

Sistema de recomendación de formación continua para docentes, directivos y asistentes

Este proyecto propone desarrollar una plataforma web (Django+Python, SQLite) que aplique aprendizaje automático para segmentar a los profesionales de la educación < y sugerirles capacitaciones personalizadas. Como fuentes de datos se usará únicamente la s tablas seleccionadas de **ENDDEIE 2023** , que caracteriza la integración de las tecnologías digitales en los establecimientos chilenos. Dicha encuesta indagó dimensiones como el marco institucional, apoyo al uso de TIC, acceso a tecnologías, actividades docentes con TIC, habilidades digitales, actitudes y efectos de las tecnologías en distintos actores. Además, cubre innovación educativa (liderazgo, prácticas innovadoras, mentalidad, promotores/ barreras). Partiendo de estos lineamientos, el proyecto formará perfiles a partir de variables clave de ENDDEIE y generará recomendaciones formativas vinculadas a esos perfiles. El sistema diferenciará dos roles: **administrador** (configura parámetros, visualiza métricas) y **usuario** (docente/directivo/asistente que completa formulario y recibe sugerencias).

Objetivos del proyecto

- **General:** Diseñar un sistema inteligente capaz de **identificar perfiles profesionales** mediante clustering y ofrecer **recomendaciones personalizadas** de formación continua, optimizando la inversión en desarrollo profesional del personal educativo.
- **Específicos:**
 - (1) Aplicar algoritmos de clustering para segmentar docentes/directivos/asistentes según competencias digitales, uso de TIC y actitudes, usando datos ENDDEIE
 - (2) Construir un motor de recomendación híbrido (filtrado basado en contenido + colaborativo) que sugiere capacitaciones acorde al perfil detectado.
 - (3) Crear una interfaz web que permita al usuario completar el formulario de perfil, visualizar su clúster asignado y acceder a cursos sugeridos.
 - (4) Incorporar un módulo administrador para ajustar parámetros (p. ej. número de clusters, umbrales) y monitorear métricas de uso.
 - (5) Evaluar la efectividad del sistema mediante métricas cuantitativas (silhouette, Precision@K, Recall@K, etc.) y retroalimentación de usuarios ³ .

Justificación técnica y metodológica

Se adopta la metodología **CRISP-DM** (Cross-Industry Standard Process for Data Mining) por ser un estándar probado en proyectos de análisis de datos . Este enfoque iterativo guiará todas las etapas:

comprensión del negocio y de los datos, preparación, modelado de clustering, construcción del sistema de recomendación, evaluación y despliegue . CRISP-DM es flexible y permite personalizar cada fase

según necesidades concretas ⁴ . En el contexto educativo, esta metodología facilita alinear el producto

con objetivos pedagógicos (diagnóstico de necesidades formativas) y técnicos (calidad de datos, validación de clusters). El uso de técnicas de **aprendizaje no supervisado** (clustering) y de **recomendación híbrida** responde a la complejidad de perfilar usuarios y sugerir recursos formativos precisos. Se aprovecharán bibliotecas Python como scikit-learn y Pandas para el modelado, y se utilizará **SQLite** como base de datos relacional ligera para un prototipo escalable.

Variables críticas de ENDDEIE para segmentación

Para crear perfiles detallados se emplearán variables representativas extraídas de la ENDDEIE. Además de las categorías de **estrato analítico** (escuela rural, urbana, liceo HC o TP) y **dependencia** (municipal, subvencionada, particular pagada) utilizadas en el estudio, se consideran las siguientes dimensiones relevantes:

- **Habilidades digitales:** autoevaluación de competencias en uso de herramientas informáticas y TIC (habilidades generales y avanzadas, ciudadanía digital, programación, etc.) ^{6 1}.
- **Uso de TIC en docencia:** frecuencia de integración de tecnologías en la planificación y clases, participación en actividades formativas asociadas a TIC, uso de plataformas educativas, desarrollo de contenidos digitales, etc.
- **Actitudes y motivación:** nivel de motivación respecto al uso de TIC, creencias y expectativas positivas hacia la innovación (sub-dimensión "Actitudes: motivación, creencias y expectativas" ⁷). Se incluyen

indicadores como interés en contenidos digitales y percepción de la importancia de las tecnologías.

