

CLASES DESBALANCEADAS Y EXPLICACIÓN DE MODELOS

Ph.D. Santiago Echeverri Arteaga



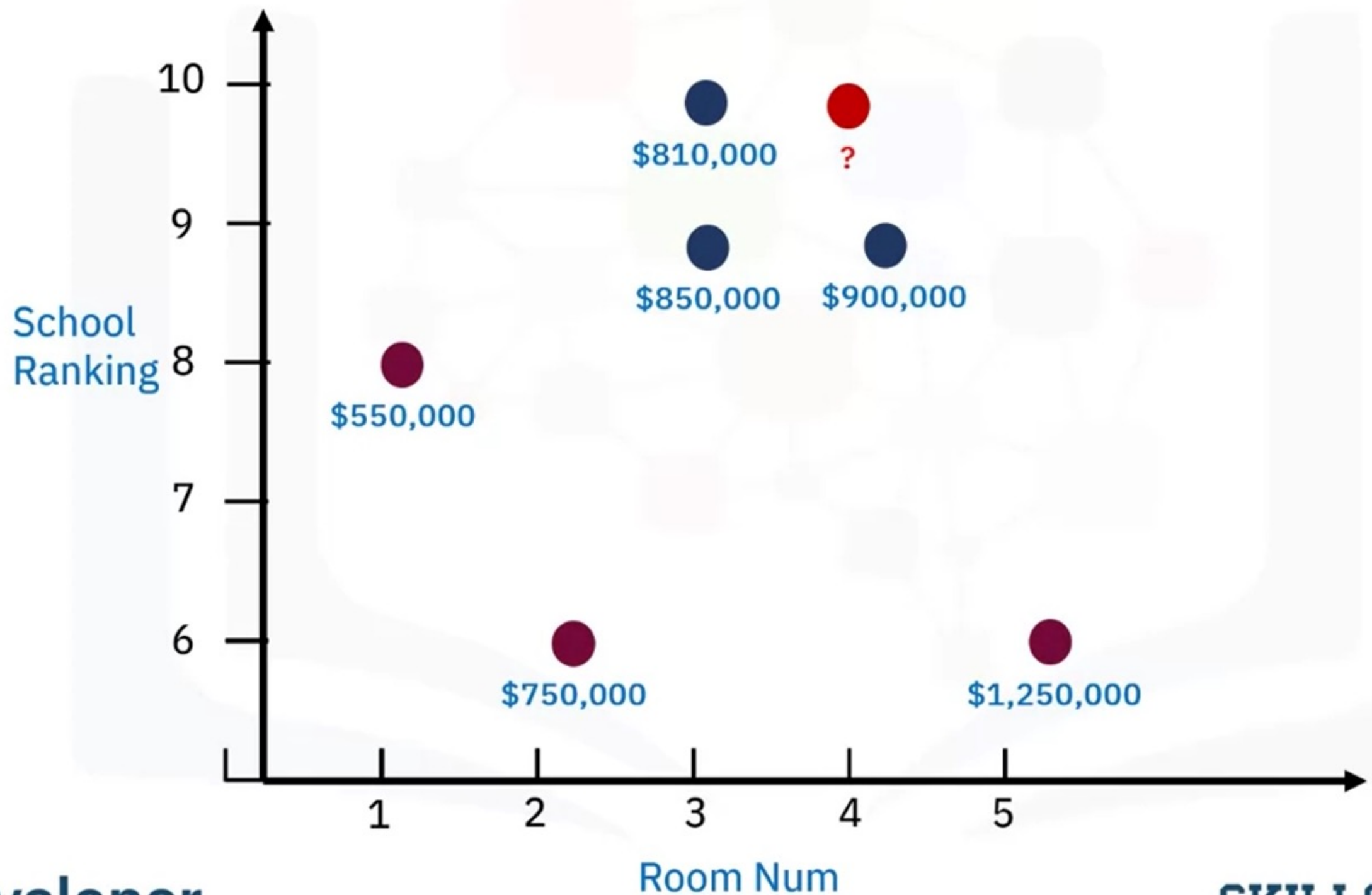
TIPOS DE MODELOS EN MACHINE LEARNING

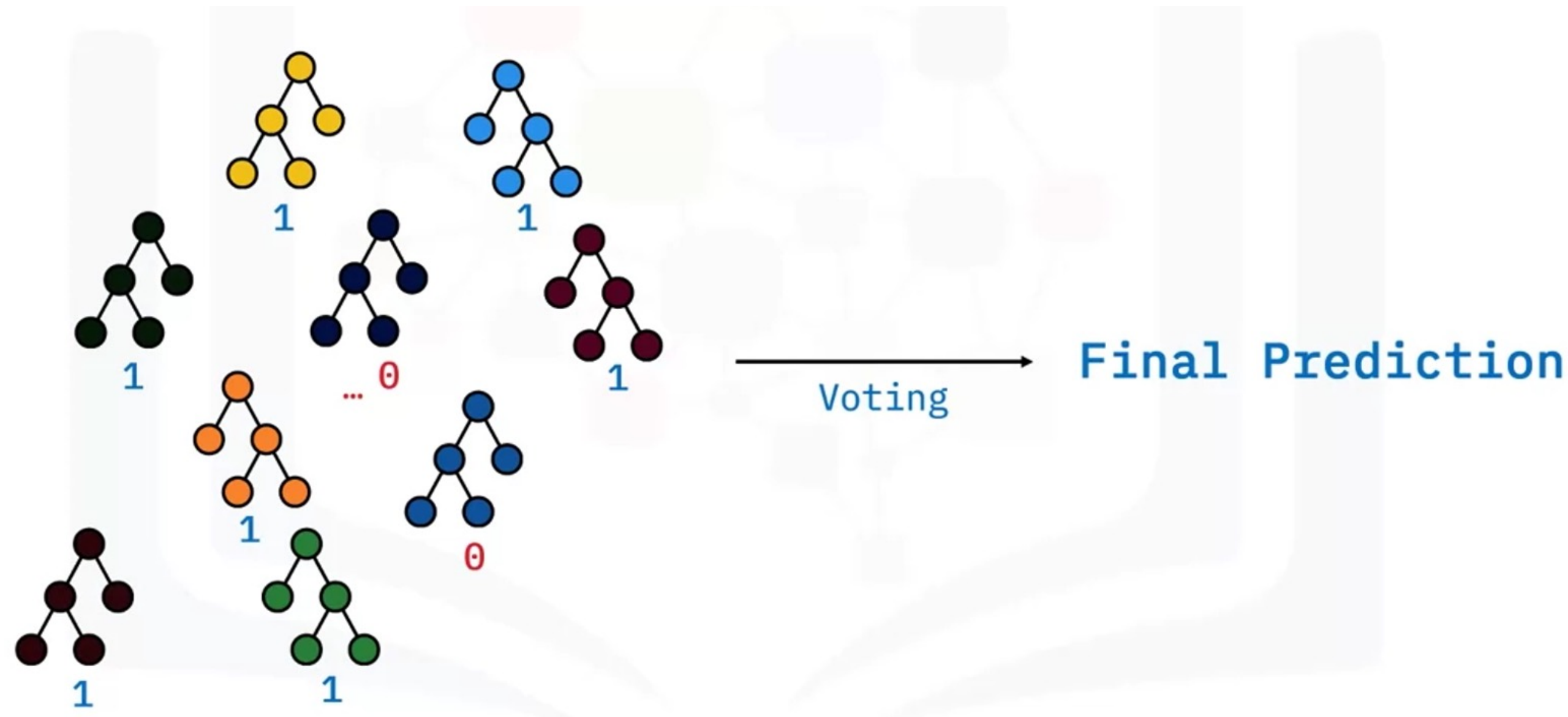
Modelos auto-interpretables

- Permiten ver la estructura del modelo, la influencia de las variables independientes y cómo hace el modelo predicciones
- No necesita métodos adicionales de explicación por su estructura sencilla

Modelos tipo caja negra

- No permite ver cómo funciona, por qué falla ni se deja monitorear
- En algunos casos son modelos con mejor rendimiento
- Procesamiento de lenguaje natural, reconocimiento de imágenes, patrones de tráfico





