Alberto Gertzain Sánchez González

Senior Data & AI ▪ Data Science ▪ Tech Lead ▪ Automation

PERFIL

Profesional senior en ingeniería de datos, ciencia de datos e IA con **5+ años en banca Tier-1** (Citibanamex, HSBC), diseñando modelos y liderando técnicamente soluciones críticas desde el análisis hasta su operación en producción.

Perfil híbrido: conecto negocio, riesgo y tecnología para convertir necesidades complejas en soluciones medibles, escalables y auditables. Combino profundidad técnica, criterio analítico y orientación a resultados.

STACK PRINCIPAL

Python	●●●●●
SQL	●●●●●
PySpark	●●●●●
scikit-learn	●●●●●
XGBoost	●●●●●
ETL / Big Data	●●●●●
MLflow	●●●●○
Docker / FastAPI	●●●●○
Airflow	●●●●○
PyTorch	●●●●○
PostgreSQL	●●●●○

Power BI • Streamlit • Kubernetes • SAS • R • Hadoop • SHAP • Optuna • Git • Jira • Alteryx

IA AGÉNTICA

En producción para uso propio:

Manejo de Claude y ChatGPT • cursos de Udemy

Desarrollo de IA local: Ollama / LLMs • RAG • generación estructurada (JSON forzado) • n8n

DOMINIO DE NEGOCIO

Riesgo crediticio • Scoring • Detección de fraude • Reporting regulatorio • Modelado actuarial • VaR / CVaR • Stress testing • Tarificación GLM • Model governance

IDIOMAS

Español (Nativo)	●●●●●
Inglés (B2)	●●●●○

CONTACTO

albertox35@yahoo.com.mx

+52 55 1350 7463

github.com/albertox35

linkedin.com/in/alberto-gertzain

CDMX • Híbrido / Remoto

Disponibilidad inmediata

FORTALEZAS SENIOR

- Ownership end-to-end: diseño, desarrollo, despliegue y operación.
- Liderazgo técnico, coordinación ágil y comunicación con negocio y auditoría.
- Priorización por impacto y riesgo; mejora continua y automatización.

$25M+	+600	95%	37
recuperados	h/mes liberadas	menos errores	soluciones en prod.

Experiencia Profesional

2022 — 2026 Application Manager · Data Engineer

Banamex (Citibanamex) • CDMX

- **Dueño técnico** de 22 soluciones regulatorias en Python y PySpark, del diseño a la operación en producción: **95% de reducción** de errores frente al proceso manual previo
- Diseñé la **arquitectura ETL de Big Data** (normalización hasta 4FN, orquestación en Lockdown Scheduler) que hoy alimenta el reporting regulatorio de la unidad
- Automaticé reportes críticos liberando **+600 horas/mes** de capacidad del equipo
- Rediseñé 4 procesos clave incorporando **modelos predictivos de riesgo crediticio** y alertas automatizadas
- Lideré el equipo en **Scrum** (Jira/Confluence); definí el estándar de documentación técnica para auditoría

2021 — 2022 Analista de Disputas · Data Analytics

HSBC • CDMX

- Recuperé **$25M+** mediante análisis de patrones de disputas y **detección de fraude**
- Construí 15 automatizaciones (VBA/Alteryx) que elevaron la eficiencia del área **45%**
- Diseñé el motor de **cálculo de interés devengado para CNR**
- Implementé dashboards en Power BI/Tableau para monitoreo de KPI en tiempo real

Lo que traigo a la mesa

- **Data Science aplicada**: modelos predictivos, validación, explicabilidad y gobernanza
- **IA agéntica**: agentes propios con RAG local y LLMs on-premise en uso real
- **De notebook a producción**: Docker, FastAPI, CI/CD, monitorización de deriva y retraining
- Enfoque en **P&L**: ahorro operativo medible y recuperación de cartera
- **Escala real**: PySpark y arquitecturas ETL sobre Big Data bancario

Formación

2021–2022 • Dipl. Finanzas Computacionales
UNAM

2016–2021 • Actuaría (Lic.)
UNAM — FES Acatlán

Bach. Tecn. Programación
Plantel Fundación Azteca

Cursos y certificaciones

✓ **Data Engineering con Python** · 2025 ✓ **Automatización y APIs REST** · 2024 ✓ **Modelado financiero y riesgo crediticio** · 2024 ✓ **SAS Programming** · 2020 ✓ **Prog. Estadística con R** · UNAM 2020 ✓ **IA para todos** · Deeplearning.ai 2020 ✓ **Pensamiento Sistémico** · UNAM 2020 ✓ **Scrum Master** · Udemy 2024 ✓ **Jira Ágil** · Udemy 2024 ✓ **Lean Startup** · Udemy 2025 ✓ **Liderazgo Estratégico** · Finnse 2024 ✓ **Competencias Digitales O365** · UAB 2020 ✓ **Seminario Retiro y Salud** · CONAC/AMA 2019 ✓ **CCNA 1: Network Fundamentals** · Cisco 2013 ✓ **CCNA 2: Routing Protocols** · Cisco 2014 **Time Series for Actuaries** · 92% **BI y DAX con Power Pivot** · 91%

