

Laboratorio 4: Métricas de Desempeño

Inteligencia Artificial - CC3085

José Antonio Mérida Castejón

4 de febrero de 2026

Task 1 - Teoría

Responda con criterio y análisis de ingeniero. No se esperan definiciones de libro, si no que analice las consecuencias de las decisiones de diseño.

La Mentira de la Independencia: Naive Bayes

En la diapositiva 6 se menciona que Naive Bayes “Asume independencia entre observaciones”. Con esto en mente, considere que usted está construyendo un filtro de spam y su modelo analiza la frase “Cuenta Bancaria”. En el mundo real, la probabilidad de que aparezca la palabra “Bancaria” aumenta drásticamente si ya apareció la palabra “Cuenta”. Responda:

- Si Naive Bayes trata estas dos palabras como eventos independientes (lanzar dos monedas separadas) en lugar de eventos dependientes, ¿está el modelo subestimando o sobreestimando la probabilidad real de la frase conjunta? Justifique su respuesta basándose en la fórmula $P(A \cap B) = P(A) \times P(B)$ vs la realidad.

Naive Bayes, al asumir que los eventos son independientes es completamente incapaz de determinar relaciones entre ellos. Desde un aspecto matemático, podemos decir que $P(B | A) > P(B)$ dado el enunciado que establece que la probabilidad que aparezca “Bancaria” aumenta al aparecer “Cuenta”. Por lo tanto, el modelado más adecuado (bajo la observación dada en el inciso) nos indica que la probabilidad real es $P(A \cap B) = P(A) \times P(B | A)$. Mientras que NB nos dice “estos eventos son completamente independientes, entonces la probabilidad que ocurra B luego de A es $P(A) \times P(B)$ “.

Como establecimos previamente, $P(B | A) > P(B) \implies P(A) \times P(B | A) > P(A) \times P(B)$. Es decir, nuestro modelo subestimaría la probabilidad de la frase real conjunta.

La Economía de los Datos (SVM)

Refierase a las slides correspondientes al tema de SVM. Con esto en mente considere, usted entrena un SVM con 1 millón de datos de partidas de League of Legends. El modelo resultante identifica 5,000 “Vectores de Soporte“. Su jefe le dice que para ahorrar espacio en la base de datos, va a eliminar los otros 995,000 puntos de datos que no son vectores de soporte y re-entrenar el modelo solo con los 5,000 restantes. Responda:

- Matemáticamente, cambiaria la frontera de decisión al hacer esto? Por qué si o por qué no?

No, la frontera sería la misma ya que esta se determina únicamente por los vectores de soporte.

- Explique la eficiencia de memoria de este algoritmo frente a KNN.

Este algoritmo es más eficiente en memoria, en este caso por ejemplo si tuviésemos un KNN deberíamos de mantener el total de 1 millón de puntos en memoria. Mientras que en SVM únicamente debemos mantener los 5,000 vectores de soporte mencionados.

La Miopía de los Árboles En las diapositivas se menciona que la construcción del árbol es un “algoritmo greedy” (codicioso / avaro). Con esto en mente considere, un algoritmo greedy toma la mejor decisión posible en el paso actual sin preocuparse por el futuro. Responda:

- Si el árbol elige el “Mejor Feature” para dividir el nodo raíz porque reduce la impureza drásticamente ahora, ¿garantiza esto que el árbol final será el más óptimo/pequeño posible?

No necesariamente, una división “mala” puede permitirnos realizar divisiones perfectas en los siguientes niveles. Al utilizar un approach greedy, descartamos completamente las posibilidades a futuro y elegimos el beneficio “local” por encima de la solución global.

- Dibuje o describa un escenario lógico donde elegir una división “sub-óptima” al inicio podría llevar a un mejor árbol al final. Qué nombre técnico recibe este fenómeno

A este fenómeno se le llama “greedy search myopia”.