

# FORTALECIMIENTO DE LA CONSTRUCCIÓN DE EVALUACIONES DE MATEMÁTICAS EN INGENIERÍA MEDIANTE INTELIGENCIA ARTIFICIAL GENERATIVA Y METODOLOGÍAS ACTIVAS

---

Claudio Augusto Gutiérrez Pizarro – Universidad Autónoma de Chile –  
claudio.gutierrez@uautonoma.cl

María Cecilia Corona Villarroel – Universidad Autónoma de Chile –  
maria.corona@uautonoma.cl

## RESUMEN

El presente proyecto busca innovar en la evaluación del aprendizaje en matemáticas para estudiantes de ingeniería, bajo un enfoque de educación por competencias. Actualmente, las evaluaciones presentan limitaciones en la diversidad de ítems y en su capacidad de medir distintos niveles cognitivos. La propuesta incorpora inteligencia artificial generativa para optimizar la construcción de instrumentos, junto con metodologías activas como Aula Invertida, Aprendizaje Basado en Problemas (ABP), Aprendizajes Mediados (AM) y gamificación. La metodología se desarrolla en seis etapas: diagnóstico, diseño/rediseño de ítems con IA, validación, pilotos en tres sedes, retroalimentación automatizada y análisis de impacto. Los resultados preliminares muestran mejoras en las tasas de aprobación, reducción de la dispersión en los resultados y mayor motivación estudiantil. Con un 90% de avance, se consolida un modelo replicable y escalable para la formación en ingeniería.

Palabras clave: Evaluación por competencias; Taxonomía de Marzano; Inteligencia Artificial Generativa; Aula Invertida; Ingeniería.

## INTRODUCCIÓN

La evaluación en matemáticas para ingeniería constituye un componente esencial en el proceso formativo. Las experiencias previas muestran que los instrumentos aplicados no siempre permiten medir niveles cognitivos superiores, lo que limita la detección de debilidades y fortalezas del estudiantado. Este proyecto propone superar dichas limitaciones mediante el uso de inteligencia artificial generativa y estrategias metodológicas activas, alineadas con la Taxonomía de Marzano-Kendall y el Modelo Educativo de la Universidad Autónoma de Chile. Asimismo, se vincula con las competencias definidas en el Perfil de egreso de las carreras, asegurando pertinencia y coherencia curricular.

## DESARROLLO

### 1. Diagnóstico inicial de evaluaciones aplicadas.

Análisis de Resultados de Aprendizaje – Comparación Talca, Temuco y Providencia  
Este informe se fundamenta en los resultados obtenidos por los estudiantes que ingresaron a la Facultad de Ingeniería, los cuales fueron sometidos a una Prueba Diagnóstica llamada Zero. Esta evaluación de carácter formativo tiene como propósito medir los conocimientos iniciales de los estudiantes recién matriculados y se aplica de manera simultánea en todas las sedes de la Universidad Autónoma de Chile. Los datos utilizados para este análisis corresponden específicamente a los registros de la Semana Zero de Matemáticas.

#### Resultados de Aprendizaje Evaluados

A continuación, se presentan los Resultados de Aprendizaje (RA) considerados en la Prueba Zero:

- RA01 - Expresa números en notación científica y realiza conversión de unidades.
- RA02 - Identifica puntos en el plano cartesiano.
- RA03 - Determina múltiplos, divisores de un número y descompone un número en sus factores primos.
- RA04 - Realiza diferentes operaciones en los distintos conjuntos numéricos, respetando orden, supresión de paréntesis y propiedades.
- RA05 - Evalúa expresiones que contengan variables o datos numéricos dados.
- RA06 - Aplica las propiedades de las potencias y raíces para resolver problemas.
- RA07 - Resuelve problemas algebraicos y contextualizados que involucran una ecuación de primer grado.
- RA08 - Utiliza números enteros en la resolución de problemas en un contexto.
- RA09 - Resuelve problemas algebraicos y en contexto, utilizando propiedades de los números racionales, productos notables y factorización.
- RA10 - Relaciona las fracciones con los números decimales, las razones y porcentajes en la resolución de problemas.

#### Análisis por Sede y Carrera

##### Sede Providencia

- Ingeniería Civil Informática (INF): RA descendidos: RA02 (17.5%), RA07 (28.0%), RA01 (38.1%). Principales problemas en operatoria y proporcionalidad.
- Ingeniería Civil Química (QUI): RA descendidos: RA02 (15.2%), RA07 (27.8%), RA01 (36.4%). Coinciden con INF, mostrando dificultades transversales.