Data Science y Riesgo

credit-portfolio-stress-testing

Riesgo de crédito • PySpark • MLflow

- Monte Carlo de escenarios macro
- Pipelines PySpark/MLflow
- Monitoreo de deriva y retraining automático

claims-fraud-detection

Fraude • scikit-learn • FastAPI

- Clasificación supervisada de fraude
- Explicabilidad SHAP por caso
- Scoring en FastAPI y Docker

portfolio-risk-lab

Riesgo de mercado • Gobernance

- VaR/CVaR y sensibilidad de portafolios
- Backtesting de pérdidas
- Scoring en contenedores

Alberto Gertzain Sánchez González

Senior Data & AI ▪ Data Science ▪ Tech Lead ▪ Automation

PERFILES SENIOR

Data Science / Applied ML: modelado predictivo, validación temporal y SHAP para scoring, riesgo y fraude.

Data Engineering / Analytics: arquitectura, ETL con PySpark, calidad e integración a escala.

IA y automatización: Ollama, RAG, salidas estructuradas, agentes y workflows con n8n.

Technical Lead: ownership de soluciones críticas, estándares de auditoría y coordinación transversal.

RESULTADOS COMPROBABLES

$25M+ recuperados; **+600 h/mes** liberadas; **95%** menos errores regulatorios; **45%** más eficiencia operativa.

LIDERAZGO Y ENTREGA

Traduzco objetivos de negocio y regulación a soluciones técnicas; alineo equipos y prioridades; entrego con trazabilidad, calidad y foco en resultados.

MÉTODOS DE DATA SCIENCE

Regresión y clasificación • árboles y boosting • series de tiempo • segmentación • selección de variables • validación temporal • tuning con Optuna • explicabilidad con SHAP.

CALIDAD Y GOBIERNO

Definición de métricas y criterios de aceptación • control de calidad de datos • reproducibilidad • versionado • monitoreo de desempeño y deriva • documentación para auditoría.

COLABORACIÓN TRANSVERSAL

Trabajo con áreas de negocio, riesgo, operaciones, tecnología y auditoría para acordar requerimientos, criterios de éxito y controles de operación.

ÁMBITOS DE NEGOCIO

Banca y servicios financieros • riesgo crediticio • fraude y disputas • reporting regulatorio • seguros y tarificación • operaciones.

GESTIÓN DE INICIATIVAS

Descomposición de problemas complejos en entregables; priorización de alcance y dependencias; seguimiento de riesgos, acuerdos y criterios de aceptación.

COMUNICACIÓN EJECUTIVA

Presento hallazgos, impacto, supuestos y limitaciones en lenguaje claro para facilitar decisiones de negocio y revisiones técnicas o de auditoría.

FORMA DE TRABAJO

Colaboración ágil con Jira y Confluence • documentación y transferencia de conocimiento • mejora continua • enfoque en servicio y resultados.

Ingeniería de datos y ML aplicado

Otros proyectos de ingeniería de datos y machine learning.

financial-data-etl

Data Engineering • Airflow

- Ingesta y limpieza de series financieras
- Airflow con reintentos
- Pruebas de transformación

insurance-pricing-glm

Actuarial • GLM • statsmodels

- Tarificación y severidad GLM
- Segmentación y regularización
- Documentación reproducible

football-match-prediction-engine

ML aplicado • Automatización de workflows

- XGBoost/scikit-learn y Optuna
- Backtesting temporal
- FastAPI, Docker y retraining

IA agéntica — en producción para uso propio

agente-generador-documentos

RAG local • Ollama • 100% on-premise

- Genera documentos desde PDF en local
- RAG propio, embeddings incrementales y OCR de respaldo
- JSON de esquema fijo, validado y con reintentos
- Informa los campos sin fuente

agente-whatsapp-crm

Agente conversacional • CRM

- Agente local para WhatsApp de negocio
- Actualiza el CRM desde la conversación

agente-operaciones-sheets

Automatización remota

- Configuración remota desde Google Sheets
- Automatiza procesos sin acceso directo al equipo

Ciclo de trabajo en Data Science

1. Problema y datos

Traduzco la necesidad de negocio a una métrica; perfilo fuentes, calidad, sesgos y variables disponibles.

2. Modelo y validación

Construyo una línea base, comparo modelos y valido con esquemas adecuados al tiempo y al costo del error.

3. Uso y seguimiento

Explico resultados, documento decisiones y preparo scoring, monitoreo y revisión periódica del desempeño.

Aplicaciones de analítica y modelado

Riesgo y fraude

Scoring, segmentación, alertas y priorización de casos con criterios interpretables.

Planeación y seguimiento

Pronósticos, escenarios y métricas para anticipar desviaciones y respaldar decisiones.

Eficiencia operativa

Automatización de tareas, reducción de errores y medición de capacidad liberada.

Impacto Data Science y automatización

Modelos de riesgo incorporados a 4 procesos; 15 automatizaciones en HSBC; $25M+ recuperados y 45% más eficiencia operativa.

Objetivo

Disponible para roles **Senior / Lead en Data Science, Data Engineering, ML/IA y analítica**. Diseño y despliego modelos confiables y lidero soluciones con foco en resultados de negocio.

Experiencia en banca, fintech, seguros, tecnología y operaciones reguladas.