MÉTODOS DE INTERPRETACIÓN DE MODELOS

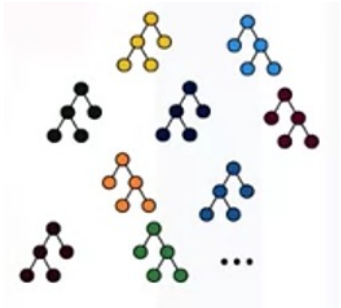
Intrínsecos

- Modelos con alta interpretabilidad
- Simplificar modelos para ganar interpretación (Regularización o poda)



Post-Hoc

- Modelos con baja interpretabilidad como máquinas de soporte vectorial con kernels no lineales, métodos de ensamble y redes neuronales
- Exigen métodos auxiliares



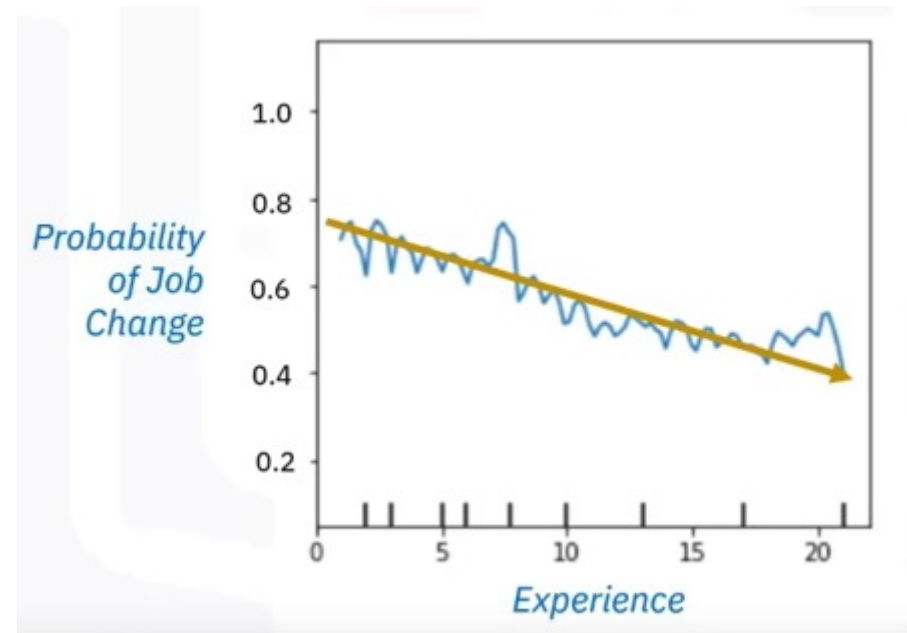
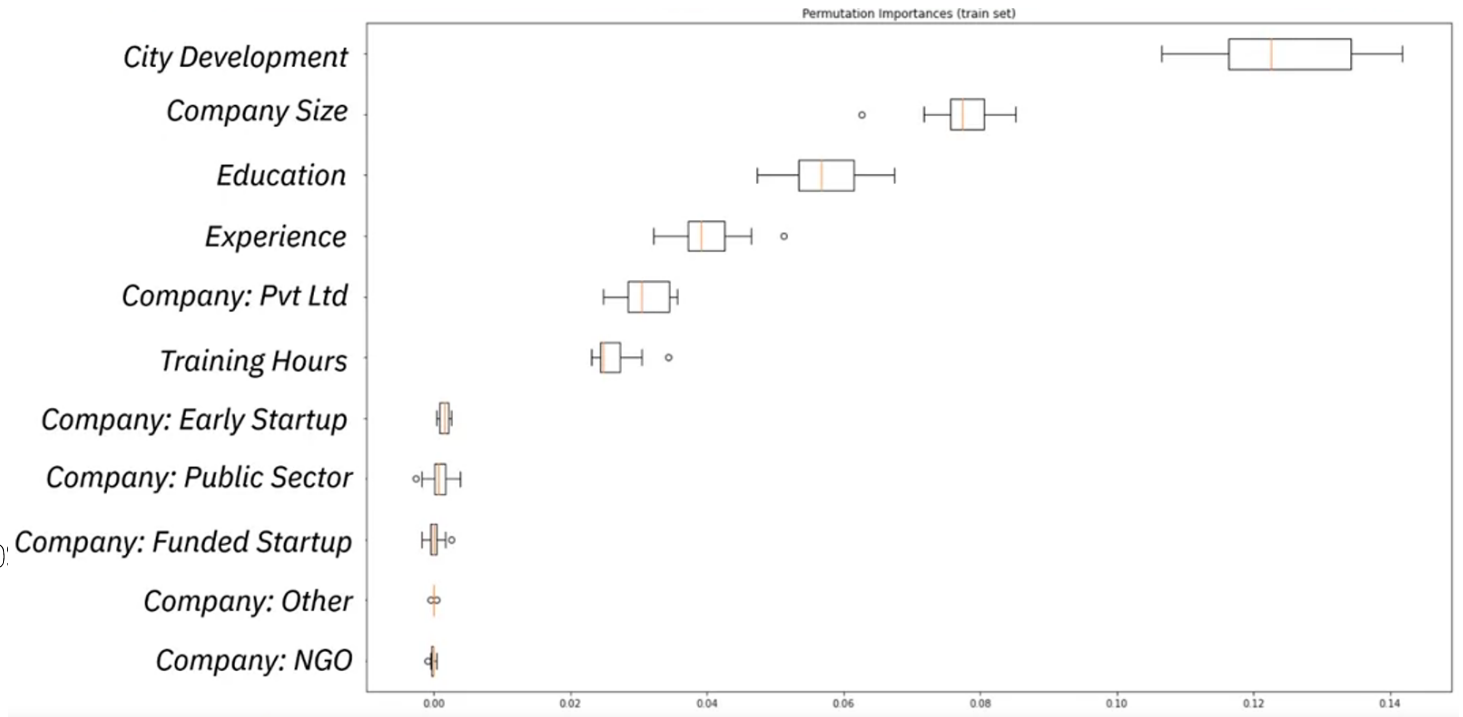
EXPLICACIÓN AGNÓSTICA DE MODELOS

- Métodos unificados para interpretar de forma unificada diferentes modelos: sin tener que acudir a diferentes métodos de explicación para cada uno

- IMPORTANCIA DE CARACTERÍSTICAS

- Importancia de características por permutación
- Importancia de características por impureza
- Explicaciones aditivas de Shapley (SHAP)
- ...