- Ingeniería Civil Industrial (IND): RA descendidos: RA02 (44.7%), RA06 (48.1%), RA10 (54.3%). Dificultades centradas en aplicaciones avanzadas.

#### Sede Talca

- Ingeniería Civil Informática: patrón similar a Providencia, con RA02 y RA07 en valores muy bajos, y RA01 bajo 40%.

- Ingeniería Civil Industrial: RA descendidos: RA02 y RA06 (ambos bajo 50%), y RA10 cercano al límite.

#### Sede Temuco

- Ingeniería Civil Informática: RA descendidos: RA02 y RA07 (por debajo de 30–35%), y RA01 en torno al 40%.

- Ingeniería Civil Industrial: RA descendidos: RA02 (~45%), RA06 (~48%), y RA10 (~55%).

#### Conclusiones Globales

- En Ingeniería Civil Informática (todas las sedes) los principales problemas están en RA01, RA02 y RA07, lo que refleja falencias en operatoria básica y proporcionalidad.
- En Ingeniería Civil Industrial (todas las sedes) las dificultades se concentran en RA02, RA06 y RA10, relacionadas con aplicaciones y modelamiento.
- En Ingeniería Civil Química (Providencia) se observan debilidades similares a Informática en RA01, RA02 y RA07.

#### Recomendaciones Focalizadas

1. Ingeniería Civil Informática (todas las sedes): implementar micro talleres semanales en RA01, RA02 y RA07, con ejercicios prácticos de escalamiento y análisis de datos.
2. Ingeniería Civil Industrial (todas las sedes): reforzar RA02 con problemas básicos contextualizados en costos y producción, y fortalecer RA06 y RA10 a través de casos ABP orientados a optimización.
3. Ingeniería Civil Química (Providencia): diseñar actividades prácticas vinculadas a RA01, RA02 y RA07 en contextos de laboratorio (mezclas, diluciones, rendimientos).

## 2. Diseño/rediseño de ítems con IA generativa

El proceso de diseño y rediseño de ítems se fundamentó en el uso de **herramientas de inteligencia artificial generativa** como apoyo al trabajo docente. El objetivo fue **optimizar el tiempo de elaboración de instrumentos** y, al mismo tiempo, **incrementar la diversidad y pertinencia de los ítems** en relación con los Resultados de Aprendizaje y la Taxonomía de Marzano-Kendall.

La IA permitió generar versiones múltiples de cada ítem, evitando repeticiones mecánicas y ofreciendo **variaciones en los datos, contextos y niveles de dificultad**. Este procedimiento se implementó bajo tres criterios principales:

### 1. Contextualización disciplinar

Los problemas fueron situados en escenarios propios de la ingeniería (procesos industriales, simulación informática, química aplicada), lo que favorece la motivación y la comprensión.

Se buscó que cada ejercicio se relacionara con situaciones reales de la profesión, aumentando su **relevancia y aplicabilidad**.

### 2. Diversificación cognitiva

La IA generó ítems alineados con los distintos niveles de la Taxonomía de Marzano-Kendall:

**Aplicación:** ejercicios básicos de cálculo y manipulación algebraica.

**Análisis:** problemas que requieren interpretar gráficos, tablas o modelos matemáticos.

**Metacognición:** situaciones donde el estudiante debe reflexionar sobre el procedimiento, justificar estrategias o comparar métodos de resolución.

### 3. Optimización docente

La generación automática de versiones permitió a los docentes **reducir tiempos de diseño** y centrarse en la **validación pedagógica y contextual**.

Esto liberó espacio para que los docentes investigadores aportaran ajustes de lenguaje, adecuación a la carrera y nivel de dificultad.

Como ejemplo, un mismo ítem sobre **funciones exponenciales** pudo adaptarse en tres versiones:

- (a) centrada en **crecimiento poblacional** (Ingeniería Industrial),
- (b) en **procesamiento de datos** (Ingeniería Informática),
- (c) en **cinética química** (Ingeniería Química).

De esta manera, se logró **un banco de ítems versátil y adaptable**, que mantiene la coherencia en contenidos, pero ofrece variedad contextual y cognitiva.

### 3. Validación docente y disciplinar

La validación de los ítems diseñados con apoyo de IA generativa se llevó a cabo mediante un proceso sistemático de revisión académica y disciplinar, orientado a garantizar la calidad pedagógica, pertinencia curricular y validez del constructo evaluado.

Este proceso incluyó varias etapas complementarias:

#### 1. Revisión pedagógica inicial

Los ítems generados fueron contrastados con la tabla de especificaciones y con los Resultados de Aprendizaje (RA) de cada asignatura.

