

Ensembles Híbridos

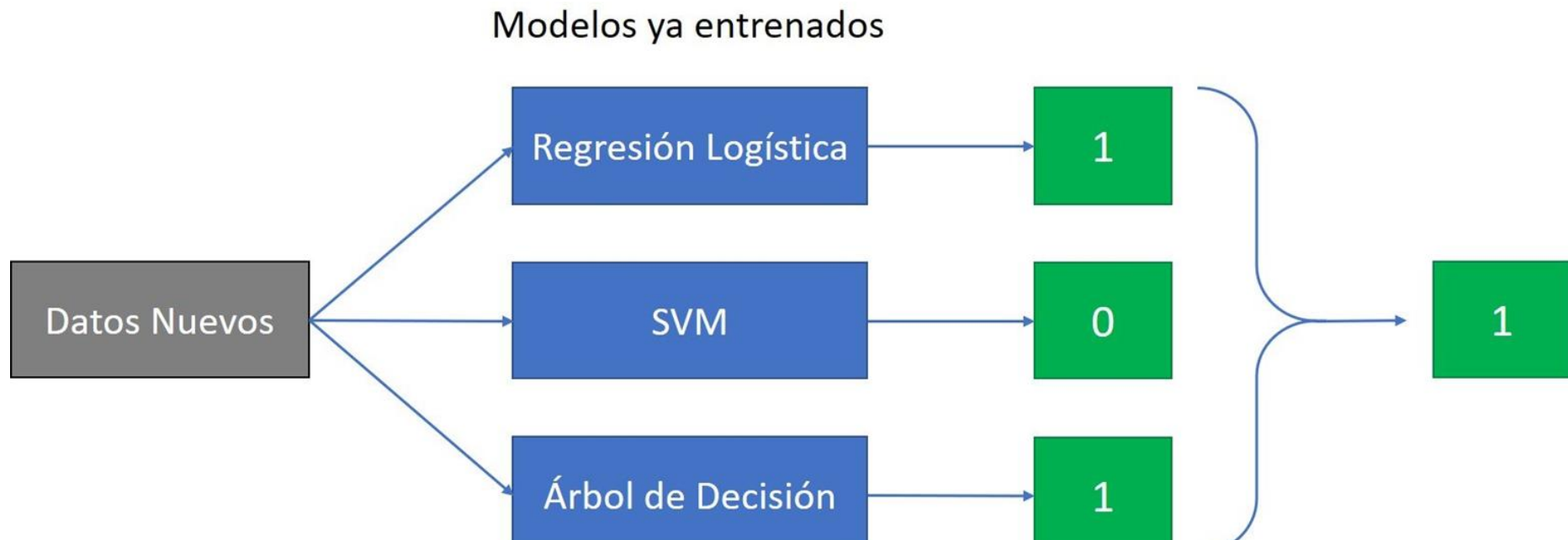
Voting, Stacking, Cascading

Ensamblas híbridos

- Ensamblas homogéneos: combinan el mismo tipo de modelo
 - Bagging
 - Boosting
- Ensamblas híbridos: combinan clasificadores de distinto tipo
 - Voting
 - Stacking
 - Cascading

Voting

Construir N modelos utilizando los mismos datos y luego tomar la predicción mayoritaria

