

Sistema de Riesgo Crediticio con RBM

Manual de Usuario

Autores:

Andrés Fernando Gómez Rojas

Carlos Andrés Gómez Vasco

Versión: 1.0.0

2025

Índice

| | |
|---|----------|
| 1. Introducción | 5 |
| 1.1. ¿Qué es este sistema? | 5 |
| 1.1.1. Características principales: | 5 |
| 1.2. ¿Para quién es este sistema? | 5 |
| 2. Flujo de Trabajo Recomendado | 5 |
| 3. Requisitos Previos | 7 |
| 3.1. Conocimientos Básicos | 7 |
| 3.2. Requisitos Técnicos | 7 |
| 3.3. Formatos de Archivo Soportados | 7 |
| 4. Convenciones de este Manual | 8 |
| 4.1. Elementos de Interfaz | 8 |
| 4.2. Código y Comandos | 8 |
| 4.3. Valores y Parámetros | 8 |
| 4.4. Iconos y Símbolos | 8 |
| 5. Guías por Módulo | 9 |
| 5.1. 1. Página de Inicio y Navegación | 9 |
| 5.1.1. Objetivo del Módulo | 9 |
| 5.1.2. Elementos Principales | 9 |
| 5.1.3. Navegación Básica | 9 |
| 5.2. 2. Generar Datos Sintéticos | 9 |
| 5.2.1. Objetivo del Módulo | 9 |
| 5.2.2. Parámetros Configurables | 10 |
| 5.2.3. Proceso de Generación | 10 |

| | | |
|--------|---|----|
| 5.2.4. | Variables Generadas | 10 |
| 5.3. | 3. Cargar Datos Externos | 10 |
| 5.3.1. | Objetivo del Módulo | 10 |
| 5.3.2. | Formatos Soportados | 11 |
| 5.3.3. | Proceso de Carga | 11 |
| 5.3.4. | Validaciones Automáticas | 11 |
| 5.4. | 4. Análisis Descriptivo | 12 |
| 5.4.1. | Objetivo del Módulo | 12 |
| 5.4.2. | Análisis Univariado | 12 |
| 5.4.3. | Análisis Bivariado | 12 |
| 5.4.4. | Visualizaciones Disponibles | 12 |
| 5.5. | 5. Ingeniería de Características | 13 |
| 5.5.1. | Objetivo del Módulo | 13 |
| 5.5.2. | Variables Derivadas Automáticas | 13 |
| 5.5.3. | Transformaciones Aplicadas | 13 |
| 5.5.4. | Selección de Características | 13 |
| 5.5.5. | Proceso Recomendado | 13 |
| 5.6. | 6. Clustering y Segmentación | 14 |
| 5.6.1. | Objetivo del Módulo | 14 |
| 5.6.2. | Algoritmos Disponibles | 14 |
| 5.6.3. | Determinación de Número de Clusters | 14 |
| 5.6.4. | Análisis de Perfiles | 14 |
| 5.7. | 7. Máquina de Boltzmann Restringida (RBM) | 15 |
| 5.7.1. | ¿Qué es una RBM? | 15 |
| 5.7.2. | Hiperparámetros Configurables | 15 |
| 5.7.3. | Proceso de Entrenamiento | 15 |
| 5.7.4. | Métricas de Evaluación | 15 |

| | | |
|---------|---|----|
| 5.7.5. | Visualizaciones | 16 |
| 5.8. | 8. Modelos Supervisados | 16 |
| 5.8.1. | Objetivo del Módulo | 16 |
| 5.8.2. | Algoritmos Disponibles | 16 |
| 5.8.3. | Métricas de Evaluación | 17 |
| 5.8.4. | Validación Cruzada | 17 |
| 5.8.5. | Optimización de Hiperparámetros | 17 |
| 5.9. | 9. Predicción de Riesgo | 17 |
| 5.9.1. | Objetivo del Módulo | 17 |
| 5.9.2. | Predicción Individual | 18 |
| 5.9.3. | Formulario de Entrada | 18 |
| 5.9.4. | Resultados de Predicción | 18 |
| 5.9.5. | Predicción Batch | 18 |
| 5.10. | 10. Reentrenamiento de Modelos | 19 |
| 5.10.1. | Objetivo del Módulo | 19 |
| 5.10.2. | ¿Por Qué Reentrenar? | 19 |
| 5.10.3. | Detección de Drift | 19 |
| 5.10.4. | Estrategias de Reentrenamiento | 19 |
| 5.10.5. | Versionado y Rollback | 19 |
| 5.11. | 11. Sistema RAG Educativo | 20 |
| 5.11.1. | ¿Qué es RAG? | 20 |
| 5.11.2. | Base de Conocimiento | 20 |
| 5.11.3. | Uso del Chat Educativo | 20 |
| 5.11.4. | Preguntas Sugeridas | 20 |
| 5.11.5. | Características del Sistema | 21 |

