

## **SERGIO LUIZ DOS SANTOS**

**Científico de Datos | Productos Analíticos Orientados al Riesgo | Machine Learning en Producción**

Brasil

Email: [santossergiorealbjj@outlook.com](mailto:santossergiorealbjj@outlook.com)

Teléfono: +55 31 99370-8399

LinkedIn: <https://www.linkedin.com/in/santossergioluiz>

Portafolio: <https://portfoliosantossergio.vercel.app>

### **PERFIL PROFESIONAL**

Científico de Datos con más de 20 años de experiencia en sistemas regulados y entornos de misión crítica. Enfocado en el desarrollo de productos analíticos orientados a la reducción de riesgo y generación de impacto financiero medible.

Experiencia integral en el ciclo completo de Machine Learning: definición del problema, análisis exploratorio (EDA), ingeniería de características, entrenamiento y validación de modelos, despliegue en producción y monitoreo continuo.

Capacidad comprobada para traducir métricas estadísticas en decisiones estratégicas de negocio, especialmente en entornos financieros y regulados.

Inglés: Intermedio-Avanzado (lectura técnica avanzada).

### **COMPETENCIAS TÉCNICAS**

#### **Lenguajes**

Python | SQL | C# | .NET

#### **Ciencia de Datos y Machine Learning**

Pandas | NumPy | Scikit-learn  
Modelos de Regresión y Árboles  
Gradient Boosting (XGBoost, LightGBM)  
Validación cruzada  
Prevención de overfitting y data leakage  
Evaluación con MAE, RMSE,  $R^2$   
Ingeniería de características  
Fundamentos de inferencia causal

### **Infraestructura y Producción**

Azure Databricks | Docker | Supabase | Render  
FastAPI | Streamlit  
Windows Server | Linux

### **Visualización**

Power BI | Matplotlib | Seaborn

### **Gobernanza**

Trazabilidad de datos  
Control de acceso  
Documentación técnica  
Trabajo en entornos regulados

### **PROYECTO DESTACADO**

#### **Plataforma Predictiva de Riesgo Operacional**

Desarrollo de solución predictiva para anticipar retrasos operativos y reducir exposición financiera.

## **Problema**

35% de operaciones con retrasos, generando penalizaciones contractuales y costos indirectos.

Decisiones tomadas de forma reactiva.

## **Estrategia**

- Arquitectura de datos en capas (raw → analytics → products)
- Ingeniería de características orientada a riesgo

Comparación entre Regresión Lineal y RandomForest

- Validación cruzada para garantizar robustez
- Evaluación de trade-offs entre interpretabilidad y desempeño

RandomForest fue seleccionado por su robustez ante variables heterogéneas y datos ruidosos.

## **Resultado Técnico**

MAE reducido de 12 días a 4,97 días

Reducción del error: 59%

## **Impacto de Negocio**

Reducción estimada de incertidumbre: ~60%

Impacto financiero anual estimado: ≈ R\$ 248.400

Transición estratégica: gestión reactiva → gestión preventiva

## **Producción**

- API REST con FastAPI
- Modelo contenerizado con Docker
- Versionado y prácticas de reproducibilidad
- Bot de consultas operativas

- Simulador ejecutivo en Streamlit

Enfoque en trazabilidad, gobernanza y usabilidad para stakeholders no técnicos.

## **EXPERIENCIA PROFESIONAL**

### **Científico de Datos – Proyectos Aplicados**

Independiente | 2023 – Actualidad

- Desarrollo end-to-end de modelos predictivos
- Despliegue en producción con enfoque en reproducibilidad
- Traducción de desempeño estadístico en impacto financiero
- Documentación alineada a gobernanza de datos
- Evaluación de riesgos y limitaciones de modelos

### **Consultor en Sistemas Críticos e Infraestructura**

Independiente | 2008 – Actualidad

- Administración de servidores Windows y Linux
- Integración de sistemas legados con soluciones modernas
- Trabajo en entornos regulados y sensibles al cumplimiento
- Soporte a áreas legales y administrativas

### **Banco Bradesco S.A.**

Administrativo → Programador → Analista de Sistemas | 1985 – 2008

- Automatización de procesos (≈ 2.920 horas/año eliminadas)
- Soporte a sistemas para más de 500 usuarios con 99,5% de disponibilidad
- Desarrollo en entorno bancario altamente regulado

- Foco en trazabilidad, confiabilidad y cumplimiento normativo

## **EDUCACIÓN Y CERTIFICACIONES**

Data Science con Python – 2025

Azure Databricks – Microsoft – 2025

Neo4J y Análisis de Grafos – 2025

IBM AI Fundamentals – 2024

Microsoft Azure Bootcamp – 2024

Power BI y Excel Avanzado – 2024

## **STACK ACTUAL**

Python | SQL | Scikit-learn | XGBoost | Azure Databricks | Docker | Power BI |  
Streamlit | Supabase