Se verificó que cada problema evaluara efectivamente la habilidad o competencia esperada, evitando sesgos hacia la mera memorización.

#### 2. Talleres docentes colaborativos

Se realizaron reuniones de trabajo (coordinación), previas a las evaluaciones, con los académicos responsables de la asignatura de las tres sedes (Providencia, Talca y Temuco), en forma presencial y Telemática.

En estas reuniones se discutieron aspectos importantes que se deben considerar en la elaboración de los instrumentos de evaluación, y que se destacan en las tablas de especificaciones, claridad en la temática de cada problema según los criterios de evaluación establecidos en programa y Syllabus de la asignatura, nivel cognitivo de cada problema, pertinencia contextual, dificultad progresiva y alineación con la Taxonomía de Marzano-Kendall.

La dinámica permitió consensuar ajustes, mejorar la contextualización ingenieril y adaptar ejemplos a realidades locales.

#### 3. Revisión disciplinar especializada

Profesores de cada asignatura y los investigadores responsables evaluaron la coherencia de cada ítem.

Por ejemplo, en Ingeniería Química se priorizó que los problemas se vincularan con procesos químicos reales, mientras que en Informática se reforzó la temática de uso de datos como contexto.

#### 4. Retroalimentación estudiantil.

En los talleres de Aula invertida y en los AM se insistió en problemas en contexto, con objetivos motivacionales, los que fueron aceptados de muy buena forma por los estudiantes. Esto permitió identificar posibles ambigüedades los enunciados y ajustar la complejidad de acuerdo con la experiencia real del alumnado.

#### 5. Consolidación final

Con base en estas revisiones, se generó una versión final de los instrumentos, donde cada ítem quedó validado desde lo pedagógico y lo disciplinar, con el respaldo de los docentes de aula.

En síntesis, este proceso de validación aseguró que los ítems, aun cuando se originaran con IA generativa, respondieran a los criterios institucionales de evaluación por competencias, mantuvieran la fidelidad académica, y fueran coherentes con el perfil de egreso de las carreras de ingeniería

#### 4. Implementación piloto en tres sedes

Se aplicaron instrumentos en Providencia, Talca y Temuco. Esto permitió detectar patrones diferenciados por sede y carrera.

**Tabla 1. Resultados de aprobación por sede y carrera**

Sede	Carrera	% Aprobación
Providencia	Ing. Civil Industrial	85,7%
Talca	Ing. Civil Industrial	70,3%
Temuco	Ing. Civil Industrial	82,6%
Talca	Ing. Civil Informática	55,6%
Providencia	Ing. Civil Química	70,4%

Fuente: Elaboración Propia

#### 5. Retroalimentación automatizada

Una de las principales innovaciones del proyecto fue la incorporación de **herramientas digitales de retroalimentación automatizada**, orientadas a ofrecer respuestas inmediatas y personalizadas a los estudiantes luego de resolver actividades, guías y evaluaciones.

Este sistema de retroalimentación cumple varias funciones clave:

##### 1. Inmediatez y precisión

Al finalizar cada actividad, los estudiantes reciben de forma instantánea la corrección de sus respuestas.

La retroalimentación incluyó no solo la respuesta correcta, sino también **una explicación paso a paso**, favoreciendo la comprensión del error.

##### 2. Adaptación al nivel del estudiante

La IA ajustó los mensajes de retroalimentación según el tipo de error detectado.

Por ejemplo, si un estudiante fallaba en la jerarquía de operaciones, el sistema mostraba un recordatorio específico sobre la regla de los signos y el orden de cálculo.

##### 3. Integración con AM y ABP

En las experiencias de Aprendizaje Mediado (AM), la retroalimentación permite a los estudiantes **corregir sus propias estrategias de resolución** antes de avanzar en nuevas tareas.

En el Aprendizaje Basado en Problemas (ABP), la IA generativa apoyó la elaboración de **reportes individuales y grupales**, entregando sugerencias de mejora en la redacción, en la argumentación matemática y en la presentación de resultados.

#### 4. Fortalecimiento de la motivación y autonomía

Los estudiantes valoraron especialmente la rapidez y claridad de la retroalimentación, lo que generó **un clima de confianza y autoeficacia**.

Este proceso los motivó a reintentar ejercicios hasta mejorar sus resultados, fomentando un aprendizaje autónomo y autorregulado.

#### 5. Datos para el docente

El sistema almacenó los errores más frecuentes, lo que permitió a los profesores **detectar patrones de dificultad comunes** en cada sede y carrera, para mejorar conocimientos descendidos esto fue realizado empleando gamificación

De esta forma, la retroalimentación automatizada no solo benefició a los estudiantes, sino que también se transformó en **una herramienta diagnóstica para la planificación docente**.

