

# Modelos de predicción y optimización de Teapot

Descripción técnica de cómo Teapot estima el aprendizaje de cada estudiante y decide qué ejercitar para maximizar su puntaje esperado en la PAES.

29 de abril de 2026 | Equipo Teapot

## Resumen ejecutivo

Teapot construye sus recomendaciones combinando dos modelos. Un **modelo de predicción**, entrenado con 2,96 millones de respuestas históricas de Filadd, estima el rendimiento esperado de un estudiante en cada subunidad del temario en función de su historial y de cuánto va a ejercitar. Un **modelo de optimización** usa esa predicción para decidir cuántos ejercicios asignar de cada subunidad, de manera que se maximice el puntaje ponderado esperado en la PAES dentro del tiempo disponible y respetando el orden curricular. Este documento describe ambos modelos.

## Visión general

El estudiante tiene un *historial* de ejercitación (qué ha hecho, cuánto, cuándo, con qué desempeño) y un *tiempo disponible* para estudiar la próxima semana. La pregunta operativa es: ¿cuántos ejercicios debe hacer de cada subunidad?

Teapot la responde en dos pasos. Primero, una función predictiva  $\hat{p}$  traduce “hacer  $x$  ejercicios de la subunidad  $s$ , dado el historial actual” a una probabilidad estimada de acierto en  $s$ . Segundo, un modelo de optimización busca la combinación de  $x$  que maximiza el puntaje ponderado esperado del estudiante, sujeto a sus restricciones de tiempo y a las dependencias del temario.

## Modelo de predicción del aprendizaje

### Qué predice

Para cada subunidad  $s$  del temario y cada periodo  $t$  (la semana en curso, por ejemplo), el modelo entrega:

$$\hat{p}(x_s^t, \vec{H}_s^t) \in [0, 1],$$

la probabilidad estimada de que el estudiante responda correctamente una pregunta de  $s$  durante  $t$ , donde:

- $x_s^t$ : cantidad de ejercicios de  $s$  que el estudiante hará en  $t$  (variable de decisión).
- $\vec{H}_s^t$ : vector con el historial del estudiante en  $s$  hasta el inicio de  $t$  (datos conocidos).

En palabras: dado lo que el estudiante ya hizo,  $\hat{p}$  estima qué tan bien le iría en  $s$  si ejercita  $x_s^t$  veces más esta semana. Esta es la pieza que el modelo de optimización utiliza para comparar planes de estudio alternativos.

## Datos de entrenamiento

El modelo se entrenó sobre el historial 2024 de Filadd: 5.002.608 registros iniciales, donde cada fila corresponde a la realización de un ejercicio bajo condiciones instruccionales estándar. Tras descartar ejercicios sin etiquetar, respuestas NO RESPONDE, intentos repetidos del mismo ejercicio y otras inconsistencias, el dataset final tiene **2.959.904 ejercicios** (40,8 % removidos en limpieza).

## VARIABLES CONSTRUIDAS

A partir de la base limpia se generó un dataset agrupado por estudiante, subunidad y periodo. Cada fila resume el contexto del estudiante en una subunidad mediante un conjunto de **21 variables** que el modelo recibe como entrada. Salvo por la variable de decisión  $x_s^t$  —la cantidad de ejercicios de  $s$  que el estudiante hará en el periodo  $t$ —, las demás cumplen distintos roles:

- Variables que permiten al modelo aprender el fenómeno del **olvido**, capturando cuánto tiempo lleva el estudiante sin practicar la subunidad.
- Variables que describen el **desempeño reciente** del estudiante en la subunidad, en ventanas temporales de distinta duración.
- Variables que describen su **trayectoria acumulada** a lo largo de toda la historia disponible.
- Variables que estiman su **nivel de dominio inicial**, antes de comenzar a practicar deliberadamente.

El detalle exacto de cada variable y su construcción es información propietaria de Teapot y no se incluye en este documento.

## Horizontes temporales

Para soportar tanto decisiones semanales como planificación a varias semanas, se entrenaron **cuatro versiones** del modelo, una por horizonte:  $t = 7, 14, 21$  y  $28$  días. Cada versión interpreta  $x_s^t$  como “cantidad de ejercicios en una ventana de esa duración” y  $MASTERY_s^t$  como la proporción de aciertos en esa misma ventana.

## Precisión

El indicador principal es el **MAE** (error absoluto medio): cuán lejos cae, en promedio, el porcentaje de aciertos predicho respecto del observado. Un MAE de 0,07 significa que la predicción se desvía en promedio 7 puntos porcentuales del rendimiento real del estudiante. La Tabla 1 resume el desempeño en la partición de prueba (30 % de las muestras).

