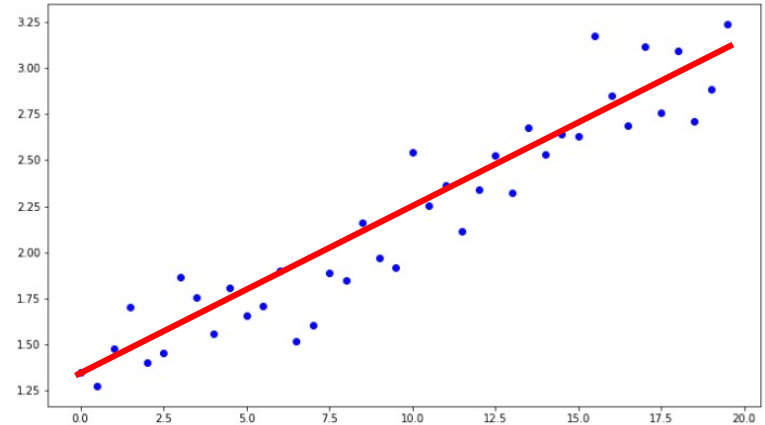
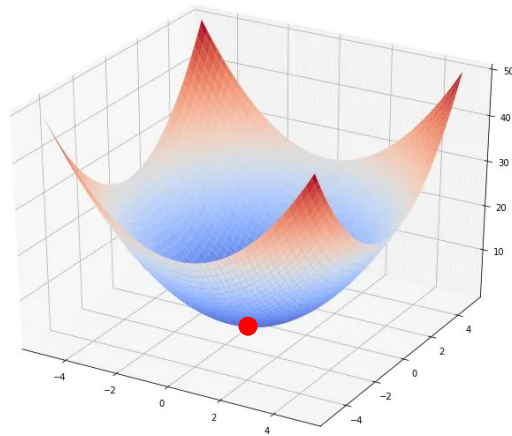
- Medición de "cuánto se desvía" el modelo de los datos
- Se comparan los valores $h(x)$ (el modelo) vs y (datos)
- Es una función de los parámetros $J(a, b)$
- Primera opción: Error cuadrático medio

$$J(a, b) = \sum_i (h(x_i) - y_i)^2 = \sum_i (a \cdot x_i + b - y_i)^2$$



Objetivo: minimizar la función de pérdida

- $\min J(a, b)$



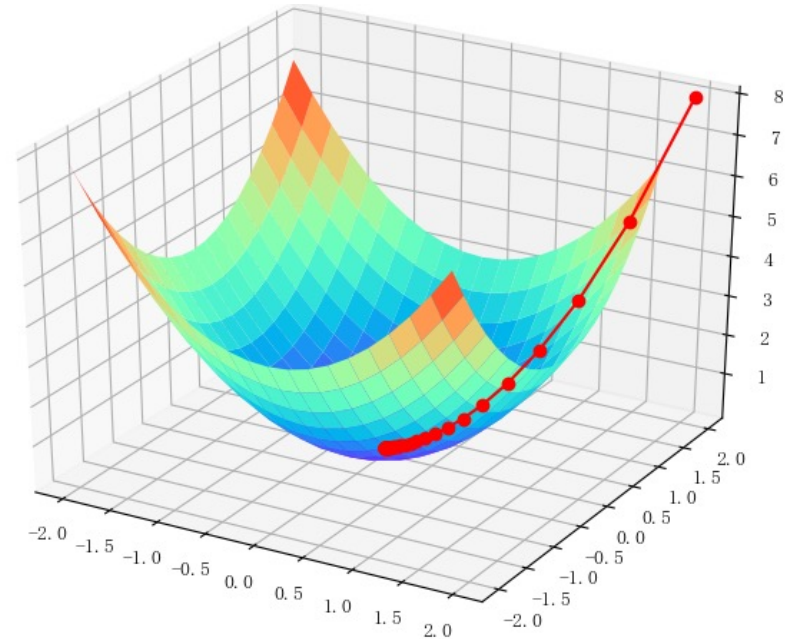
Cada par de parámetros (a, b) tiene asociado un valor de la función pérdida $J(a, b)$ y una recta que ajusta los datos. El valor mínimo de la función de pérdida corresponde a la recta que "mejor ajusta" los datos