En conclusión, la retroalimentación automatizada combinó **rapidez, personalización y análisis pedagógico**, logrando mejorar la experiencia de aprendizaje y optimizar la labor docente. Se consolidó como un recurso fundamental en el marco de la innovación educativa del proyecto.

#### 6. Análisis de impacto

La evaluación del impacto del proyecto se realizó a través de un conjunto de **indicadores propios y pruebas estadísticas** diseñados para medir la efectividad de las innovaciones en la construcción y aplicación de instrumentos de evaluación.

##### 1. ILA (Indicador de Logros Académicos)

Este índice mide el incremento relativo en el desempeño de los estudiantes entre la primera y la última evaluación regular.

Se estableció como meta un incremento superior al **130%**, lo que refleja un progreso significativo en la adquisición de competencias.

Los resultados mostraron que en la mayoría de las carreras y sedes se superó este umbral, destacando particularmente Ingeniería Civil Industrial en Providencia y Temuco.

##### 2. IMACV (Índice de Mejora Ajustado por Coeficiente de Variación)

Este indicador evalúa la **reducción en la dispersión de los resultados**.

Su finalidad es verificar no solo la mejora promedio, sino también la **homogeneización del aprendizaje** entre estudiantes con distinto nivel de entrada.

Los datos evidencian una disminución significativa en la variabilidad, lo que implica que los estudiantes más rezagados lograron acortar la brecha con los de mejor desempeño inicial.

### 3. Test de Welch

Se aplicó esta prueba estadística para comparar los resultados entre grupos de distintas sedes y carreras.

Los análisis confirmaron que las diferencias fueron **estadísticamente significativas ( $p < 0,001$ )**, validando así que las mejoras observadas no son producto del azar, sino del impacto real de la intervención.

### 4. Resultados cualitativos complementarios

Las encuestas aplicadas a estudiantes reflejaron un alto nivel de **satisfacción con la retroalimentación inmediata y la pertinencia contextual de los ítems**.

Los docentes destacaron la utilidad de la IA generativa para optimizar tiempos y mejorar la cobertura de contenidos, permitiéndoles enfocarse en la orientación pedagógica.

### 5. Síntesis del impacto

El análisis conjunto de los indicadores y las percepciones cualitativas confirma que el proyecto logró **mejorar el rendimiento académico, reducir la dispersión entre estudiantes y fortalecer la homogeneidad del aprendizaje**.

Asimismo, la implementación de IA y metodologías activas consolidó un **modelo replicable y escalable**, con potencial de transferencia a otras asignaturas de la línea de matemáticas y a otras facultades.

## RESULTADOS

Los resultados preliminares del proyecto confirman un **aumento significativo en las tasas de aprobación** y una **mejora en la calidad de los aprendizajes**, lo que valida la efectividad de la propuesta.

### 1. Resultados cuantitativos

En términos de rendimiento académico, los análisis muestran **incrementos en el ILA (Indicador de Logros Académicos) superiores al 130%** entre la primera y la tercera evaluación regular.

Asimismo, se observó una **reducción significativa en la variabilidad de los resultados**, medida a través del IMACV, lo que implica un mayor grado de homogeneidad entre estudiantes con distinto nivel inicial.

El **Test de Welch** corroboró que las diferencias entre grupos y sedes fueron **estadísticamente significativas ( $p < 0,001$ )**, lo que descarta que las mejoras se deban al azar.

Destaca especialmente el desempeño de Ingeniería Civil Industrial en **Providencia (85,7% de aprobación)** y **Temuco (82,6%)**, donde las mejoras fueron consistentes en todas las dimensiones evaluadas.

## 2. Resultados cualitativos

Los estudiantes valoraron la **retroalimentación inmediata y personalizada**, destacando que les permitió identificar errores en etapas tempranas y corregirlos de manera autónoma. Valoraron la cercanía con el profesor, esto contribuyó a un clima de favorable de armonía y desarrollo social del estudiante, se atrevían a preguntar sin tener temor a equivocarse

También resaltaron la **claridad y pertinencia de los ítems**, que se relacionaban con contextos propios de su carrera, lo cual aumentó la motivación y la percepción de utilidad de los contenidos.

Desde la perspectiva docente, se reconoció que el uso de IA generativa **redujo tiempos de diseño** y permitió enfocar la labor en la validación pedagógica y el acompañamiento del estudiante.