- GRÁFICA DE DEPENDENCIA PARCIAL



TOY MODEL

O también conocidos como modelos sustitutos

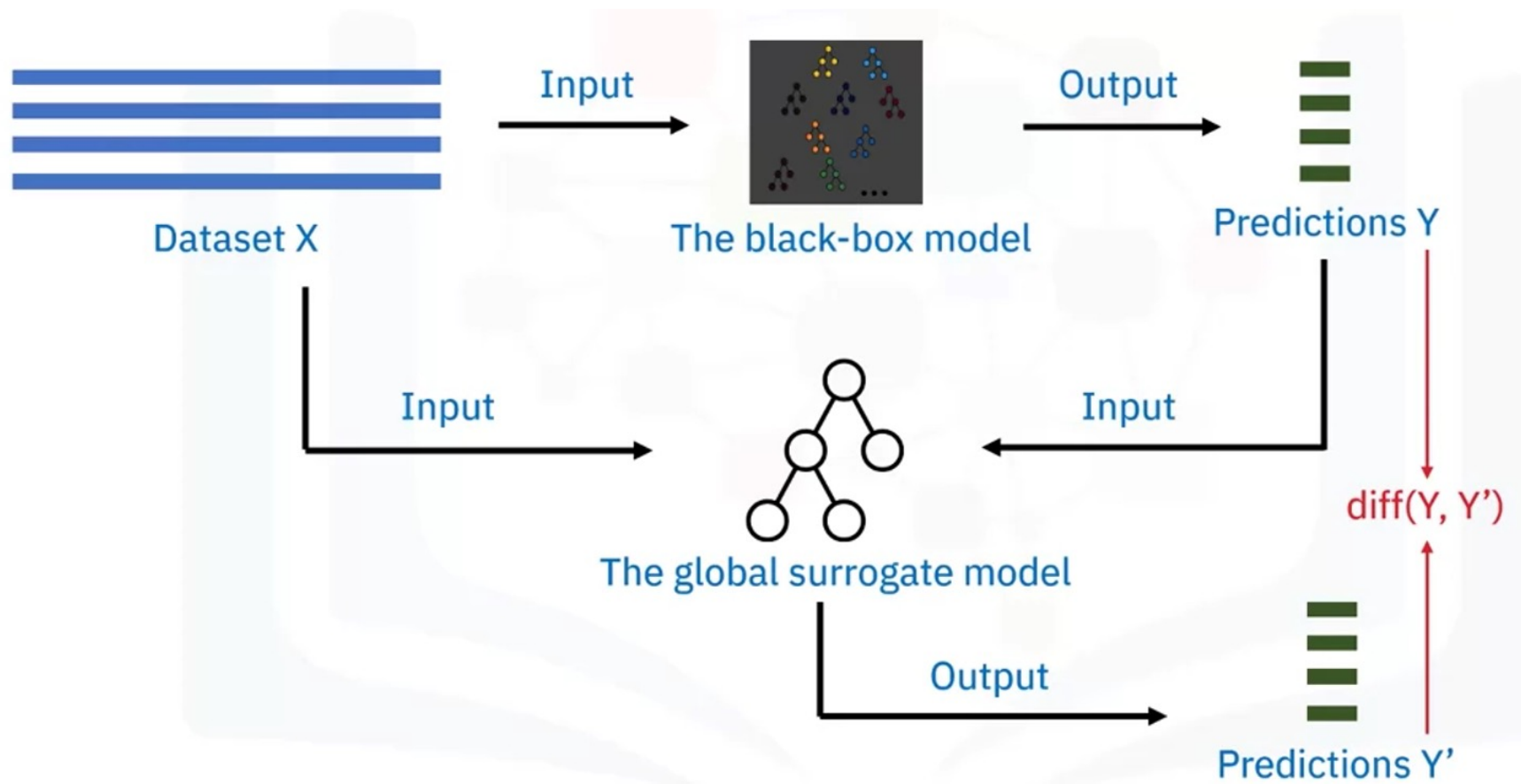


MODELOS SUSTITUTOS

Si se tiene un modelo muy complejo para explicar se puede intentar crear un modelo a partir de los inputs y outputs del modelo original. Si las predicciones del modelo sustituto son similares, se asume que la INTERPRETACIÓN es la del sustituto

Modelo sustituto globales vs Modelo sustituto local

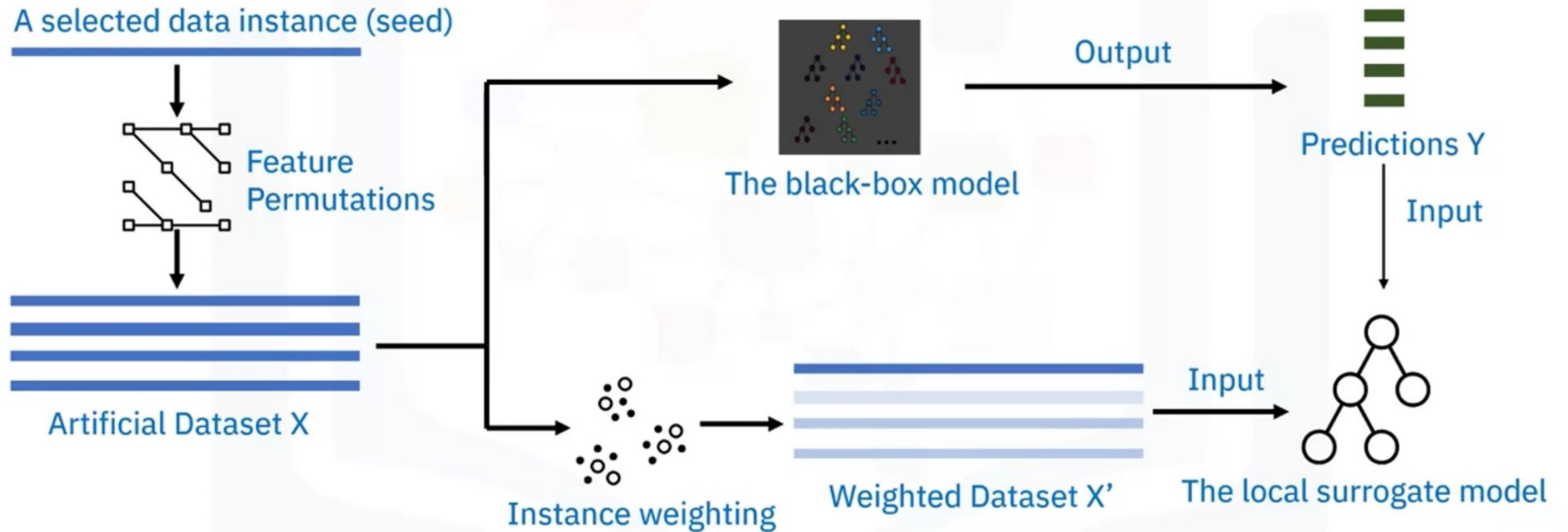
MODELOS SUSTITUTOS GLOBALES

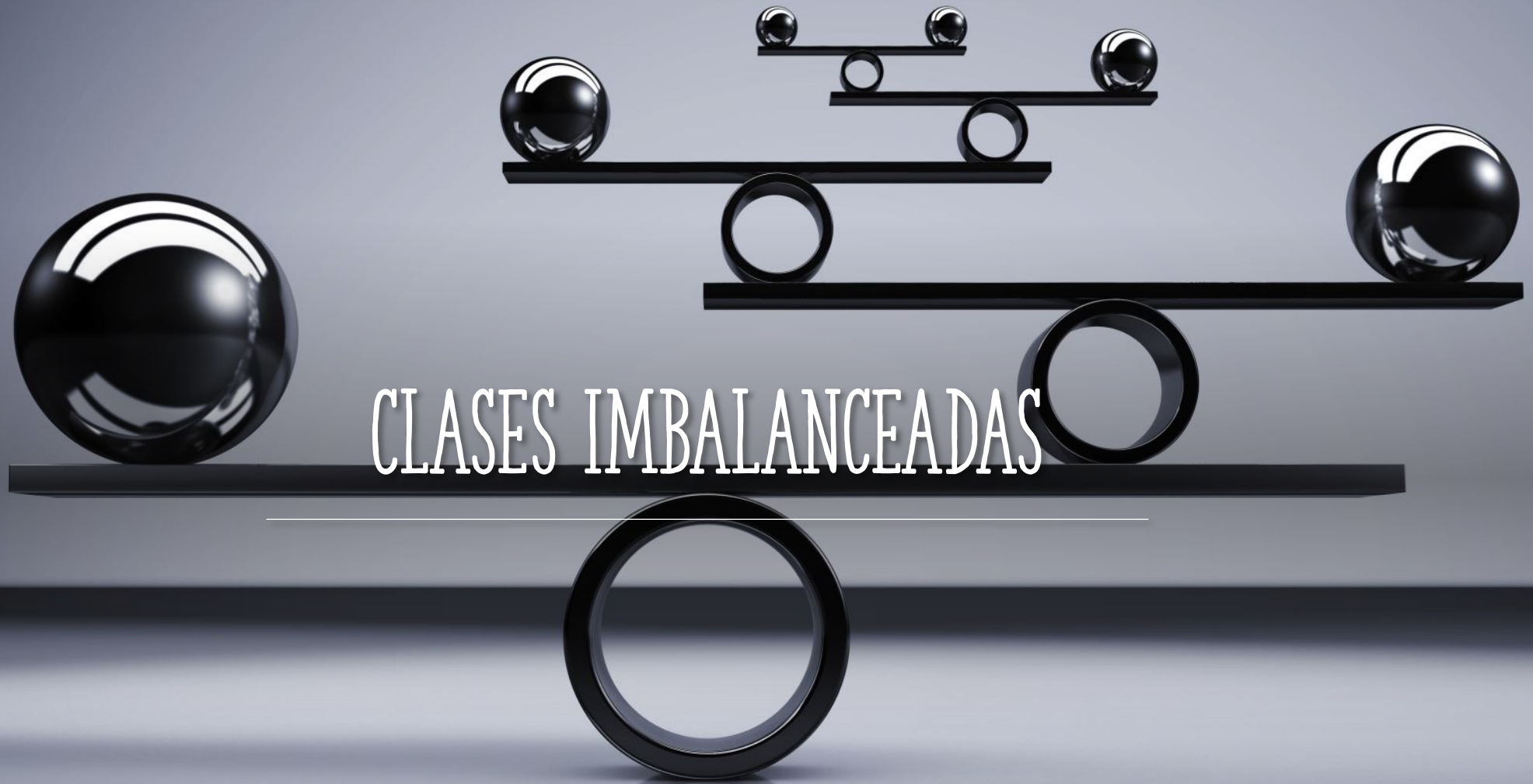


EXPLICACIONES LOCALES INTERPRETABLES DE MODELOS AGNÓSTICOS (LIME)

