

Divulga UAB

Revista de difusió de la recerca de la
Universitat Autònoma de Barcelona

23/09/2025

¿Qué etiqueta le asignamos? Clasificación en contextos de incertidumbre



En el ámbito del aprendizaje automático supervisado, la clasificación es una tarea central que asigna etiquetas según su distribución de probabilidades. A menudo se clasifica según el criterio de máxima probabilidad (MAP), pero este puede ser inadecuado cuando las categorías tienen un orden, como en las escalas de satisfacción. Este artículo propone Ord-MAP, una alternativa óptima que abre la puerta a una mejor práctica en la clasificación ordinal.

Clasificar es una de las tareas centrales del aprendizaje automático supervisado, una rama de la inteligencia artificial que construye, a partir de datos previos, sistemas capaces de asignar etiquetas (clasificadores) a nuevos ítems. Los clasificadores no asignan etiquetas de manera determinista, sino que lo hacen en un contexto de incertidumbre: para cada nuevo ítem, proporcionan una distribución de probabilidades sobre las posibles etiquetas.

Cuando las etiquetas son binarias (sí/no, bueno/malo, ...), es habitual asignar la etiqueta con mayor probabilidad —la **moda** de la distribución. Este criterio, conocido como **MAP** (Maximum A Posteriori), no sólo es intuitivo, sino también óptimo si asumimos que todos los

errores de clasificación tienen la misma pérdida asociada, en el sentido de que minimiza la pérdida esperada al clasificar.

El criterio MAP también se utiliza en clasificación multiclase nominal (más de dos etiquetas no ordenadas), bajo la misma hipótesis de simetría en las pérdidas por error de clasificación. Sin embargo, en muchos problemas reales, las etiquetas tienen un orden intrínseco, como en las escalas de satisfacción (muy insatisfecho – insatisfecho – neutro – satisfecho – muy satisfecho), conocidas como escalas de Likert. En este caso, no todos los errores son igual de graves: clasificar a un cliente “muy insatisfecho” como “insatisfecho” es menos grave que hacerlo como “muy satisfecho”.

Cuando las etiquetas están ordenadas, una alternativa más coherente a penalizar todos los errores por igual es tener en cuenta la **distancia** entre etiquetas. En un artículo reciente, se ha propuesto un nuevo criterio de decisión para la clasificación ordinal: **Ord-MAP**, que consiste en asignar la **mediana** de la distribución de probabilidades, es decir, la primera etiqueta para la cual la suma acumulada de probabilidades supera 0.5.

Por ejemplo, si un clasificador devuelve las siguientes probabilidades: 0.35, 0.05, 0.05, 0.30 y 0.25, correspondientes a las cinco categorías de satisfacción, el criterio MAP asignaría la etiqueta “muy insatisfecho” (la de mayor probabilidad), mientras que Ord-MAP elegiría “satisfecho”, porque es la primera categoría para la cual la suma acumulada supera 0.5.

Este criterio, sorprendentemente sencillo, se demuestra matemáticamente en el artículo que es **óptimo** en el sentido de que minimiza la pérdida esperada cuando ésta se define como la distancia entre la etiqueta real y la asignada. Los experimentos realizados con distintos clasificadores y bases de datos reales, así como las simulaciones, muestran claramente la superioridad del criterio Ord-MAP respecto al MAP utilizado habitualmente.

Esta aportación abre la puerta a una mejor práctica en la clasificación ordinal, tan presente en aplicaciones reales como las recomendaciones automáticas, las encuestas o la evaluación de servicios.

Rosario Delgado

Departamento de Matemáticas

Universitat Autònoma de Barcelona

Rosario.Delgado@uab.cat

Referencias

Delgado, R. (2025). **Ord-MAP criterion: Extending MAP for ordinal classification**. *Knowledge-Based Systems* 324, 113837, <https://doi.org/10.1016/j.knosys.2025.113837>

[View low-bandwidth version](#)