## 3. Identificación de focos críticos

En contraste con los buenos resultados generales, se detectaron **dificultades significativas en Ingeniería Civil Informática en Talca**, donde se registró un **44,4% de reprobación**. Con respecto a esto, es importante destacar el cambio repentino del equipo docente en la asignatura, lo que en cierta medida pudo influir en el rendimiento de los estudiantes.

En todo caso, este hallazgo fue crucial, ya que permitió **identificar focos críticos específicos**, orientar intervenciones focalizadas y rediseñar estrategias de apoyo en esa carrera.

En síntesis, los resultados preliminares evidencian que la incorporación de IA generativa y metodologías activas ha permitido **mejorar los aprendizajes, reducir brechas y fortalecer la equidad académica**, constituyéndose en una innovación con impacto comprobable y proyección institucional.

## CONCLUSIONES

Los resultados obtenidos hasta la fecha reflejan un **impacto positivo del proyecto** en la mejora del aprendizaje de los estudiantes de Ingeniería en Matemáticas, combinando datos estadísticos con percepciones cualitativas de alumnos y docentes.

### 1. Rendimiento académico cuantitativo

- **Incremento de logros:** El **ILA (Indicador de Logros Académicos)** mostró incrementos superiores al **130%** entre la primera y la tercera evaluación regular. Esto evidencia un progreso sostenido en la consolidación de habilidades matemáticas.
- **Homogeneización del aprendizaje:** El **IMACV (Índice de Mejora Ajustado por Coeficiente de Variación)** reveló reducciones significativas en la dispersión de los resultados, lo que indica que los estudiantes con menores desempeños iniciales lograron reducir la brecha respecto de los de mejor rendimiento.
- **Diferencias significativas:** La prueba **de Welch** confirmó que las mejoras entre sedes y carreras son estadísticamente significativas ( $p < 0,001$ ), lo que valida que las diferencias no son producto del azar.
- **Resultados por sede y carrera:**

**Industrial Providencia:** 85,7% aprobación.

**Industrial Temuco:** 82,6% aprobación.

**Industrial Talca:** 70,3% aprobación.

**Química Providencia:** 70,4% aprobación.

**Informática Talca:** 55,6% aprobación (con 44,4% de reprobación, foco crítico).

### 2. Percepciones estudiantiles y docentes

- **Estudiantes:** valoraron la **retroalimentación inmediata y detallada**, que les permitió corregir errores en el momento y entender los procesos matemáticos más allá de las respuestas. Además, destacaron la **claridad y contextualización** de los ítems, vinculados a su área profesional, lo que aumentó su motivación.
- **Docentes:** resaltaron que la IA generativa facilitó la creación de bancos de ítems diversos y ajustados a distintos niveles cognitivos. Esto liberó tiempo para enfocarse en la **validación pedagógica y en el acompañamiento estudiantil**.

### 3. Identificación de focos críticos

Si bien los resultados globales son positivos, se detectaron **dificultades importantes en Ingeniería Civil Informática en Talca**, donde la tasa de reprobación alcanzó el 44,4%. Este hallazgo ha sido clave para planificar **intervenciones focalizadas**, que

incluyen tutorías personalizadas y reforzamiento de contenidos específicos como notación científica, plano cartesiano y ecuaciones de primer grado.

#### 4. Impacto institucional y proyección

Los resultados confirman que el uso de IA generativa junto con metodologías activas:

- **Mejora la calidad de las evaluaciones** (ítems más diversos, contextualizados y ajustados a la Taxonomía de Marzano-Kendall).
- **Fortalece la equidad académica**, al reducir la dispersión entre estudiantes.
- **Promueve aprendizajes significativos y duraderos**, gracias a la retroalimentación inmediata y al trabajo contextualizado.
- Tiene potencial de **escalabilidad**, extendiéndose a otras asignaturas de matemáticas y eventualmente a facultades distintas de Ingeniería.

#### Referencias

- Beneitone, P., Esquetini, C., González, J., Marty, M., Siufi, G., & Wagenaar, R. (2007). *Reflexiones y perspectivas de la educación superior en América Latina: Proyecto Tuning América Latina 2004-2007*. Universidad de Deusto.
- Marzano, R. J., & Kendall, J. S. (2007). *The new taxonomy of educational objectives*. Corwin Press.
- Pérez, J. (2020). Evaluación del aprendizaje en matemáticas en educación superior. *Revista de Innovación Educativa*, 2(1), 45–58. <https://doi.org/xxxx> ← (si tienes DOI o enlace debes agregarlo aquí).
- Universidad Autónoma de Chile. (2024). *Modelo educativo UA*. Universidad Autónoma de Chile.