En la práctica, el MAE general se mantiene entre 6,8 % y 7,1 % en todos los horizontes, y es razonablemente homogéneo entre asignaturas. Esto significa que, al evaluar el efecto esperado de un plan de estudio, la predicción introduce un error promedio de  $\approx 7$  puntos porcentuales por subunidad, suficientemente bajo para que la comparación entre planes alternativos sea informativa.

## Modelo de optimización

El modelo decide, para un periodo  $t$  (típicamente la próxima semana), cuántos ejercicios asignar de cada subunidad para maximizar el puntaje ponderado esperado del estudiante en la PAES.

## Notación

- $P$ : conjunto de pruebas PAES relevantes para el estudiante (según su carrera/universidad objetivo).
- $U_p$ : unidades de la prueba  $p \in P$ .
- $S_u$ : subunidades de la unidad  $u \in U_p$ .
- $\omega_p$ : ponderación que la carrera/universidad objetivo asigna a la prueba  $p$ .

Tabla 1: MAE de  $\hat{p}(x_s^t, \vec{H}_s^t)$  por horizonte y prueba (partición de prueba).

Horizonte	Prueba	Muestras totales	Muestras de prueba	MAE
7 días	Matemática M1	11.270	3.387	0,073
	Matemática M2	1.221	371	0,070
	Ciencias	31.949	9.589	0,071
	Competencia Lectora	1.467	441	0,074
	Historia y Cs. Sociales	5.104	1.534	0,061
	<b>General</b>	<b>51.011</b>	<b>15.322</b>	<b>0,071</b>
14 días	Matemática M1	12.372	3.719	0,070
	Matemática M2	1.253	380	0,076
	Ciencias	29.825	8.952	0,069
	Competencia Lectora	1.467	441	0,074
	Historia y Cs. Sociales	4.832	1.451	0,063
	<b>General</b>	<b>49.749</b>	<b>14.943</b>	<b>0,069</b>
21 días	Matemática M1	13.001	3.907	0,070
	Matemática M2	1.268	385	0,074
	Ciencias	28.232	8.476	0,069
	Competencia Lectora	1.425	428	0,071
	Historia y Cs. Sociales	4.568	1.373	0,063
	<b>General</b>	<b>48.494</b>	<b>14.569</b>	<b>0,069</b>
28 días	Matemática M1	13.399	4.028	0,072
	Matemática M2	1.301	395	0,073
	Ciencias	27.104	8.137	0,066
	Competencia Lectora	1.382	415	0,073
	Historia y Cs. Sociales	4.470	1.343	0,061
	<b>General</b>	<b>47.656</b>	<b>14.318</b>	<b>0,068</b>

- $n_s$ : número de preguntas de  $s$  que aparecen en la PAES.
- $\text{puntaje}_p(\cdot)$ : función oficial que convierte aciertos en la prueba  $p$  a puntaje PAES (escala 100–1000).
- $\hat{p}(x_s^t, \vec{H}_s^t)$ : probabilidad de acierto en  $s$  predicha por el modelo de la sección anterior.
- $x_s^t$ : variable de decisión, ejercicios a realizar de  $s$  en  $t$ .

### Derivación de la función objetivo

**Paso 1: aciertos esperados en una subunidad.** Sea  $Z_s$  la variable aleatoria que cuenta los aciertos del estudiante en las preguntas de  $s$  que aparecerán en la PAES, y  $p_s$  la probabilidad real (desconocida) de que acierte una de esas preguntas. Modelando cada pregunta como un ensayo independiente con probabilidad  $p_s$ :

$$\mathbb{E}[Z_s \mid n_s, p_s] = n_s \cdot p_s.$$

**Paso 2: estimaciones operativas.** Como  $n_s$  y  $p_s$  no se conocen, se reemplazan por estimaciones:

$$n_s \approx \hat{n}_s, \quad p_s \approx \hat{p}(x_s^t, \vec{H}_s^t),$$

donde  $\hat{n}_s$  es el promedio histórico de apariciones de  $s$  en versiones pasadas de la PAES y  $\hat{p}$  es la salida del modelo de predicción. Reemplazando:

$$\mathbb{E}[Z_s] \approx \hat{n}_s \cdot \hat{p}(x_s^t, \vec{H}_s^t).$$

**Paso 3: aciertos esperados en una prueba completa.** Sea  $Y_p$  el total de aciertos del estudiante en la prueba  $p$ . Como  $Y_p$  es la suma de los aciertos por subunidad,

$$Y_p = \sum_{u \in U_p} \sum_{s \in S_u} Z_s,$$

y por linealidad de la esperanza:

$$\mathbb{E}[Y_p] \approx \sum_{u \in U_p} \sum_{s \in S_u} \hat{n}_s \cdot \hat{p}(x_s^t, \vec{H}_s^t).$$