Proceso de optimización

- Caso general: Procedimiento iterativo
- **Método del descenso del gradiente**

Casos particulares: Solución directa

- Ej. Regresión lineal



Datos (\mathbf{X} , y)

- Para que un modelo puede ser entrenado, necesita tuplas de información: datos de entrada y salida

Modelo $h(\mathbf{x})$

- Función utilizada para describir un modelo

Parámetros Θ

- Un modelo posee parámetros, los cuales son ajustados durante el proceso de la optimización (entrenamiento)

Función de pérdida $J(\Theta)$

- Función que se busca minimizar, de manera que el modelo se ajuste a los datos

Optimización

- Técnica utilizada para realizar la optimización

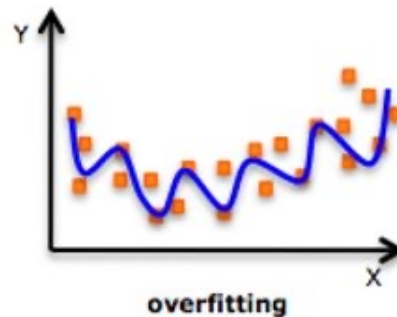
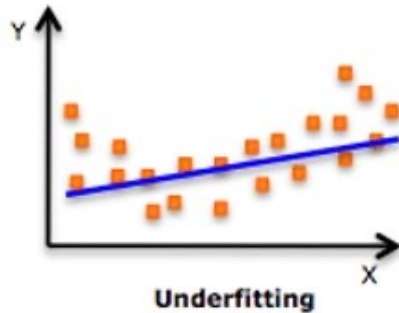
Existen dos etapas diferentes en los modelos ML:

Entrenamiento / Ajuste

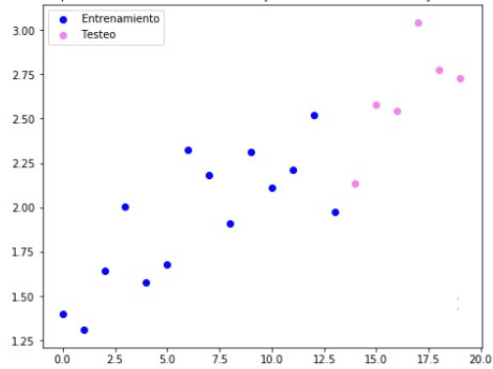
- Momento en que se ajustan los parámetros del modelo en base a un proceso de optimización

Predicción

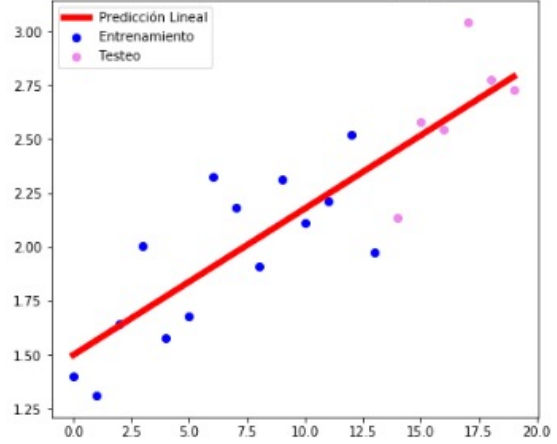
- Momento en que se verifica la respuesta de un modelo (sin modificar sus parámetros)



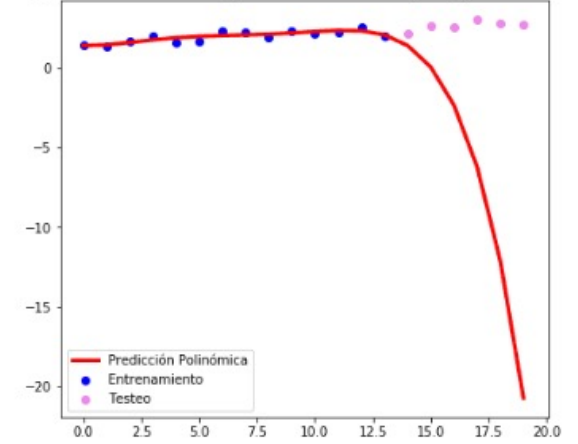
Separación de los datos en los conjuntos de Entrenamiento y Testeo