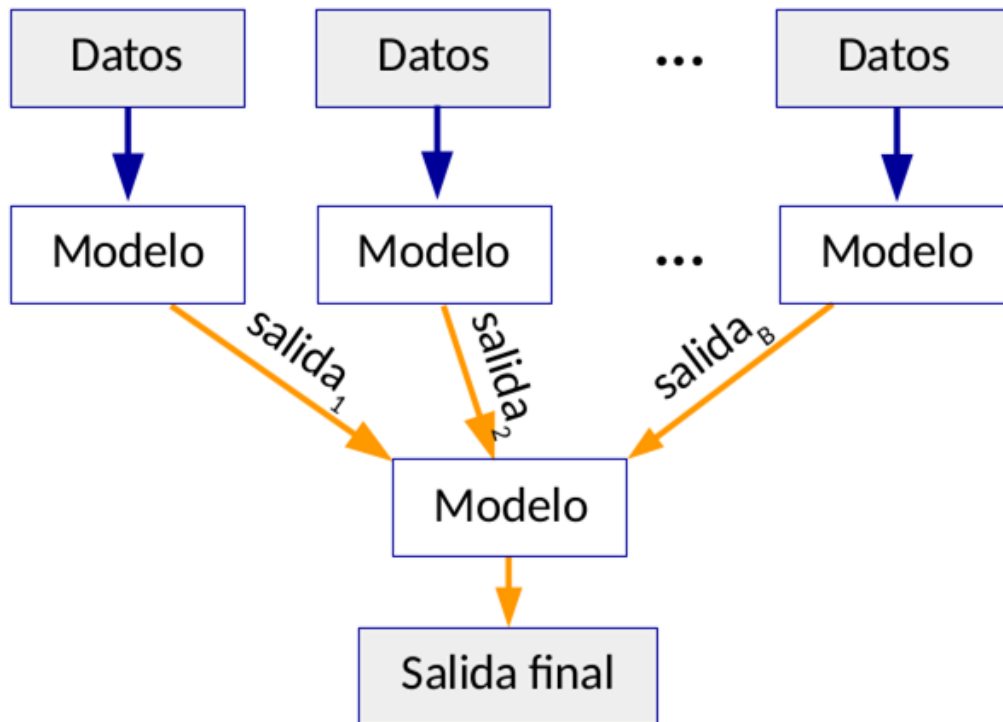


Stacking

- Entrenar diferentes modelos (modelos base) y un modelo más, que decide, dada una instancia nueva, qué modelo usar.
- Concepto de meta-aprendizaje para reemplazar el mecanismo de voto
- Pueden apilarse tantas capas de modelos como se desee

Clasificación de una nueva instancia

Nueva instancia



Entrenamiento de B modelos distintos

Stacking características

- Los modelos para el meta-aprendizaje suelen ser: árboles, NB, SVM o Perceptrón (lo vemos en Redes Neuronales)
- Para los otros puede usarse cualquiera
- Menos popular que boosting, bagging
 - Dificultad de análisis teórico: caja negra
 - Múltiples variantes

Stacking características

- Se puede interpretar como una mejora (generalización) del método de votación (Voting)
- Si los clasificadores base pueden generar medidas de certeza, suele funcionar mejor