1. Seleccionar datos (semilla)
2. Construir un set de datos artificial lo suficientemente grande como para entrenar un modelo
3. Se pesa cada observación con base en que tan cerca está a la semilla y a partir de ahí se determina nuestro \mathbf{x}'
4. Se introduce el conjunto de datos generados al modelo caja negra para hacer las predicciones \mathbf{y}
5. Con \mathbf{x}' y \mathbf{y} se entrena un modelo sustituto local

MODELOS SUSTITUTOS LOCALES

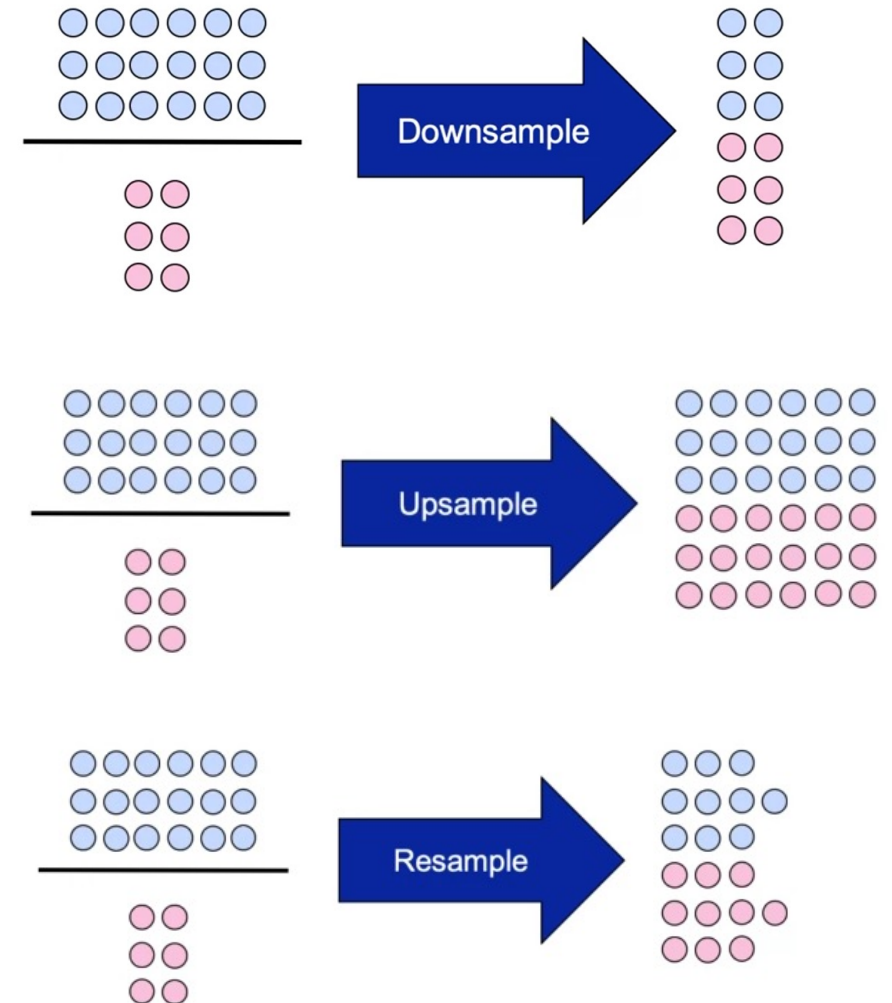




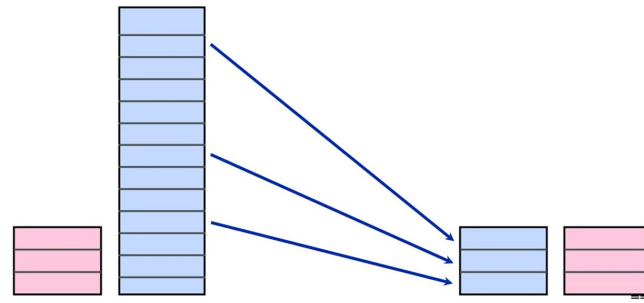
CLASES IMBALANCEADAS

COMO ABORDAR CLASES IMBALANCEADAS

- Los métodos están diseñados para optimizar la exactitud (accuracy)
- **DOWNSAMPLING**: Tomar tantos datos de la clase mayoritaria como haya de la clase minoritaria
- **UPSAMPLING**: Crear copias de las filas minoritarias hasta que se tenga una muestra equilibrada
- **REMUESTREO**: Downsampling + Upsampling hasta llegar a un punto medio
- SIEMPRE se debe hacer una división del set de datos ESTRATIFICADA
- Suele implicar "sacrificios" en una de las clases al aumentar las predicciones de las clases minoritarias positivas (Aumento de Falsos Positivos *Aumento de hombres embarazados*). Por lo que se suele aumentar la sensibilidad y disminuir la precisión



DOWNSAMPLING



- Aumenta la importancia de las clases minoritarias. Sensibilidad mucho mayor que precisión.