| | |
|--|-----------|
| 6.1. Primeros Pasos | 21 |
| 6.2. Análisis de Datos | 21 |
| 6.3. Modelado Avanzado | 21 |
| 6.4. Producción | 21 |
| 6.5. Aprendizaje | 22 |
| 7. Soporte y Recursos | 22 |
| 7.1. Documentación Adicional | 22 |
| 7.2. Recursos de Aprendizaje | 22 |
| 7.3. Soporte Técnico | 22 |
| 8. Próximos Pasos | 23 |
| 8.1. Para Nuevos Usuarios | 23 |
| 8.2. Para Usuarios Intermedios | 23 |
| 8.3. Para Usuarios Avanzados | 23 |
| 8.4. Recursos de Formación | 23 |
| 9. Glosario | 24 |
| 10. Información de Contacto | 24 |
| 10.1. Desarrolladores | 24 |
| 10.2. Correo Electrónico | 24 |
| 10.3. Repositorio del Proyecto | 25 |
| 10.4. App desplegada | 25 |
| 10.5. Documentacion | 25 |
| 10.6. Licencia | 25 |
| 10.7. Versión | 25 |

1 Introducción

1.1 ¿Qué es este sistema?

Este sistema integral permite analizar y predecir el riesgo crediticio hipotecario utilizando técnicas avanzadas de Machine Learning, específicamente **Máquinas de Boltzmann Restringidas (RBM)** combinadas con modelos supervisados.

Nota

Este manual está diseñado para usuarios no técnicos. Si buscas documentación técnica del código, consulta la documentación de la API.

1.1.1 Características principales:

- **Generación de datos sintéticos** realistas para Colombia
- **Carga y validación** de datos externos
- **Análisis exploratorio** completo
- **Ingeniería de características** automática
- **Entrenamiento de RBM** para extracción de características
- **Múltiples modelos** de clasificación supervisados
- **Sistema de predicción** en tiempo real
- **Reentrenamiento automático** de modelos
- **Asistente educativo RAG** con papers científicos

1.2 ¿Para quién es este sistema?

Estudiantes de Física: Aprende sobre Máquinas de Boltzmann y sus aplicaciones en finanzas mediante el sistema RAG educativo.

2 Flujo de Trabajo Recomendado

Para obtener los mejores resultados, sigue este flujo de trabajo estructurado:

1. Generar o Cargar Datos

- Comienza generando un dataset sintético o cargando tus propios datos de crédito
- Asegúrate de que los datos tengan la estructura requerida

2. Validar y Limpiar

- Utiliza las herramientas de validación automática
- Corrige valores atípicos y datos faltantes
- Genera reportes de calidad

3. Análisis Exploratorio

- Comprende las distribuciones de las variables
- Identifica correlaciones y patrones
- Visualiza los datos con gráficos interactivos

4. Ingeniería de Características

- Crea variables derivadas automáticamente
- Aplica transformaciones necesarias
- Selecciona las características más relevantes

5. Entrenar RBM

- Configura los hiperparámetros de la RBM
- Extrae características latentes
- Monitorea el proceso de entrenamiento

6. Entrenar Modelos Supervisados

- Selecciona entre múltiples algoritmos
- Compara el rendimiento de los modelos
- Optimiza hiperparámetros

7. Realizar Predicciones

- Evalúa el riesgo de nuevos solicitantes
- Interpreta los resultados con explicaciones
- Genera reportes detallados

8. Reentrenar Modelos

- Actualiza los modelos con nuevos datos
- Monitorea el drift de datos
- Mantén la precisión predictiva

9. Aprender sobre RBMs

- Utiliza el asistente RAG educativo
- Consulta papers científicos
- Profundiza en la teoría