Cascading

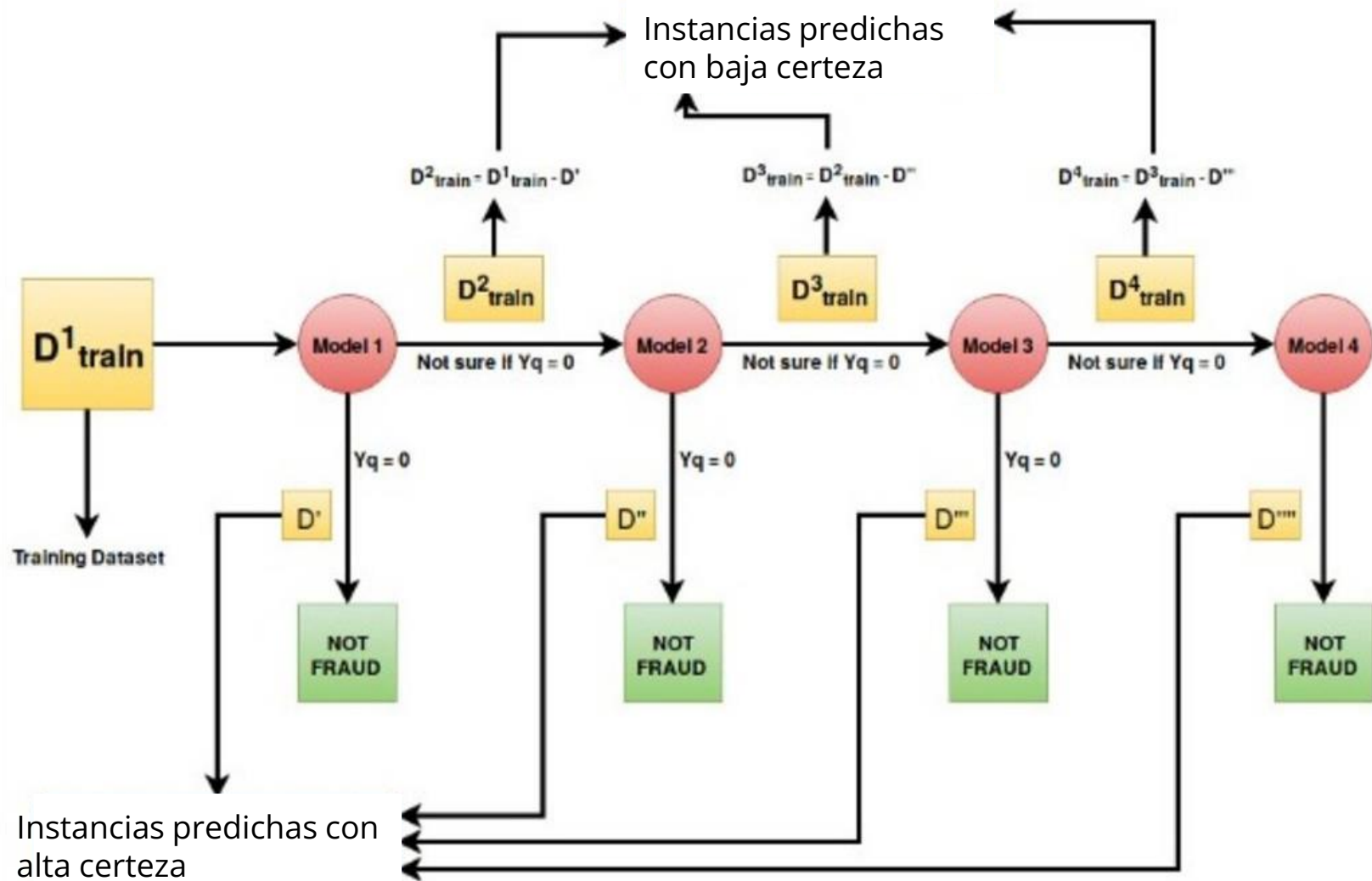
- Enfoque en el que se pasa sucesivamente los datos de un modelo a otro
- A diferencia de Stacking cada “capa” tiene un sólo modelo
- El input de cada modelo son las instancias predichas con poca certeza por el modelo anterior
- Suele utilizarse cuando se necesita una alta certeza en la predicción

Cascading Ejemplo

- Queremos un modelo que prediga si una transacción con tarjeta de crédito es fraudulenta
- Necesitamos una alta certeza para definir si no lo es, en caso de error las pérdidas pueden ser millonarias
- Construimos entonces una secuencia (o cascada) de modelos de ML