Regresión Lineal - MSE_entrenamiento: 0.0515, MSE_testeo: 0.0444



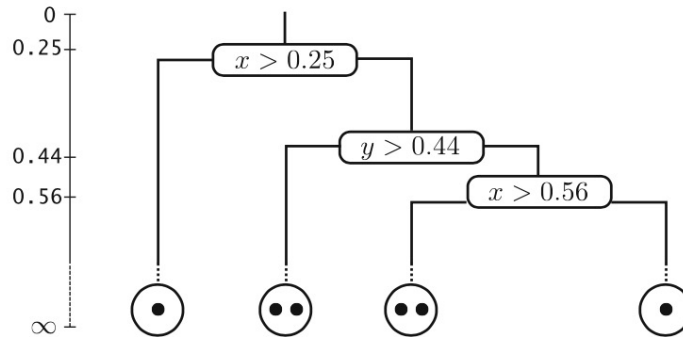
Regresión Polinómica - MSE_entrenamiento: 0.0367, MSE_testeo: 148.293



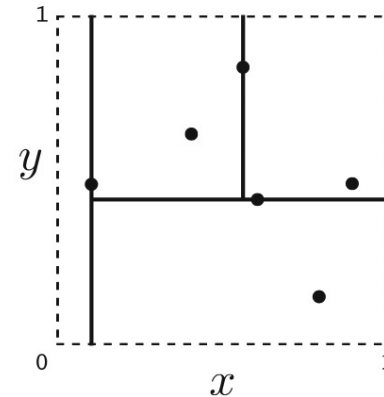
- **¿Cómo se propaga la predicción al modelo de bloques?**
 - Necesitamos las variables “de entrada” definidas o medidas en el dominio de predicción
 - Como alternativa, generar co-variables en el dominio:
 - Distancias al contacto
 - Modelos geológicos
 - Variables geofísicas
 - etc

- Árboles aleatorios de particiones del dominio, tanto de las muestras como del dominio a estimar.
- Se procesan con una “función de agregación”

$$\text{MF}(\Theta, \lambda | \mathbf{m}) = \{T_1, \dots, T_m\}, T_k \sim \text{MT}(\Theta, \lambda)$$