3 Requisitos Previos

3.1 Conocimientos Básicos

- Conceptos financieros básicos (crédito, tasa de interés, score crediticio)
- Familiaridad con navegadores web y aplicaciones web
- Comprensión básica de estadística (opcional pero útil)
- Conocimientos elementales de machine learning (deseable)

3.2 Requisitos Técnicos

- **Navegador web moderno:** Chrome 90+, Firefox 88+, Safari 14+, Edge 90+
- **Conexión a internet:** Requerida para el sistema RAG educativo
- **Resolución mínima:** 1280x720 píxeles
- **JavaScript:** Debe estar habilitado
- **Cookies:** Requeridas para mantener la sesión

3.3 Formatos de Archivo Soportados

- **CSV:** Archivos separados por comas
- **Excel:** Formatos .xlsx y .xls
- **Parquet:** Formato columnar eficiente
- **JSON:** Para datos estructurados

Advertencia

Asegúrate de que los archivos no excedan el límite de tamaño de 100MB y que tengan la codificación UTF-8 para caracteres especiales.

4 Convenciones de este Manual

4.1 Elementos de Interfaz

- Los **botones** se muestran como: [Generar Datos]
- Los **campos de entrada** se muestran como: *Número de registros*
- Los **menús** se muestran como: **Sidebar** → **Generar Datos**
- Los **valores predeterminados** se muestran como: `valor = 1000`

4.2 Código y Comandos

Los comandos de terminal se muestran en recuadros de código:

```
1 streamlit run app.py
```

Listing 1: Ejecutar la aplicación

4.3 Valores y Parámetros

- Los **valores numéricos** se muestran como: 100
- Los **nombres de variables** se muestran como: `nivel_riesgo`
- Los **archivos** se muestran como: `datos_credito.csv`
- Los **directorios** se muestran como: `src/generar_datos.py`

4.4 Iconos y Símbolos

- Análisis de datos
- Procesamiento rápido
- Configuración
- Inteligencia artificial
- Visualización
- Predicción
- Educación
- Advertencia

5 Guías por Módulo

5.1 1. Página de Inicio y Navegación

5.1.1 Objetivo del Módulo

Familiarizarse con la interfaz principal, aprender a navegar entre módulos y comprender la estructura general de la aplicación.

5.1.2 Elementos Principales

- **Barra lateral:** Navegación principal entre módulos
- **Área de contenido:** Visualización de resultados y controles
- **Header:** Información del sistema y estado
- **Footer:** Información de versión y autores

5.1.3 Navegación Básica

1. Abre tu navegador y ve a: `http://localhost:8501`
2. En la barra lateral, selecciona el módulo deseado
3. Utiliza los controles deslizantes y botones para interactuar
4. Los resultados se actualizan automáticamente

Consejo

Usa el botón R.^{en} la esquina superior derecha para recargar la aplicación si experimentas problemas de rendimiento.

5.2 2. Generar Datos Sintéticos

5.2.1 Objetivo del Módulo

Crear datasets sintéticos realistas para Colombia que simulen solicitudes de crédito hipotecario.

5.2.2 Parámetros Configurables

- **Número de registros:** 1,000 - 50,000 (valor recomendado: 10,000)
- **Semilla aleatoria:** Para reproducibilidad (valor predeterminado: 42)
- **Distribución de riesgo:** Bajo (60 %), Medio (25 %), Alto (15 %)
- **Exportar formatos:** CSV, Excel, Parquet

5.2.3 Proceso de Generación

1. Accede al módulo: **Sidebar** → **Generar Datos**
2. Configura los parámetros deseados
3. Haz clic en [Generar Dataset]
4. Espera la confirmación de generación
5. Descarga los archivos generados

5.2.4 Variables Generadas

| Variable | Tipo | Rango |
|---------------------|------------|------------------|
| valor_inmueble | Numérico | 50M - 2,000M COP |
| monto_credito | Numérico | 20M - 1,500M COP |
| puntaje_datacredito | Numérico | 150 - 950 |
| salario_mensual | Numérico | 1M - 50M COP |
| nivel_riesgo | Categorico | Bajo/Medio/Alto |

Cuadro 1: Variables principales del dataset sintético

Importante

Los datos sintéticos incluyen correlaciones realistas entre variables y validaciones de consistencia para simular escenarios del mundo real.

5.3 3. Cargar Datos Externos

5.3.1 Objetivo del Módulo

Cargar, validar y preparar datasets externos para su análisis en el sistema.