- **Acceso y recursos tecnológicos:** disponibilidad de equipos (computadores, tablets, internet) en el establecimiento y en el hogar, cobertura de conectividad, así como soporte técnico y pedagógico recibido (apoyo al uso de TIC).

- **Formación previa en TIC:** antecedentes de capacitación o talleres recibidos (incluso por rol: docente o directivo) en tecnologías, robótica, IA, etc. ENDDEIE registra formación/capacitación del personal educativo ⁸.

- **Liderazgo e innovación:** medidas de liderazgo educativo para la innovación (visiones compartidas, fomento de prácticas innovadoras) y de gestión de cambio (promotores/barreras para innovar) ^{6 1}.

- **Contexto institucional:** aspectos como prioridad dada a TIC en el plan institucional, existencia de proyectos digitales, y estado socio-económico del establecimiento (vulnerabilidad, nivel SEP, etc.) — variables disponibles o inferibles en ENDDEIE.

Estas variables permiten caracterizar perfiles formativos con base en competencias y necesidades digitales ^{6 7}. Por ejemplo, dos docentes del mismo estrato/dependencia podrían diferir en su

motivación por las TIC o en sus **habilidades digitales**, lo que justificaría recomendaciones distintas.

Estructura recomendada del formulario

El formulario web para usuarios (docentes, directivos o asistentes) debe ser breve pero representativo, con secciones claras:

- **Identificación del perfil:** cargo (docente/directivo/asistente), asignatura o área que enseña (para docentes), nivel educacional (básica/media), años de experiencia.
- **Competencias digitales:** preguntas de escala Likert sobre autopercepción en habilidades TIC (similar a indicadores de ENDDEIE), experiencia en herramientas digitales (uso de software educativos, plataformas online, etc.).
- **Uso actual de tecnologías:** frecuencia de actividades con TIC en el aula, niveles de integración en la planificación, tipo de actividades digitales realizadas.
- **Formación previa:** listado de capacitaciones recientes en tecnología educativa o pedagogía digital que ha recibido (sí/no).
- **Motivación y actitudes:** ítems breves para medir interés en innovación y predisposición a usar tecnología, importancia percibida de las TIC (p.ej., escala de acuerdo con afirmaciones sobre TIC en educación).
- **Recursos y soporte:** preguntas sobre acceso a internet/dispositivos en el establecimiento y autopercepción del apoyo institucional (técnico y pedagógico).

El formulario utilizará preguntas cerradas (selección múltiple y escalas) derivadas de la ENDDEIE, pero reducido a ~10–15 variables clave. Esto facilita el agrupamiento sin sobrecargar al usuario. Además, permitirá al sistema asignar el perfil en un clúster basado en respuestas cuantificables.

Flujo de usuario y arquitectura del sistema

El sistema sigue un flujo sencillo: el **usuario** (docente, directivo o asistente) accede al portal web, crea una cuenta y completa el formulario de perfil. El servidor procesa las respuestas (preprocesamiento y asignación de clúster) y retorna en pantalla una lista de sugerencias de capacitación, con enlaces a recursos formativos externos (videos, cursos, etc.). El usuario puede explorar estas recomendaciones y marcar favoritas.

El **administrador** se autentica con permisos especiales para acceder a un dashboard con métricas de uso (número de usuarios, clusters activos, recomendaciones más vistas) y a una sección de configuración. Ahí puede ajustar parámetros del algoritmo (p.ej. número de clusters en K-Means, umbral de similitud) y editar el contenido del formulario (agregar/quitar preguntas clave).

En cuanto a arquitectura técnica, se sugiere un esquema en capas:

- **Presentación:** interfaz web implementada en Django, usando su motor de plantillas y formularios. Para prototipado rápido se puede usar Django nativo o front-end ligero (Bootstrap).
- **Lógica de negocio:** código Python (Django views y modelos) que integra el pipeline de datos. Incluye módulos para el preprocesamiento de respuestas, modelado de clustering (scikit-learn) y recomendación.
- **Base de datos:** SQLite (integrada con Django) almacena datos de usuarios, respuestas del formulario, clusters asignados y catálogo de cursos. SQLite es suficiente para prototipos; en producción puede migrarse a PostgreSQL sin cambiar código significativamente.
- **Modelos analíticos:** bibliotecas Python (scikit-learn, pandas) hospedadas en el servidor procesan los datos del formulario. Por ejemplo, tras obtener la respuesta de un usuario, el sistema cargará el vector de características y aplicará el modelo de clustering entrenado.