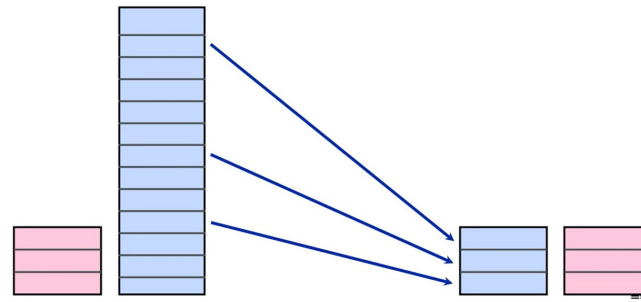
Near miss 1

Near miss 3

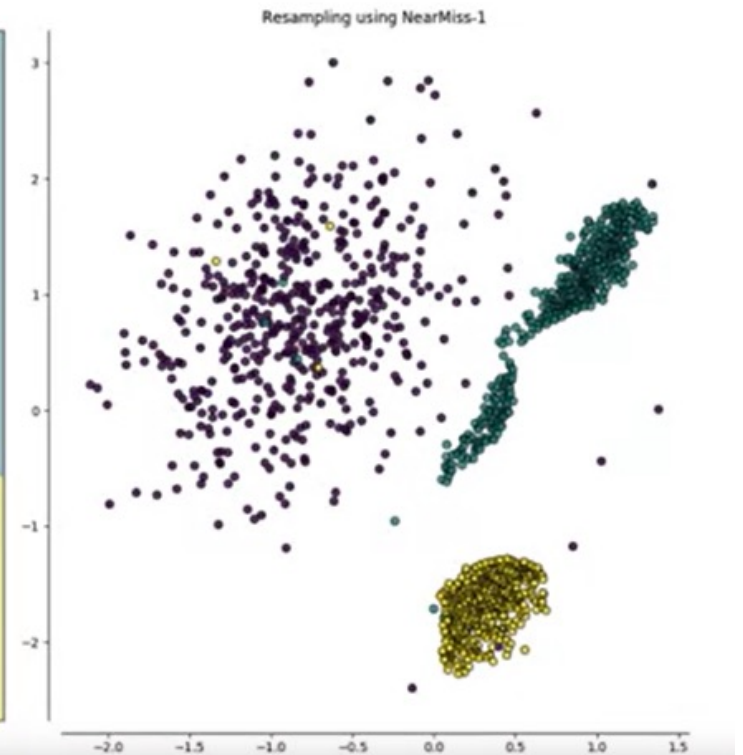
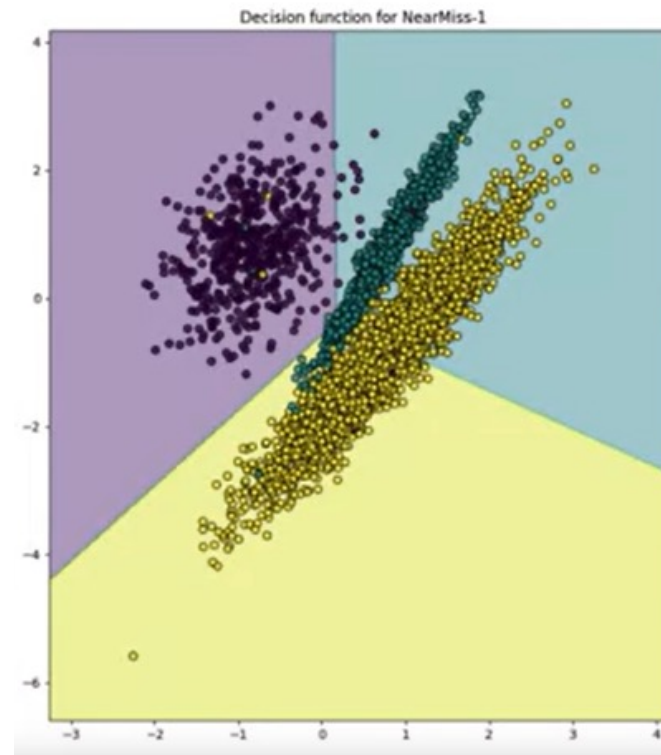
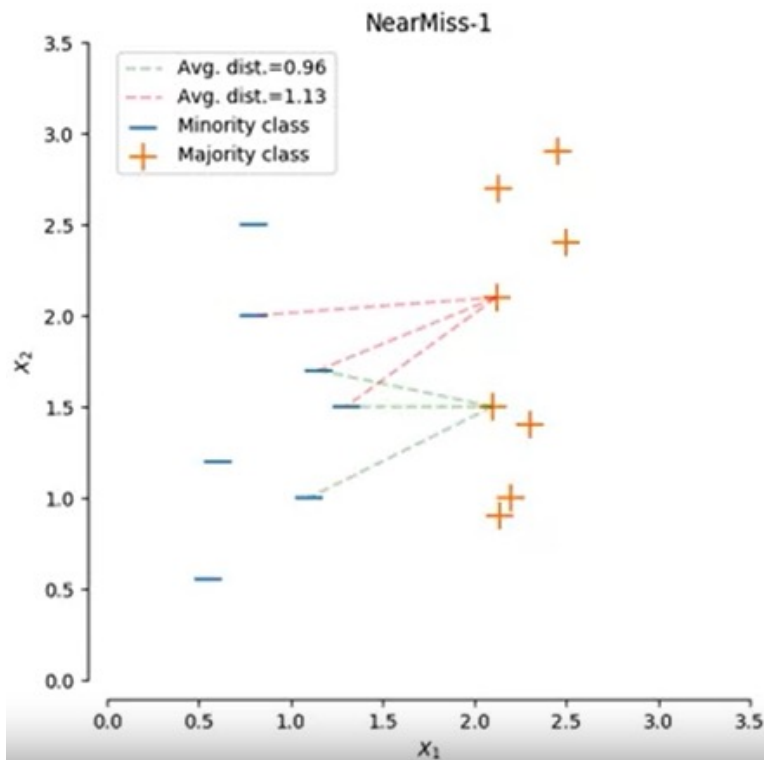
Near miss 2

Links de Tomek

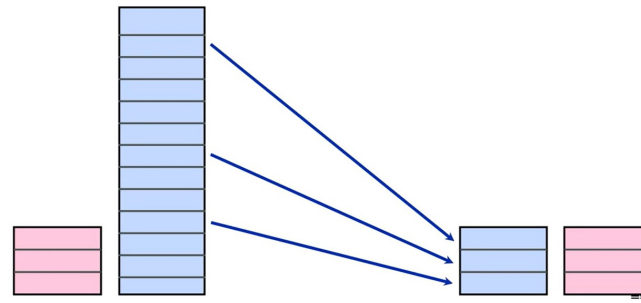
NEAR MISS 1



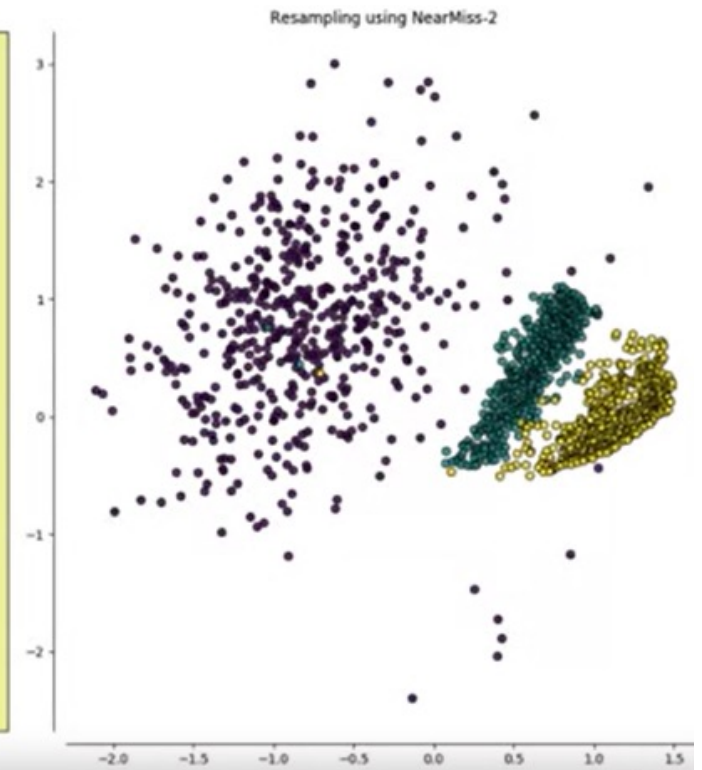
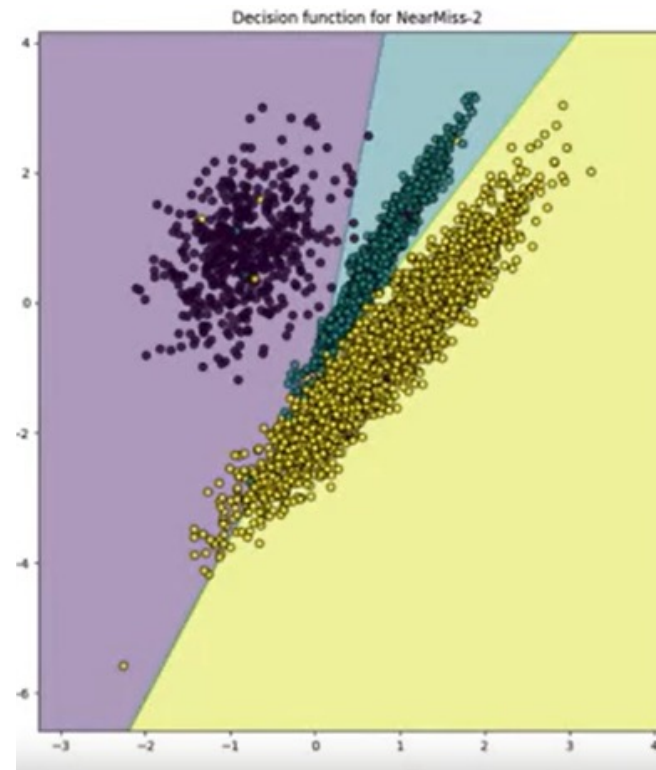
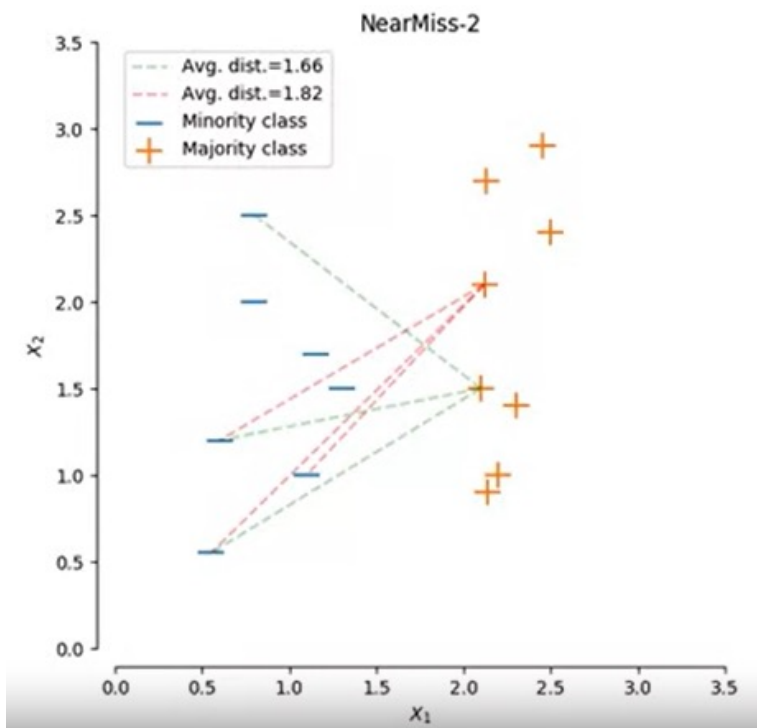
- Seleccionar los puntos más cercanos a la clase minoritaria (posible sesgo por outliers)