5.3.2 Formatos Soportados

- **CSV:** Separado por comas, punto y coma o tabs
- **Excel:** Hojas de cálculo (.xlsx, .xls)
- **Parquet:** Formato columnar eficiente
- **JSON:** Datos estructurados

5.3.3 Proceso de Carga

1. Accede al módulo: **Sidebar** → **Cargar Datos**
2. Selecciona [Seleccionar archivo]
3. Elige el archivo desde tu dispositivo
4. Espera la validación automática
5. Revisa el reporte de calidad
6. Confirma la carga con [Usar este Dataset]

5.3.4 Validaciones Automáticas

El sistema realiza las siguientes validaciones:

- **Estructura:** Columnas requeridas presentes
- **Tipos de datos:** Coherencia en tipos de variables
- **Rangos:** Valores dentro de límites razonables
- **Valores nulos:** Identificación de datos faltantes
- **Duplicados:** Detección de registros repetidos
- **Consistencia:** Relaciones lógicas entre variables

Advertencia

Si el archivo no pasa las validaciones, el sistema proporcionará un reporte detallado de los problemas encontrados y sugerencias para corregirlos.

5.4 4. Análisis Descriptivo

5.4.1 Objetivo del Módulo

Realizar análisis exploratorio de datos para comprender distribuciones, correlaciones y patrones en los datos de crédito.

5.4.2 Análisis Univariado

- **Estadísticas descriptivas:** Media, mediana, desviación estándar
- **Visualizaciones:** Histogramas, boxplots, gráficos de densidad
- **Tests de normalidad:** Shapiro-Wilk, Kolmogorov-Smirnov
- **Análisis de outliers:** Detección y cuantificación

5.4.3 Análisis Bivariado

- **Matrices de correlación:** Pearson, Spearman, Kendall
- **Gráficos de dispersión:** Relaciones entre variables numéricas
- **Tablas de contingencia:** Relaciones entre variables categóricas
- **ANOVA y tests Chi-cuadrado:** Significancia estadística

5.4.4 Visualizaciones Disponibles

- **Heatmaps:** Correlaciones entre variables
- **Scatter plots:** Relaciones bivariadas
- **Bar charts:** Distribuciones categóricas
- **Box plots:** Distribuciones y outliers
- **Violin plots:** Densidades de distribución
- **Q-Q plots:** Comparación con distribuciones teóricas

Consejo

Utiliza el filtrado interactivo para explorar subconjuntos específicos de datos y comparar diferentes segmentos de solicitantes.

5.5 5. Ingeniería de Características

5.5.1 Objetivo del Módulo

Crear, transformar y seleccionar características que mejoren el poder predictivo de los modelos de machine learning.

5.5.2 Variables Derivadas Automáticas

- **Ratios financieros:** LTV, DTI, capacidad de ahorro
- **Indicadores de riesgo:** Scores de edad, estabilidad, situación legal
- **Interacciones:** Educación×Salario, Edad×Tiempo empleo
- **Agregaciones:** Promedios móviles, tasas de cambio

5.5.3 Transformaciones Aplicadas

- **Escalado:** Min-Max, Standard, Robust
- **Normalización:** Log, raíz cuadrada, Box-Cox
- **Codificación:** One-Hot, Label, Target encoding
- **Discretización:** Binning equal-width, equal-frequency

5.5.4 Selección de Características

- **Métodos filter:** Correlación, mutual information
- **Métodos wrapper:** Forward selection, backward elimination
- **Métodos embedded:** Lasso, Random Forest importance
- **Reducción dimensionalidad:** PCA, t-SNE, UMAP

5.5.5 Proceso Recomendado

1. Generar todas las características derivadas
2. Evaluar la importancia de cada característica
3. Seleccionar las más relevantes

4. Validar con múltiples modelos
5. Iterar y refinar

Importante

Guarda el pipeline de transformaciones para aplicarlo consistentemente a nuevos datos durante la predicción.

5.6 6. Clustering y Segmentación

5.6.1 Objetivo del Módulo

Identificar grupos homogéneos de solicitantes de crédito mediante algoritmos de clustering no supervisado.