Esta arquitectura modular facilita escalabilidad y mantenimiento. Django provee gestión de usuarios, formularios y sesiones; el algoritmo de clustering y recomendación se encapsula en servicios o scripts separados. La integración con SQLite permite almacenar y recuperar eficientemente la información de los usuarios y los parámetros del sistema ⁹.

Algoritmos de agrupamiento recomendados

Para segmentar los perfiles se valoran métodos de clustering sin supervisión adaptados al tipo de datos (numéricos o categóricos). Se recomiendan algoritmos clásicos como **K-Means** (óptimo para datos numéricos bien distribuidos) y **DBSCAN** o **clustering jerárquico aglomerativo** (útiles si los perfiles no son esféricos o hay ruido) ³. K-Means++ es una opción inicial para definir clusters sin partir de

centroides arbitrarios ³. Si la cantidad de datos es moderada (miles de usuarios), K-Means puede

ajustarse dinámicamente (p. ej. optimizando el número K con el método del codo o Silhouette). DBSCAN identifica agrupaciones por densidad, lo cual ayuda si existen perfiles mayoritarios y minoritarios atípicos. El clustering jerárquico (enlace completo o promedio) puede revelar subgrupos o similitud gradual entre perfiles.

Como alternativa, se puede usar **Mixture Models** gaussianos si los datos se asumen multimodales, o algoritmos de **redes neuronales autoorganizadas** (SOM) para visualización, aunque K-Means y DBSCAN suelen ser suficientes en un primer prototipo educativo ³. En cualquier caso se validarán las

particiones con métricas como el **Silhouette Score** y se revisarán interpretabilidad y cohesión de los clusters.

Consideraciones técnicas con Django, Python y SQLite

- **Django:** aprovecha su sistema de *models, views y templates* para construir la aplicación. Se definen modelos de datos para usuarios, respuestas del formulario y recomendaciones. Se utiliza el sistema de autenticación incorporado para manejar perfiles de **usuario vs administrador**. El panel de administración de Django puede habilitarse para gestionar contenidos (texto del formulario, recursos de capacitación) sin codificar. Los formularios web se construyen con `forms.ModelForm` o `forms.Form` de Django, validando las entradas del usuario.
- **Python & análisis de datos:** el backend ejecuta scripts Python que procesan los datos recibidos. Se usará **Pandas** para manipulación inicial y **scikit-learn** para algoritmos de clustering y recomendación de contenido (p.ej. implementar filtros basados en similitud de texto de cursos). El procesamiento debe realizarse en el servidor Django en segundo plano (podría usarse Django-rq/Celery si los cálculos tardan).
- **SQLite:** base de datos ligera que Django configura por defecto. Es útil para prototipos y entornos de desarrollo, y almacena la tabla de respuestas procesables y un registro de los cursos recomendados. Si el sistema crece, se recomienda migrar a PostgreSQL; Django facilita este cambio en configuración. Para la persistencia de modelos (clusters entrenados) se pueden almacenar parámetros en la base de datos o en archivos serializados.
- **Escalabilidad y rendimiento:** aunque el proyecto inicial es a nivel de prueba de concepto, se diseñará considerando que puede crecer nacionalmente. Django junto con SQLite puede atender cientos de consultas simultáneas; en caso de mayor demanda se puede implementar caching, balanceo de carga o microservicios (por ejemplo, un servicio aparte de Python puro para el cálculo de recomendaciones). El uso de SQLite implica limitaciones de concurrencia, por lo que para muchos usuarios simultáneos se debería configurar una base de datos más robusta.
- **Seguridad y despliegue:** es importante proteger datos personales (LDAP/autenticación segura, HTTPS) y desplegar la aplicación en un servidor web confiable. Django tiene buenas prácticas incorporadas (protección CSRF, gestión de sesiones). Finalmente, se documentarán e implementarán tests unitarios y un plan de pruebas general para asegurar la calidad del sistema.

En conjunto, esta solución integral permitirá agrupar a los profesionales según sus necesidades de formación identificadas por las respuestas a variables clave (habilidades, uso de TIC, motivación, etc.); y con base en el perfil resultante proveerá cursos y recursos formativos adecuados, facilitando la toma de decisiones en la formación docente continua.