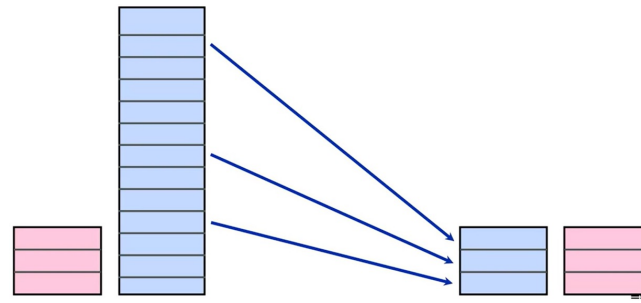
NEAR MISS 2



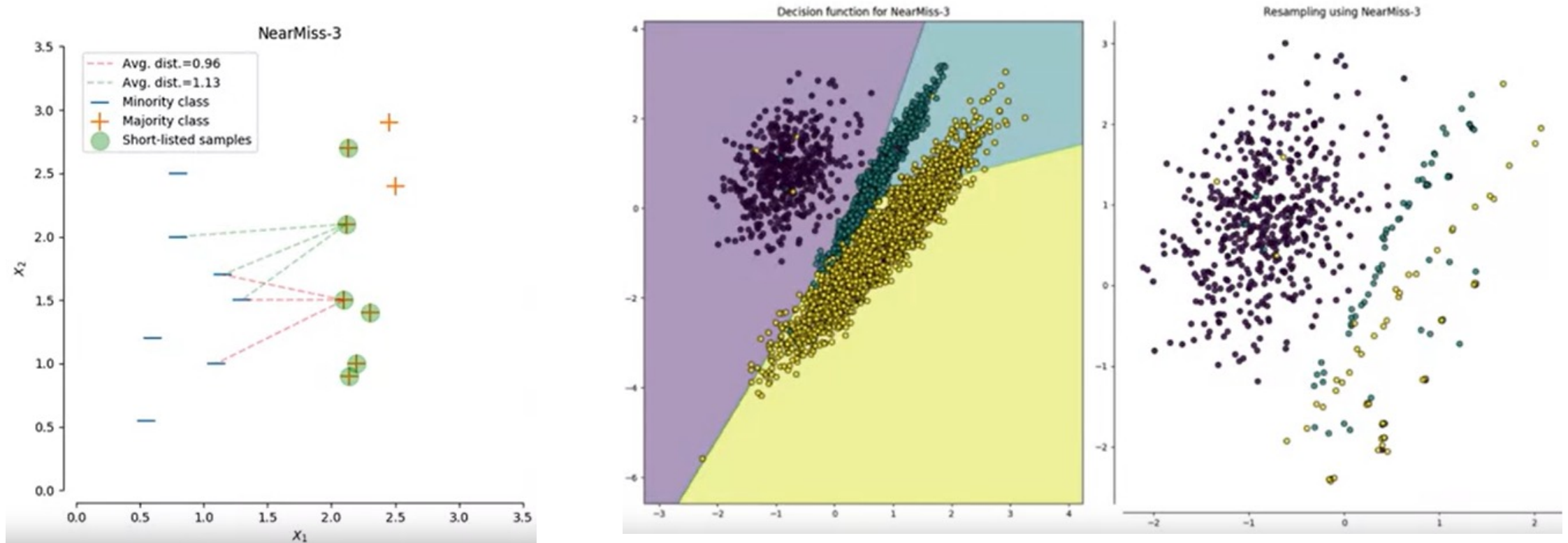
Selecciona valores de la clase mayoritaria que tienen menor distancia a los puntos más alejados de la clase minoritaria (en promedio)
Menos sensible a outliers pero se puede afectar por outliers marginales



