

# Qualification Round

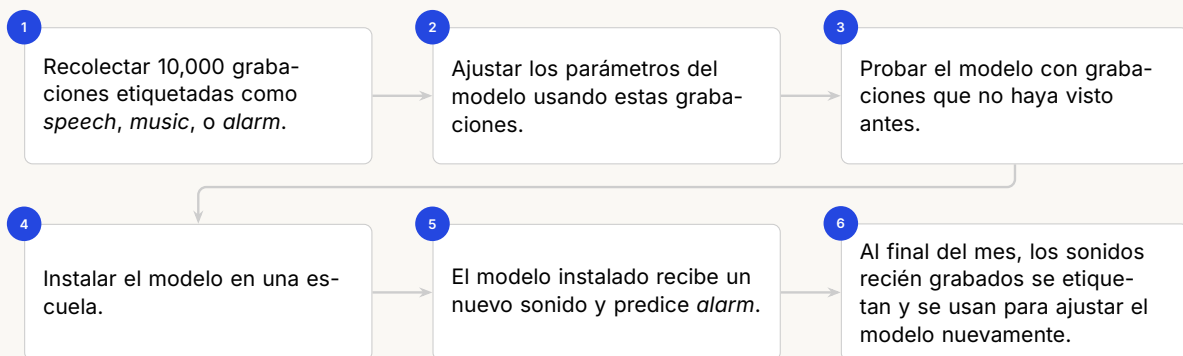
EDITION 2026



## Problem A

[5 points]

Se desarrolla un sistema de reconocimiento de sonido a través de los seis pasos que se muestran a continuación.



- (a) Etiqueta cada paso como data collection, training, evaluation, deployment o inference. Un paso puede tener más de una etiqueta.
- (b) ¿En qué pasos está aprendiendo realmente el sistema? Explica qué significa aprender en este contexto.

## Problem B

[5 points]

Un invernadero controla su techo con este árbol de decisión:



- (a) ¿Qué acción predice el árbol cuando la temperatura es de 26 °C y la humedad es del 68%?

- (b) Se instala un nuevo sensor de CO<sub>2</sub> y el encargado comienza a registrar las condiciones junto con lo que debería haber hecho el techo:

| Temperatura | Humedad | CO <sub>2</sub> | Techo            |
|-------------|---------|-----------------|------------------|
| 31 °C       | 60%     | 700 ppm         | ABRIR            |
| 26 °C       | 75%     | 650 ppm         | ABRIR            |
| 25 °C       | 60%     | 800 ppm         | MANTENER CERRADO |
| 27 °C       | 68%     | 1,100 ppm       | MANTENER CERRADO |
| 24 °C       | 55%     | 1,400 ppm       | ABRIR            |
| 22 °C       | 50%     | 1,550 ppm       | ABRIR            |

Ajusta el árbol de decisión para que modele correctamente cada observación del registro.

### Problem C

[5 points]

Un modelo predice  $y$  a partir de  $x$ . Los datos de entrenamiento son:

| $x$ | 0 | 1   | 2   | 3 |
|-----|---|-----|-----|---|
| $y$ | 1 | 3.2 | 4.8 | 7 |

Se proponen dos modelos:

$$M_1(x) = 0.2x^3 - 0.9x^2 + 2.9x + 1, \quad M_2(x) = 2x + 1.$$

Los modelos se evalúan usando

$$J = \sum (y - \hat{y})^2 + \lambda \sum_i a_i^2,$$

donde la segunda suma recorre los coeficientes del modelo para  $x$ ,  $x^2$  y  $x^3$ .

- (a) Para  $\lambda = 1$ , calcula la puntuación  $J$  de cada modelo. ¿Qué modelo selecciona la regla de puntuación?
- (b) ¿Qué modelo se ajusta mejor a los datos de entrenamiento? Explica por qué la puntuación  $J$  selecciona de todos modos al otro modelo y para qué sirve el término  $\lambda \sum_i a_i^2$ .

### Problem D

[5 points]

Un modelo de lenguaje se entrena para maximizar una recompensa fija de modo que sus resultados se ajusten mejor a las cualidades deseadas, como ser útil y estar alineado con las preferencias humanas. Sin embargo, sin una restricción, el modelo puede sufrir un drift alejándose mucho de su punto de partida confiable, el modelo de referencia, y aprender a manipular la recompensa en lugar de producir respuestas genuinamente buenas.<sup>1</sup>

<sup>1</sup>Esta es una explicación simplificada de RLHF (reinforcement learning from human feedback). Para una introducción, consulta por ejemplo <https://huggingface.co/blog/rlhf>.

Sea  $t \geq 0$  la medida del drift respecto al modelo de referencia. Supongamos que la recompensa es  $rt$  y la penalización por drift es  $\beta t^2$ , con  $r, \beta > 0$ . La pérdida resultante es

$$L(t) = -rt + \beta t^2.$$

y el entrenamiento elige el valor de  $t$  que la minimiza. En sistemas reales, el drift se mide típicamente usando la divergencia KL; aquí usamos  $t^2$  por simplicidad.

- Encuentra el desplazamiento  $t^*$  que minimiza la pérdida, y el valor de la pérdida que alcanza.
- ¿Qué sucede con  $t^*$  cuando  $\beta \rightarrow 0$  y cuando  $\beta \rightarrow \infty$ ? Explica en una o dos frases qué significa cada caso para el modelo de lenguaje.
- La estimación de la recompensa  $r$  tiene ruido: los ingenieros solo saben que  $0 < r \leq r_{\max}$ . Ellos requieren que el desplazamiento se mantenga por debajo de un límite de seguridad  $T$ , sin importar cuál sea el valor de  $r$ . Demuestra que esto está garantizado exactamente cuando  $\beta \geq \frac{r_{\max}}{2T}$ .
- ¿Qué tiene en común el término  $\lambda \sum_i a_i^2$  del Problema C con el término de penalización  $\beta t^2$  de aquí?

## Problem E

[5 points]

Una organización quiere introducir un sistema de aprendizaje automático para agricultores en las zonas rurales de Nepal. Un agricultor proporciona información sobre un cultivo y el sistema recomienda si debe regar, aplicar un tratamiento, esperar o contactar a un experto. La organización quiere implementar un modelo de lenguaje extenso ya existente, como ChatGPT, Claude o Gemini. Analiza tres oportunidades y tres riesgos de introducir este sistema.



## Instrucciones de Participación

- ✓ Escribe tus soluciones a mano en hojas de papel o escríbelas en una computadora.
- ✓ Envía tus soluciones en línea antes de la fecha límite que aparece en el sitio web.  
Sitio web: <https://imlco.org/submission>
- ✓ No es necesario que incluyas el enunciado de los problemas en tu documento de solución.
- ✓ Muestra tu procedimiento para recibir el puntaje completo. Puntos máximos: 25.

- ✓ Etiqueta claramente cada problema y resalta tus respuestas finales.
- ✓ Necesitas obtener al menos **15/17/20 puntos** como Junior/Youth/Senior para clasificar a la Pre-Final Round. Consulta <https://imlco.org/age-groups> para más detalles.
- ✓ Si tienes preguntas, contáctanos en: [info@imlco.org](mailto:info@imlco.org)