Conclusiones y Recomendaciones Informe ENDDIE 2023

La Encuesta Nacional De Desarrollo Digital Escolar e Innovación Educativa (ENDDEIE) 2023 se desarrolló con el objetivo de “caracterizar la transformación digital y la innovación educativa en los establecimientos educacionales en Chile como un medio para que las y los/as estudiantes se desarrollen de forma integral y colaborativa en las sociedades del siglo XXI”. La encuesta indagó la dimensión de desarrollo digital, con temáticas de marco institucional, apoyo al uso, acceso, actividades, habilidades, actitudes y efectos de las tecnologías digitales en distintos actores de los establecimientos educacionales. También se incorporó la dimensión de innovación educativa, con temáticas de liderazgo educativo, prácticas y procesos, mentalidad, promotores/barreras y la innovación en el proceso de enseñanza y aprendizaje.

La ENDDEIE fue administrada a directores/as de establecimientos urbanos y rurales, docentes de establecimientos urbanos, estudiantes de establecimientos rurales y urbanos, coordinadores de informática de establecimientos urbanos, más una pauta de infraestructura administrada a todos los establecimientos participantes. Su administración se llevó a cabo a partir de seis instrumentos auto aplicados de manera online, con excepción de la pauta de infraestructura la cual fue completada por un miembro del equipo en terreno. La muestra total de participantes que fue utilizada para los análisis y reporte de resultados correspondió a 1.206 establecimientos educacionales, organizados en los estratos analíticos de escuela rural, escuela urbana, liceo humanista científico y liceo técnico profesional, y de dependencias municipal, particulares subvencionados y particulares pagados.

Se revisan los principales resultados a partir de las dimensiones indagadas: desarrollo digital e innovación educativa.

10.1 Desarrollo digital escolar

Esta dimensión abordó el conjunto de elementos que promueven la integración efectiva de las tecnologías digitales (o Tecnologías de la Información y Comunicación - TIC) en el proceso de enseñanza y aprendizaje. Este concepto abarca diversos aspectos: en primer lugar, el **marco institucional**, que engloba la gestión de estas tecnologías, la innovación institucional, la comunicación y la vinculación. En segundo lugar, el **apoyo al uso**, que incluye la formación en el uso de tecnologías digitales y programación, así como el soporte técnico y pedagógico. En tercer lugar, el **acceso**, que se refiere a la disponibilidad de recursos digitales y la infraestructura necesaria. En cuarto lugar, las **actividades**, que abarcan desde el uso recreativo hasta las prácticas de enseñanza y aprendizaje con dichas tecnologías. En quinto lugar, las **habilidades**, que comprenden la madurez y competencias digitales. En sexto lugar, las **actitudes**, que involucran la motivación y expectativas hacia las tecnologías digitales; y finalmente, los **efectos**, que se refieren a la percepción del impacto en el aprendizaje de los/as estudiantes y el desempeño profesional.

En primer lugar, se aprecia que ha habido **avances en la implementación de infraestructura digital en los establecimientos**, donde el acceso a estas tecnologías presenta un porcentaje de logro del 87,9%, sin embargo, los liceos HC están aún algo atrás (86,1%). Estos resultados podrían ser reflejo del énfasis en la provisión de infraestructura de la política de informática educativa de los últimos años, así como de la prioridad que los establecimientos han dado a la adquisición de infraestructura digital. Si bien, esta es una condición necesaria para la integración de tecnologías digitales en la enseñanza y aprendizaje (Ertmer et al., 2012; Kim et al., 2013; Tondeur et al., 2012), no resulta suficiente. Al respecto, tal como muestran los otros resultados, estos esfuerzos no necesariamente están acompañados de las condiciones necesarias para lograr una utilización efectiva de esta infraestructura para la enseñanza y aprendizaje.