**Paso 4: del número de aciertos al puntaje PAES.** El DEMRE no usa los aciertos directamente, sino una conversión no lineal  $\text{puntaje}_p(\cdot)$  que transforma aciertos en puntaje (100–1000). Aplicándola al total de aciertos esperado se obtiene el puntaje esperado en la prueba  $p$ :

$$\text{puntaje}_p \left( \sum_{u \in U_p} \sum_{s \in S_u} \hat{n}_s \cdot \hat{p}(x_s^t, \vec{H}_s^t) \right).$$

**Paso 5: puntaje ponderado de ingreso.** El estudiante ingresa a una carrera/universidad cuyo puntaje de selección pondera las pruebas con  $\omega_p$ . La función objetivo de Teapot suma esos puntajes ponderados sobre todas las pruebas relevantes:

$$\boxed{\max_{x \in X} \sum_{p \in P} \omega_p \cdot \text{puntaje}_p \left( \sum_{u \in U_p} \sum_{s \in S_u} \hat{n}_s \cdot \hat{p}(x_s^t, \vec{H}_s^t) \right)}$$

donde  $x = (x_s^t)_s$  es el vector de asignaciones y  $X$  es el conjunto factible definido por las restricciones que se describen a continuación.

## Restricciones

### Datos adicionales.

- $\tau_s^{\text{solución}}$ : tiempo promedio para resolver un ejercicio de  $s$ .
- $\tau_s^{\text{corrección}}$ : tiempo promedio para revisar un ejercicio incorrecto de  $s$ .
- $\tau^{\text{estudio}}(x^t)$ : tiempo estimado para estudiar la teoría de las subunidades nuevas incluidas en  $x^t$ .
- $\delta_t$ : tiempo total disponible del estudiante en el periodo  $t$ .
- $\text{prerrequisitos}(s)$ : subunidades que deben haberse estudiado antes de  $s$ .
- $\text{min\_ejercicios}$ : mínimo de ejercicios para considerar una subunidad estudiada.
- $s_{\text{nuevas}}(t)$ : cantidad mínima de subunidades nuevas a abrir en  $t$ , calibrada para que el estudiante alcance a cubrir todo el temario antes de la PAES.

**(1) Tiempo disponible.** La asignación no puede consumir más tiempo que el que el estudiante declaró tener. El tiempo se descompone en resolver, corregir errores, y estudiar teoría de contenidos nuevos:

$$\sum_{u \in U} \sum_{s \in S_u} \left( \tau_s^{\text{solución}} x_s^t + \tau_s^{\text{corrección}} (1 - \hat{p}(x_s^t, \vec{H}_s^t)) x_s^t \right) + \tau^{\text{estudio}}(x^t) \leq \delta_t.$$

El término con  $(1 - \hat{p})$  refleja que solo los ejercicios incorrectos generan tiempo adicional de revisión.

**(2) Prerrequisitos curriculares.** Una subunidad  $s$  solo puede recibir ejercicios si sus prerrequisitos ya se trabajaron lo suficiente:

$$x_s^t > 0 \Rightarrow \sum_{t' < t} x_{s'}^{t'} \geq \text{min\_ejercicios}, \quad \forall s' \in \text{prerrequisitos}(s).$$

**(3) Subunidades nuevas mínimas.** Para garantizar avance curricular, en cada periodo se debe abrir un mínimo de subunidades que el estudiante todavía no ha tocado:

$$\left| \{s : x_s^t > 0 \wedge \sum_{t' < t} x_s^{t'} = 0\} \right| \geq s_{\text{nuevas}}(t).$$

### Formulación compacta

Reuniendo todo, el problema que resuelve Teapot en cada ejecución es:

$$\begin{aligned} & \max_{x \in \mathbb{Z}_+^{|S|}} \sum_{p \in P} \omega_p \cdot \text{puntaje}_p \left( \sum_{u \in U_p} \sum_{s \in S_u} \hat{n}_s \cdot \hat{p}(x_s^t, \vec{H}_s^t) \right) \\ & \text{sujeto a} \quad (1) \text{ tiempo disponible,} \\ & \quad \quad \quad (2) \text{ prerrequisitos curriculares,} \\ & \quad \quad \quad (3) \text{ subunidades nuevas mínimas.} \end{aligned}$$

En palabras: *elegir cuántos ejercicios hacer de cada subunidad de manera que la combinación maximice el puntaje ponderado esperado del estudiante, sin exceder el tiempo disponible, respetando el orden del temario y abriendo suficientes contenidos nuevos cada periodo.* La salida  $x^*$  de este problema es la recomendación que Teapot entrega al estudiante.