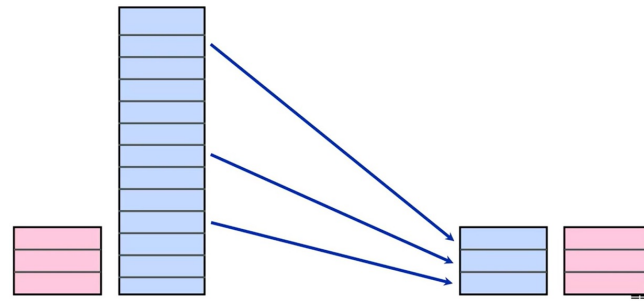
NEAR MISS 3



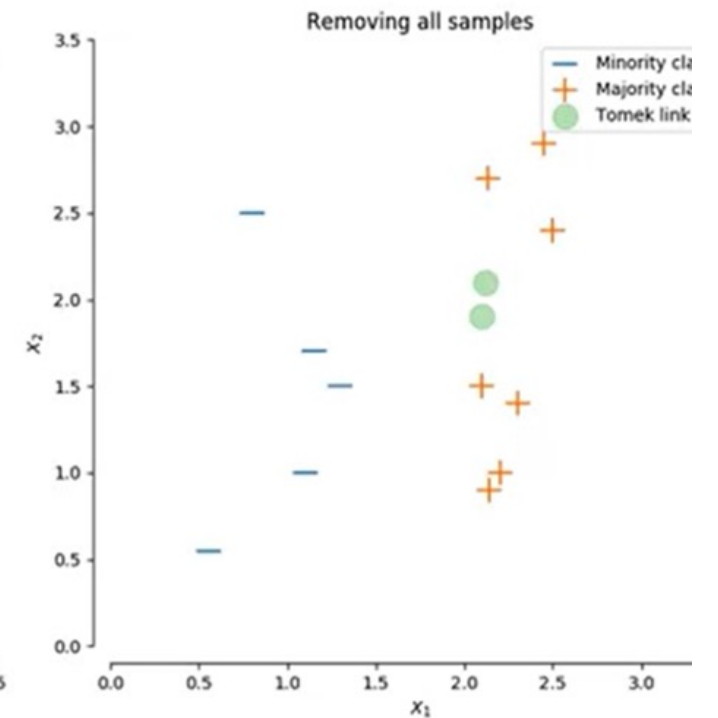
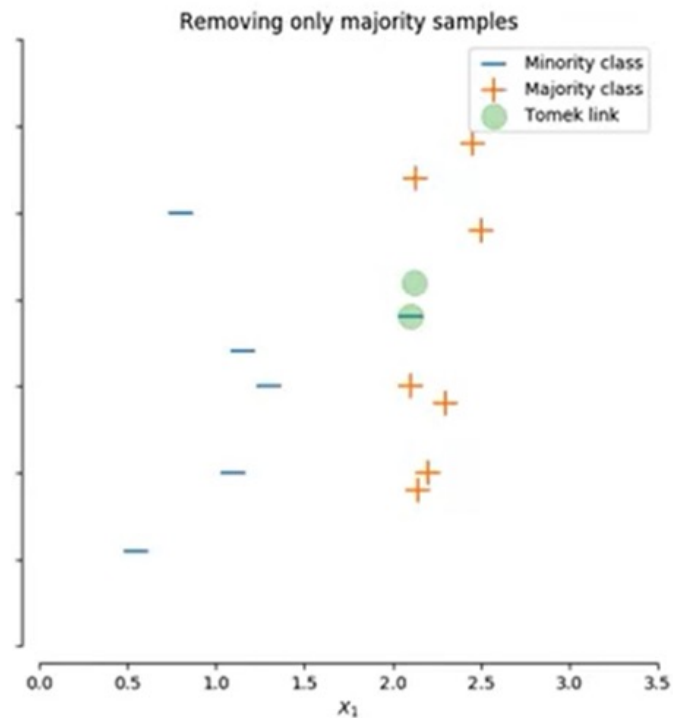
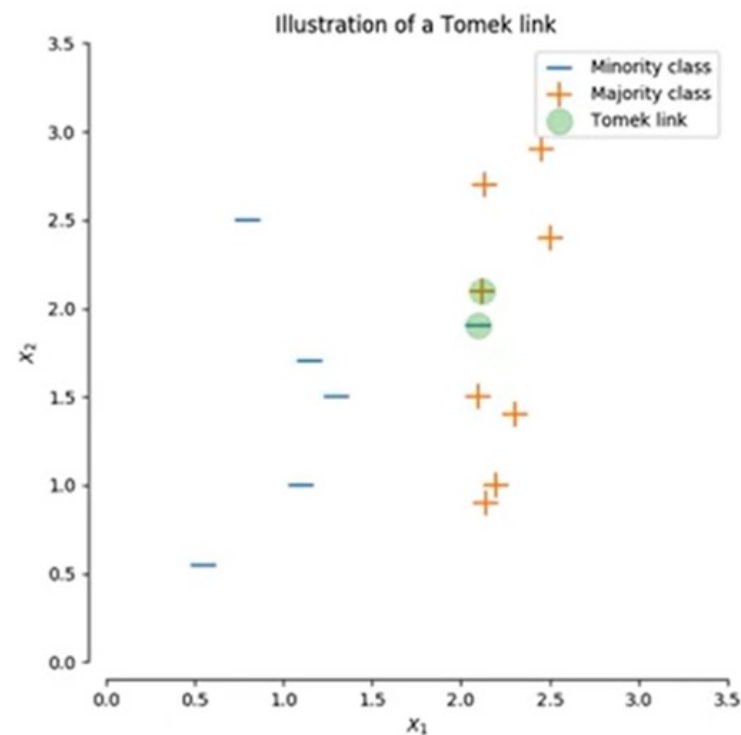
Para cada minoritario se encuentran los KNN de la clase mayoritaria. Se seleccionan las muestras mayoritarias para las que la distancia promedio a los KNN es la más grande. Menos afectada por el ruido y por outliers por tomar puntos distantes entre sí



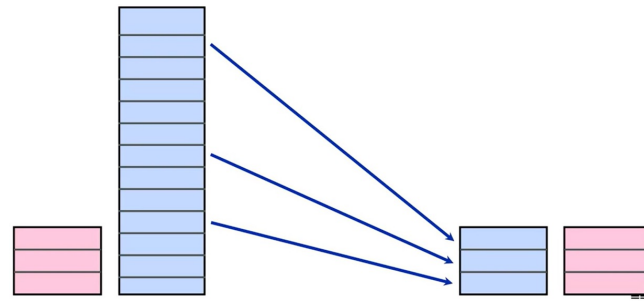
LINK DE TOMEK



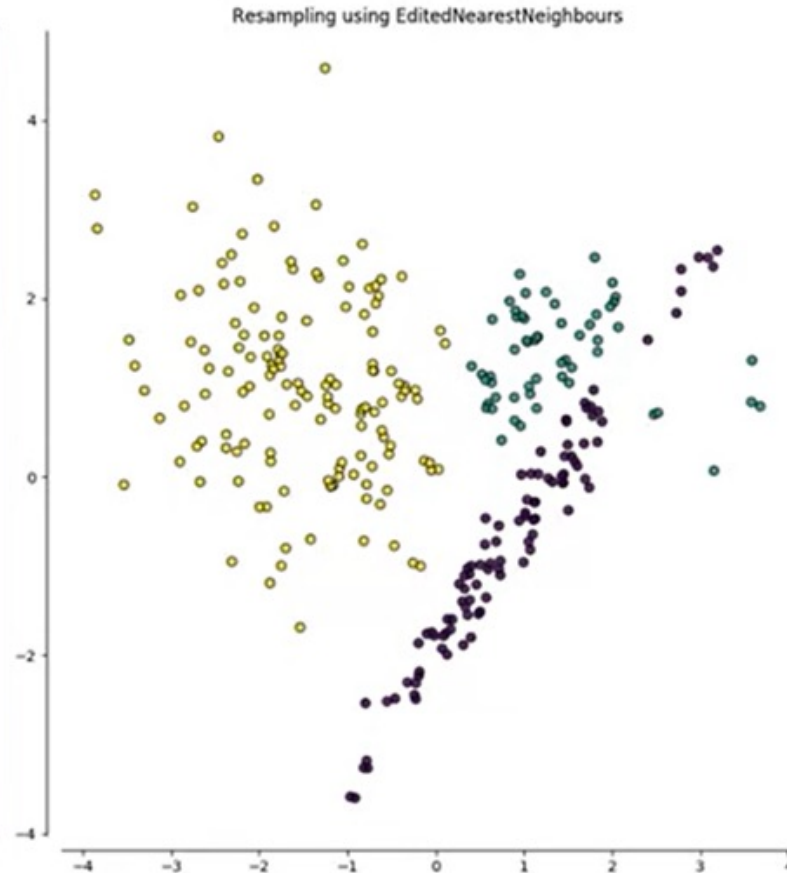
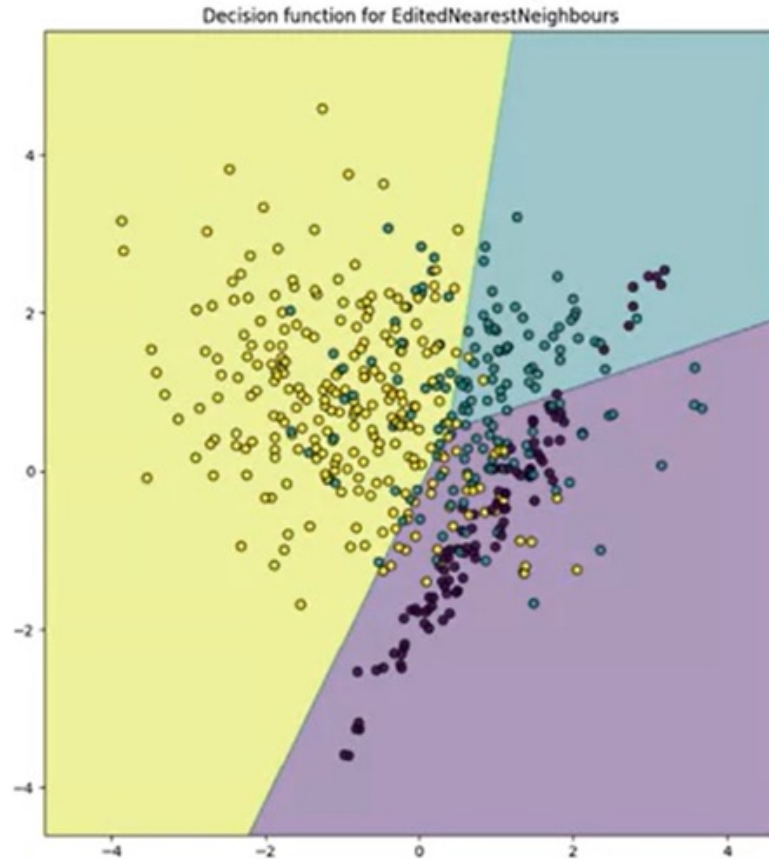
Un link de Tomek existe si dos vecinos de clases diferentes son los KNN.