Asimismo, se observan diferencias relevantes en la **formación de habilidades digitales**, donde los/as estudiantes de establecimientos particulares pagados declaran haber participado de más actividades de formación que sus pares de otros establecimientos. Dada la importancia de estas habilidades a nivel nacional e internacional, esto da cuenta de una falencia que resulta necesario abordar para asegurar que los/as estudiantes participen de manera efectiva en la sociedad digital. Consistentemente, respecto a la autopercepción de las habilidades digitales, resulta evidente la brecha entre los/as docentes y estudiantes de establecimientos subvencionados (municipales y particulares) respecto a los de establecimientos particulares pagados. Esta sigue siendo una dimensión que requiere atención, ya que, de mantenerse esta tendencia, podría implicar una ampliación de la brecha asociada a competencias digitales, por lo que resulta un aspecto crítico a abordar, tanto como una competencia necesaria para desempeñarse en la sociedad (OECD, 2021b; UNESCO, 2023), como para el uso efectivo de estas tecnologías en educación (Lin et al., 2023; Pedaste et al., 2023; Timotheou et al., 2023).

En **términos institucionales**, los resultados muestran que el equipo directivo presenta interés para favorecer el uso de tecnologías digitales en sus establecimientos, otra condición que según diversos autores (Olszewski & Crompton, 2020; Wu et al., 2019), favorece la apropiación de estas tecnologías en educación. Sin embargo, las medidas tomadas son cuestionadas por los/as docentes, quienes muestran respuestas divididas respecto a que en sus establecimientos existan las condiciones necesarias para el uso de tecnologías digitales. Así, se observa que a pesar de que existe un alto grado de integración del uso de tecnologías digitales en las definiciones estratégicas de los establecimientos, no necesariamente éstas van acompañadas de las habilidades digitales requeridas para aprovecharlas, así como de las condiciones institucionales para la integración de tecnologías digitales en la enseñanza, tal como lo perciben los/as docentes. Al respecto, según los resultados de PISA 2022, sólo un 36,8% de los/as estudiantes chilenos asiste a establecimientos que cuenta con un programa específico para preparar a los/as estudiantes para un comportamiento responsable en Internet, en comparación a un promedio del 64,3% de los/as estudiantes en los países de la OCDE (OECD, 2023). Asimismo, sólo un 40,6%

de los/as estudiantes asiste a establecimientos que cuentan con un programa específico para fomentar la colaboración en el uso de dispositivos digitales entre los/as docentes, en comparación con un 54,6% en los países de la OCDE (OECD, 2023).

En relación con las **actividades de los/as estudiantes con tecnologías digitales en el establecimiento educacional**, su frecuencia es relativamente baja, esto es, sólo algunas clases al mes, lo cual coincide con los resultados de PISA 2022, que muestran que un 48,3% de los/as estudiantes declara que usa las tecnologías digitales para actividades de aprendizaje 10 horas a la semana o menos, en comparación con el 34,4% de los/as estudiantes de los países de la OCDE. En contraste el 52,9% afirma que las usa para entretenerse más de 20 horas a la semana (promedio OCDE: 62,1%) (OECD, 2023). Este escenario es consistente con lo declarado por los/as docentes, que mencionan utilizar tecnologías digitales en el 61% de su tiempo de clases, pero afirman que sus estudiantes lo hacen sólo el 37% del tiempo. De manera similar, menos de la mitad de los/as docentes utiliza estas tecnologías para labores de apoyo profesional tres o más veces a la semana.

Adicionalmente, es necesario tener presente que algo más de un tercio de los/as estudiantes declara hacer un uso indebido de internet (ej. copiar trabajos) y un cuarto, menciona que se han realizado acciones asociadas a ciberbullying frecuentemente o siempre. Esto merece ser abordado ya que podría ser una tendencia creciente.

Si bien en base a estos resultados, es posible concluir que hay una relativa baja frecuencia de uso de tecnologías digitales como apoyo a la enseñanza y el aprendizaje, también hay datos interesantes respecto al **uso de celulares en clases**, ya que cerca del 60% de estudiantes declara que lo utiliza para tomar registro de la clase o realizar actividades educativas a solicitud de los/as docentes en algunas o todas las clases de la semana. Consistentemente, entre los/as docentes, un 75% declara que permite el uso de celulares para buscar información o datos relacionados con el contenido de la clase. Estos resultados están en línea con los datos PISA 2022, que muestran que el 82,5% de los/as estudiantes chilenos está en desacuerdo o muy en desacuerdo con que se les prohíba llevar su celular a clases (promedio de la OCDE es 79,1%) (OECD, 2023). Esto abre un interesante espacio de exploración respecto al aprovechamiento de estos dispositivos en el contexto escolar. Sin embargo, es necesario considerar que hay importantes diferencias respecto al uso de celulares en cuanto al grupo etario, ya que en las escuelas el uso es mucho menos frecuente que en liceos, y en cuanto a dependencia, en los establecimientos particulares subvencionados es más frecuente que en los demás tipos de establecimientos.