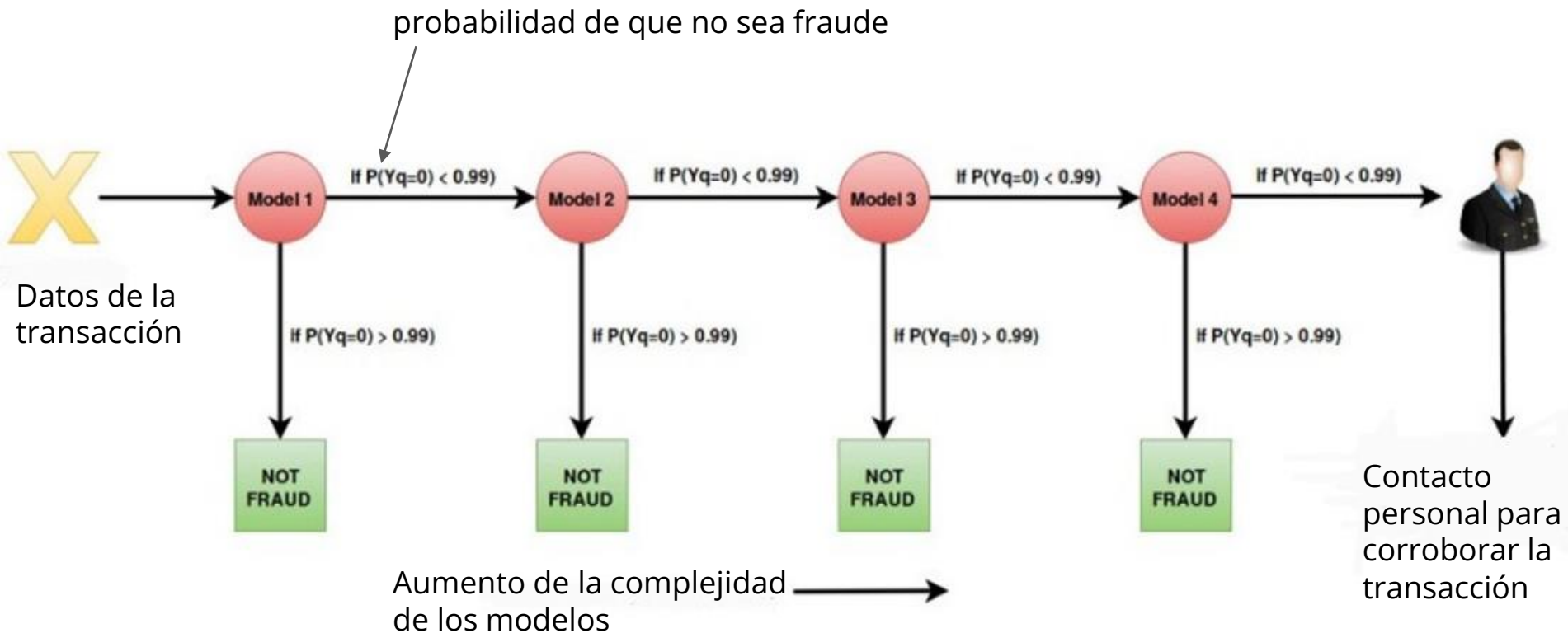
Cascading Entrenamiento

- Entrenamos N modelos distintos utilizando nuestro set de entrenamiento
- Cada modelo lo entrenamos sobre las instancias predichas con baja certeza del modelo anterior
- Suele hacerse de forma tal de empezar por modelos simples y a medida que se entrenan nuevos los mismos sean de mayor complejidad



Cascading predicción

- El primer modelo recibe los datos de la transacción
- Si la probabilidad de que no sea fraude es menor a 0.99, se pasa al siguiente modelo
- Caso contrario se descarta la posibilidad de fraude
- Si ningún modelo descarta el fraude con $p < 0.99$ se procede a confirmar la operación de forma personal



Cascading más características

- A diferencia de Voting y Stacking que tienen un enfoque de modelos “Multi-expertos”, Cascading tiene un enfoque “Multi-estado”
- Inicialmente creados para computer vision
- Cascadas muy profundas pueden producir overfitting

Resumen

- La combinación de clasificadores permite generar clasificadores más precisos a cambio de menor comprensión de los mismos
- Métodos homogéneos básicos: bagging, boosting
 - Combinan mismo tipo de clasificador
 - Manipulación del conjunto de entrenamiento
- Método heterogéneo básico: Voting
 - Combinan distintos tipos de clasificador

Propiedades de los Ensamblados

- Ventajas
 - Generalmente logran mejores predicciones que los modelos vistos hasta ahora
 - Buen *trade-off* sesgo varianza
- Desventajas
 - Se pierde interpretabilidad
 - Tienen una complejidad computacional mayor

Referencias

L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123– 140, 1996.

R. E. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. In 11 th Annual Conference on Computational Learning Theory (COLT1998), pages 80–91.ACM, 1998.

I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2nd edition, 2005.

D. Wolpert. Stacked generalization. *Neural networks*, 5:241–260, 1992.

Greedy Function Approximation: A Gradient Boosting Machine, by Friedman

Heitz, G., Gould, S., Saxena, A., & Koller, D. (2009). Cascaded classification models: Combining models for holistic scene understanding. In *Advances in Neural Information Processing Systems* (pp. 641-648).

[XGBoost en mas profundidad](#)