

Caso de éxito

Análisis automatizado de causas de deserción de aprendices mediante Inteligencia Artificial – SENA CTA Cartago Valle

Programa de Fortalecimiento de Habilidades y Herramientas de Inteligencia Artificial para el Sector Público

Nombre Autor (a): Edison de Jesús Durango León
2025

1. Nombre del Caso/Iniciativa

Análisis automatizado de causas de deserción de aprendices mediante Inteligencia Artificial – SENA CTA Cartago Valle

2. Entidad(es) Responsable(s)

Servicio Nacional de Aprendizaje (SENA), Regional Valle del Cauca, específicamente en el Centro de Tecnología Agroindustriales (CTA) – área de formación académica en Electrónica y Automatización

3. Sector Administrativo (Salud, Justicia, Educación, etc.)

Educación

4. Área de Aplicación

Soporte a la Toma de Decisiones y Planificación Estratégica

5. Problema Público Abordado

El proceso que se buscó optimizar fue el análisis de los motivos de deserción de aprendices. Antes de la intervención, las razones de retiro se registraban de forma manual en campos de texto libre y en entrevistas, lo que obligaba al personal a revisar uno a uno cientos de registros para intentar identificar causas predominantes. Esto implicaba cerca de 12 horas semanales de trabajo manual, generaba demoras, inconsistencias entre analistas, falta de estandarización en las categorías y hacía muy difícil detectar patrones o cruzar información con variables como género, programa o tipo de deserción. El problema principal era, por tanto, la ineficiencia y la falta de datos estructurados para comprender las causas reales de abandono y diseñar estrategias de permanencia basadas en evidencia

6. Solución de IA Implementada

La solución consistió en construir un flujo de trabajo apoyado en Inteligencia Artificial para transformar esos registros cualitativos en datos estructurados. En una primera etapa se exportaron los registros históricos de aprendices desertores, incluyendo el campo “Motivo de retiro”. Luego se realizó una limpieza y anonimización de la información, eliminando nombres y elementos irrelevantes para proteger datos personales y mejorar la calidad del insumo. Posteriormente se utilizó el modelo de lenguaje ChatGPT mediante un prompt específico que indicaba que actuara como analista del SENA y que clasificara cada motivo en categorías predefinidas

(económicas, laborales, académicas, personales/familiares, traslado/residencia, falta de motivación u otros), devolviendo siempre una estructura del tipo “Categoría | Subcategoría | Resumen breve”. La herramienta procesó los textos y entregó resultados tabulares, listos para ser llevados a una hoja de cálculo. Después se verificó manualmente una muestra del 10 % de los registros, con una precisión reportada del 88 %, y se ajustó el prompt cuando fue necesario. Finalmente, se utilizaron esos datos clasificados para generar tablas dinámicas y gráficos que mostraron la distribución de causas de deserción y sirvieron como base para la toma de decisiones en bienestar y seguimiento académico

7. Tecnologías Utilizadas (ej. PLN, Visión por Computador, Machine Learning)

La tecnología central fue el procesamiento de lenguaje natural (PLN) mediante un modelo de lenguaje generativo (ChatGPT, versión GPT-4/GPT-5), utilizado para leer y clasificar textos en lenguaje natural. Como soporte se emplearon herramientas de ofimática como Microsoft Excel y Google Sheets para organizar, limpiar, analizar y visualizar los datos ya clasificados. Todo el flujo se apoya en IA generativa para la clasificación automática y en hojas de cálculo para el análisis estadístico y la representación gráfica de resultados.

8. Resultados

Antes de la intervención, el análisis manual de deserción demandaba alrededor de 12 horas semanales; con la solución de IA el mismo trabajo se realiza en menos de una hora, lo que equivale a una optimización operativa cercana al 91 %. La herramienta permitió procesar y clasificar más de 600 registros históricos, con una precisión del 88 % frente a la validación humana. Los resultados mostraron que aproximadamente el 37 % de las deserciones obedecen a causas económicas, el 25 % a motivos personales o familiares, el 20 % a razones laborales y el porcentaje restante a factores académicos o de traslado de residencia. Estos datos facilitaron la elaboración automática de tablas y gráficos, mejoraron la coherencia y estandarización de los registros y redujeron de forma importante los errores de interpretación derivados de criterios subjetivos. En términos cualitativos, la solución fortaleció la calidad del seguimiento académico, hizo viable el análisis de tendencias y permitió que bienestar y formación diseñaran acciones más focalizadas (apoyos económicos, acompañamiento psicosocial, refuerzo académico, etc.)

9. Factor de Sostenibilidad y Escalabilidad

La solución aporta valor público al mejorar la capacidad del SENA para comprender por qué el alumnado abandona los programas y, con ello, orientar estrategias para la permanencia y el éxito formativo. La entidad obtiene información estructurada para decidir con mayor fundamento dónde invertir esfuerzos de bienestar, apoyos económicos y acompañamiento académico. Para la ciudadanía, esto se traduce en mayores oportunidades de culminar la formación técnica o tecnológica y en un uso más eficiente de los recursos públicos destinados a formación para el trabajo. Desde la perspectiva de la sostenibilidad y la escalabilidad, el caso muestra que el modelo es replicable en otros centros del SENA y en otras entidades, porque se basa en herramientas de bajo costo o ya disponibles (ChatGPT, Excel, Google Sheets), no exige conocimientos avanzados de programación y se apoya en un flujo metodológico documentado. Bastan acceso a internet, una herramienta de IA generativa, una hoja de cálculo y una capacitación básica en formulación de prompts y validación de resultados para que pueda escalarse a nivel regional o nacional.