En cuanto a las **actitudes y percepciones de los efectos del uso de las tecnologías digitales**, a diferencia de los/as directores/as y docentes que se muestran muy motivados por el uso y reconocen su potencial impacto positivo, los/as estudiantes se muestran más moderados en su interés respecto al uso de tecnologías digitales en el contexto escolar, lo que, según las teorías de adopción de tecnologías digitales (Davis et al., 1989; Venkatesh et al., 2003),

podría atentar en contra de su uso para actividades de aprendizaje. Asimismo, resulta interesante que los/as directores/as y docentes de establecimientos particulares pagados tienden a expresar una percepción más crítica respecto a los impactos potenciales de estas tecnologías.

10.2 Innovación educativa

En cuanto a los aspectos asociados al **liderazgo educativo para la innovación**, los resultados dan cuenta de que, en general, los equipos directivos generan las condiciones necesarias para innovar. Esto se traduce en que la implementación de innovaciones en los establecimientos educativos ha sido frecuente en los últimos dos años, con énfasis en fortalecer aprendizajes, condiciones para la enseñanza y convivencia. Sin embargo, persisten desafíos en la participación de familias y colaboración entre docentes de los diferentes establecimientos. Asimismo, destaca la alta valoración del rol de las tecnologías digitales para innovar en educación, particularmente entre los/as docentes de establecimientos municipales y particulares subvencionados.

Respecto a los **promotores y barreras para innovar**, destacan los **factores externos**, como principales barreras para los docentes, incluyendo la evaluación docente, las políticas educativas del Mineduc y los resultados de mediciones de logro, especialmente por parte de los docentes de liceos TP y establecimientos municipales.

En cuanto a la **innovación en la enseñanza y aprendizaje**, destaca la aparente contradicción entre la alta percepción de innovación de los/as directivos y docentes, respecto a las prácticas innovadoras que se implementan en el aula, especialmente lo relacionado con metodologías activas. Esto sugiere la necesidad de profundizar y operacionalizar este tipo de conceptos, desarrollando y promoviendo modelos concretos de lo que resulta ser innovador en la enseñanza y aprendizaje. En esta dimensión resulta interesante que, en general, los/as estudiantes de escuelas rurales tienden a expresar mayores grados de acuerdo respecto a su protagonismo y al logro de aprendizajes en clases, así como una mayor frecuencia de actividades de aprendizaje activo. Por el contrario, los/as estudiantes de establecimientos particulares pagados tienden a estar más bajos en estas dimensiones que sus pares de otras dependencias.

Asimismo, los resultados del IIE muestran la necesidad de **fortalecer la mentalidad frente a la innovación**, es decir, la disposición de los distintos actores con respecto a la implementación de cambios y mejoras en sus establecimientos. Al respecto, si bien, la disposición del liderazgo para la implementación de cambios y mejoras de los/as directores, la autopercepción de la forma de pensar de los/as estudiantes y la percepción del rol de la tecnología en la innovación de los/as docentes resultan con niveles de logro altos, la percepción de la disponibilidad de las condiciones necesarias para innovar de los/as directores y la flexibilidad en las prácticas de docentes, resultan con niveles de logro más

bajos, especialmente entre las escuelas rurales, liceos TP y los establecimientos municipales.

En términos generales, los resultados muestran que **el IIE es mayor en escuelas rurales y establecimientos municipales**. Por el contrario, los establecimientos particulares pagados tienden a mostrar menores niveles de logro en casi todas las dimensiones de innovación educativa.

Desde un punto de vista metodológico, en términos generales, considerando los altos grados de acuerdo de los distintos actores respecto a los temas planteados, sería recomendable profundizar en la manera en que los actores comprenden los diversos aspectos consultados, ya que, en muchas preguntas, la alta concentración de respuestas en los tramos superiores de la escala podría estar reflejando ciertos sesgos de los actores, especialmente entre directivos y docentes.