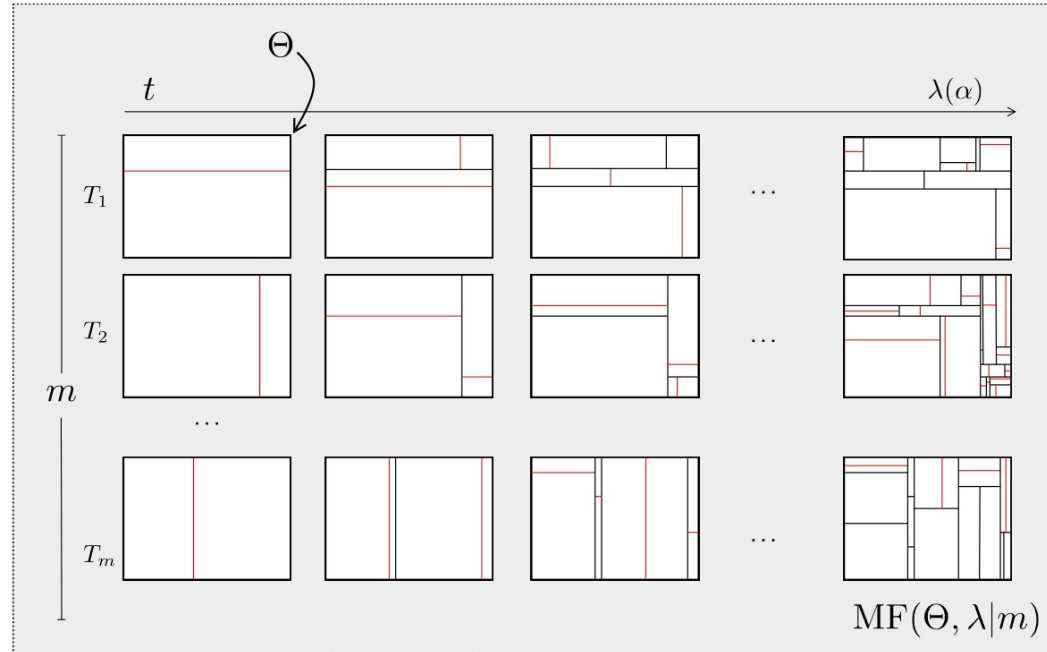


(a) Mondrian tree

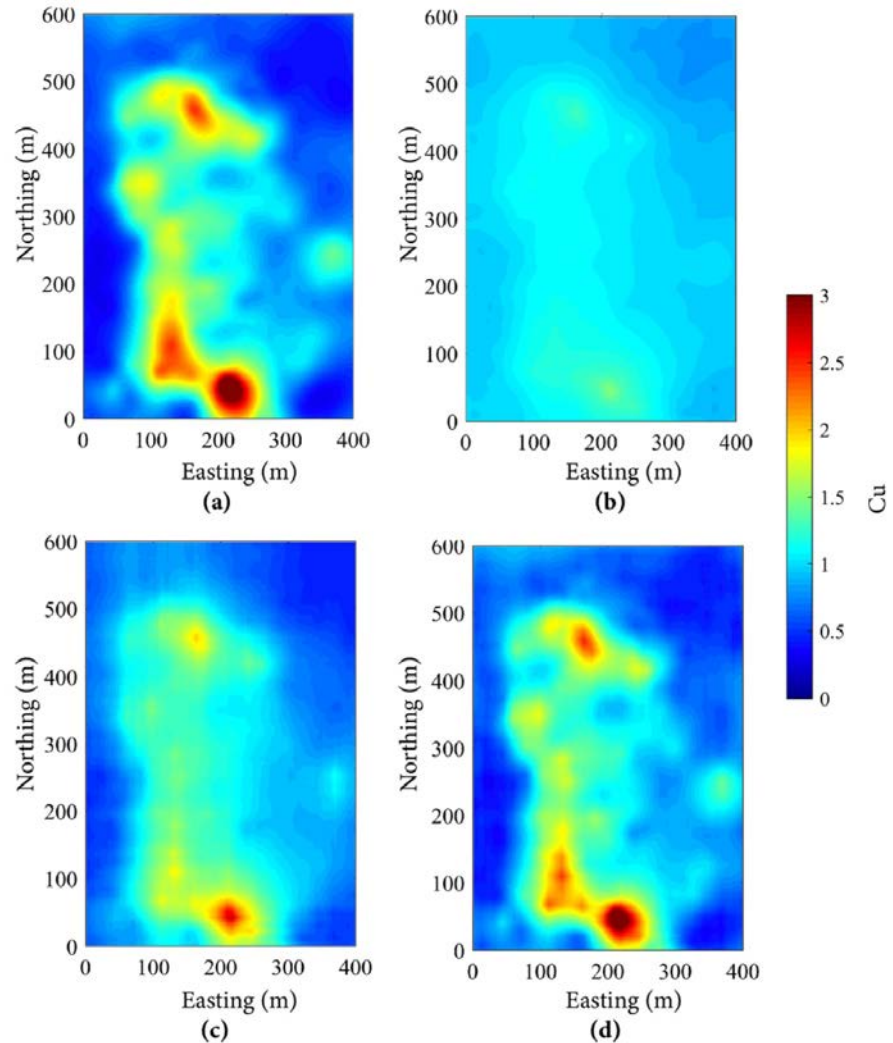


(b) Conditioning data

- Proceso generativo para dibujar un conjunto de partición estocástica dependiente del tiempo de tamaño m . El tiempo va a lo largo del eje x , terminando en el tiempo dado por ρ . Las líneas rojas indican el corte de la partición temporal actual.



- a) Kriging Global Ordinario (GOK)
- b) Inverso de la distancia (IDW)
- c) ESI con IDW
- d) ESI con GOK



- Los métodos de ML son una alternativa válida en presencia de datos abundantes
- Es necesario escoger el mejor modelo para el problema a resolver
- En el futuro, los métodos computacionales reducirán el costo de resolver “problemas pequeños”:

- Los métodos de ML son una alternativa válida en presencia de datos abundantes
- Es necesario escoger el mejor modelo para el problema a resolver
- En el futuro, los métodos computacionales reducirán el costo de resolver “problemas pequeños”:
- ¿Estamos camino a una estimación **sin variograma**?



Comisión Calificadora de
Competencias en Recursos
y Reservas Mineras



Ingeniería de Minas
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Uso de machine learning en la estimación de recursos minerales como alternativa a kriging

Felipe Navarro

fnavarro@alges.cl