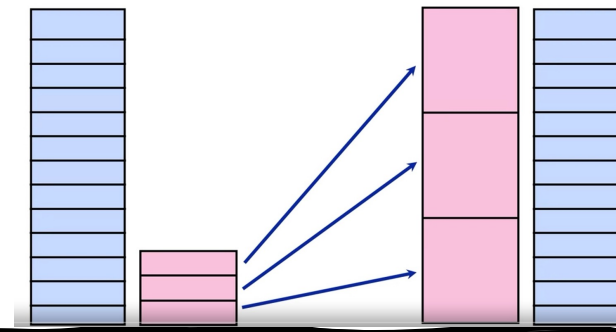
KNN MODIFICADO



KNN con $K=1$ para remover los datos mal clasificados



UPSAMPLING



- Mitiga el peso excesivo de la clase mayoritaria. El desbalance sensibilidad/precisión es menor que en downsampling
- Se puede usar el kwarg class weight → "balanced" que lo que hace es balancear el error entre las clases mayoritaria y minoritaria
- Sobremuestreo aleatorio: Remuestreo con remplazo de las filas de la clase minoritaria. (Bueno para datos categóricos)
- Sobremuestreo sintético: Crear nuevos datos que no existían.
 - Seleccionar una observación
 - Elegir uno de los k-vecinos cercanos
 - Poner un punto nuevo entre ellos
 - Repetir para los vecinos cercanos



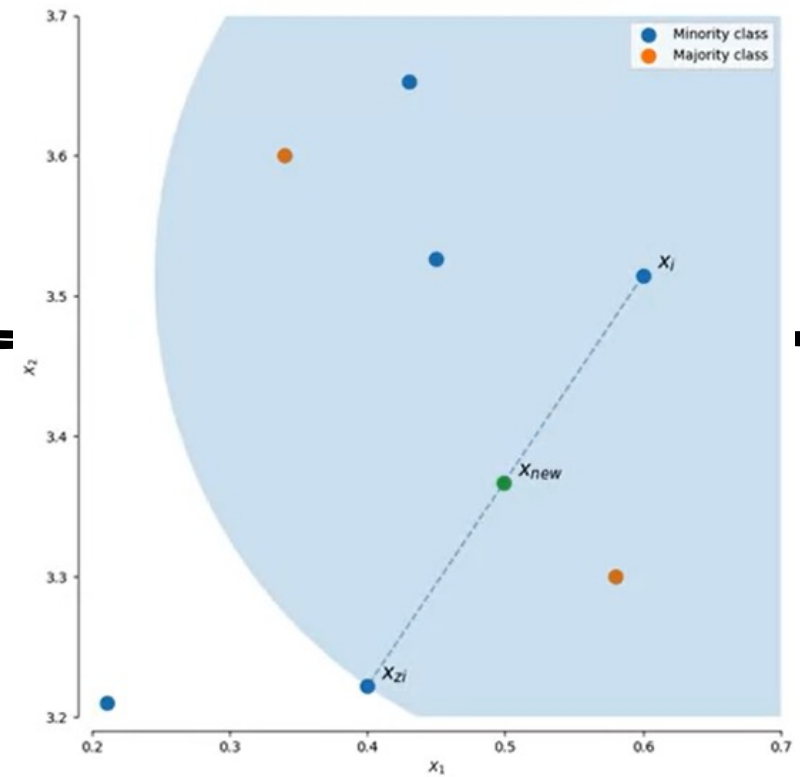
Técnica de sobremuestreo sintético minoritario SMOTE

Muestreo sintético adaptativo (ADASYN)

UPSAMPLING

SMOTE

- Regular: Conexión entre los puntos de la clase minoritaria con cualquier vecino (Incluso otras clases)
- Borderline: Analiza si el punto minoritario es **outlier**, **seguro** o en **peligro**
 - 1: Conecta en **peligro** solo con clase minoritaria.
 - 2: Conecta en **peligro** con todos.
- SVM: Usa máquina de soporte vectorial para generar nuevos puntos.



ADASYN

- Mira las clases alrededor del punto minoritario. El número de muestras generadas por cada punto es proporcional al número de datos que no son de la misma clase en el vecindario

Da mas peso a los mal clasificados (motivado por KNN)

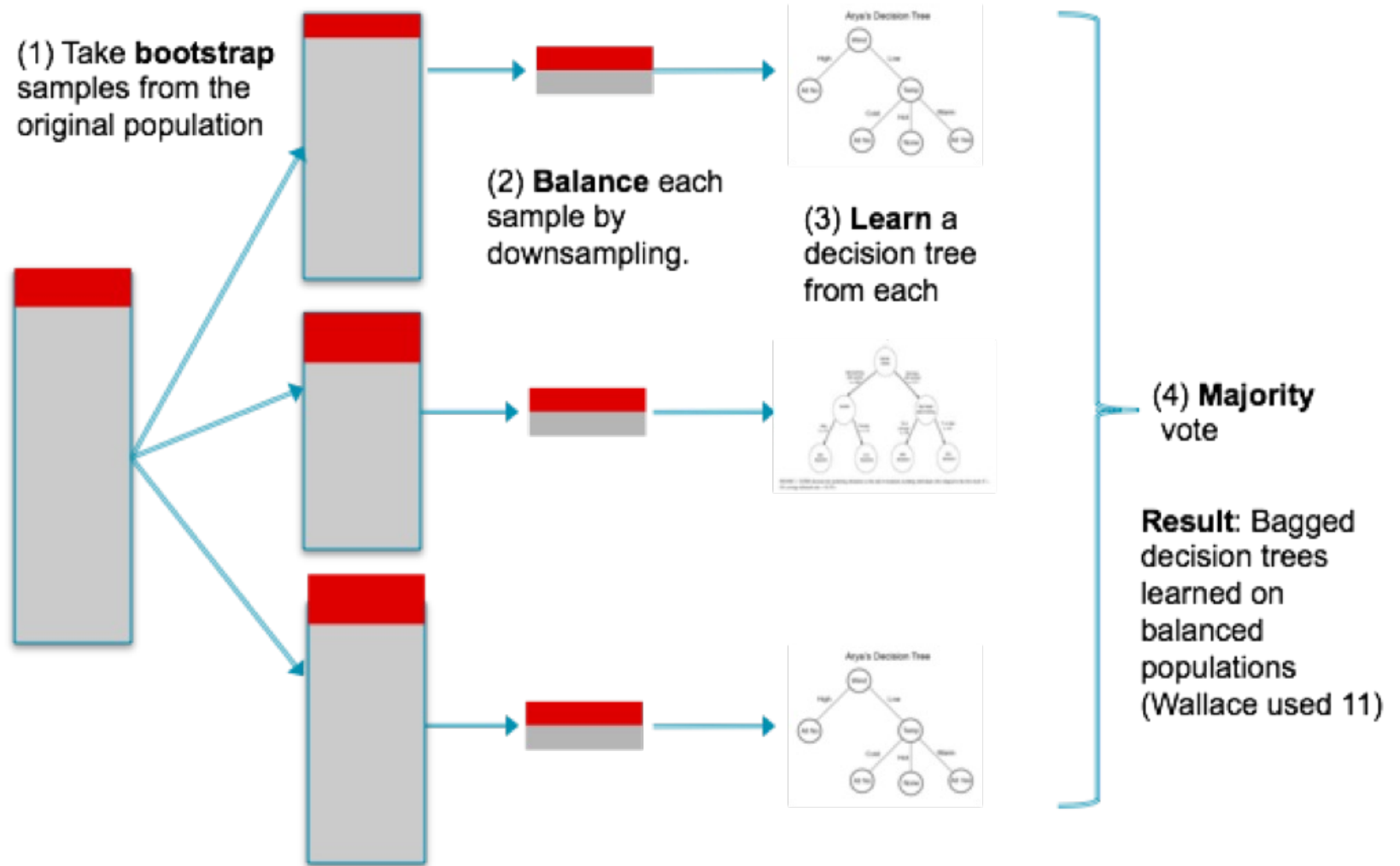
REMUESTREO

SMOTE + Links de
Tomek

SMOTE + KNN
modificado

Blagging

BLAGGING: BALANCED BAGGING



A TENER EN CUENTA

- Hacer division entrenamiento/prueba estratificada
- Usar métricas sensibles
- NO usar accuracy
- Undersampling, oversampling o remuestreo