5.6.2 Algoritmos Disponibles

- **K-Means:** Rápido y eficiente para clusters esféricos
- **DBSCAN:** Detecta clusters de forma arbitraria y maneja outliers
- **Jerárquico:** Proporciona dendrogramas para análisis visual
- **Gaussian Mixture:** Modelo probabilístico para clusters superpuestos

5.6.3 Determinación de Número de Clusters

- **Método del codo:** Suma de cuadrados intra-cluster
- **Silhouette score:** Cohesión y separación de clusters
- **Gap statistic:** Comparación con distribución de referencia
- **Análisis de estabilidad:** Consistencia en múltiples ejecuciones

5.6.4 Análisis de Perfiles

- **Características por cluster:** Medias y distribuciones
- **Perfiles de riesgo:** Comportamiento crediticio por segmento
- **Visualización:** PCA 2D/3D, t-SNE, UMAP

- **Interpretación:** Significado comercial de cada cluster

Consejo

Combina el clustering con el análisis de riesgo para identificar segmentos de alto riesgo que requieran políticas crediticias específicas.

5.7 7. Máquina de Boltzmann Restringida (RBM)

5.7.1 ¿Qué es una RBM?

Una Máquina de Boltzmann Restringida es un modelo generativo no supervisado que aprende representaciones latentes de los datos mediante una arquitectura de dos capas (visible + oculta).

5.7.2 Hiperparámetros Configurables

- **Unidades ocultas:** 50 - 500 (recomendado: 100-200)
- **Tasa de aprendizaje:** 0.001 - 0.1 (recomendado: 0.01)
- **Épocas:** 50 - 500 (recomendado: 100-200)
- **Batch size:** 32 - 256 (recomendado: 64)
- **Pasos CD-k:** 1 - 10 (recomendado: 1)

5.7.3 Proceso de Entrenamiento

1. Preparar los datos (normalizar y binarizar si es necesario)
2. Configurar los hiperparámetros de la RBM
3. Inicializar pesos y sesgos aleatoriamente
4. Ejecutar Contrastive Divergence (CD-k)
5. Monitorear error de reconstrucción
6. Extraer características latentes

5.7.4 Métricas de Evaluación

- **Error de reconstrucción:** Disminución durante el entrenamiento

- **Log-verosimilitud:** Medida de ajuste del modelo
- **Características aprendidas:** Interpretación de unidades ocultas
- **Calidad de generación:** Muestras generadas por la RBM

5.7.5 Visualizaciones

- **Evolución del error:** Curva de aprendizaje
- **Matriz de pesos:** Patrones aprendidos
- **Activaciones:** Distribución de unidades ocultas
- **Reconstrucciones:** Comparación datos originales vs reconstruidos

Importante

Las características extraídas por la RBM pueden utilizarse como input para modelos supervisados, mejorando significativamente su rendimiento predictivo.

5.8 8. Modelos Supervisados

5.8.1 Objetivo del Módulo

Entrenar, evaluar y comparar múltiples algoritmos de clasificación para predecir el riesgo crediticio.

5.8.2 Algoritmos Disponibles

- **Regresión Logística:** Modelo lineal interpretable
- **Random Forest:** Ensemble robusto y preciso
- **XGBoost:** Gradient boosting de alto rendimiento
- **LightGBM:** Gradient boosting eficiente
- **SVM:** Máquinas de vectores de soporte
- **Red Neuronal:** Perceptrón multicapa

5.8.3 Métricas de Evaluación

- **Accuracy:** Precisión general
- **Precision:** Capacidad de no falsos positivos
- **Recall:** Capacidad de encontrar todos los positivos
- **F1-Score:** Media armónica de precision y recall
- **ROC-AUC:** Capacidad discriminativa general
- **Matriz de confusión:** Errores por clase

5.8.4 Validación Cruzada

- **K-Fold:** 5 o 10 folds para estimación robusta
- **Stratified:** Mantiene proporción de clases
- **Repeated:** Múltiples ejecuciones para estabilidad
- **Time Series Split:** Para datos temporales

5.8.5 Optimización de Hiperparámetros

- **Grid Search:** Búsqueda exhaustiva en grid
- **Random Search:** Muestreo aleatorio eficiente
- **Bayesian Optimization:** Búsqueda inteligente
- **Genetic Algorithms:** Algoritmos evolutivos

Consejo

Utiliza la validación cruzada anidada para obtener una estimación no sesgada del rendimiento del modelo con optimización de hiperparámetros.

5.9 9. Predicción de Riesgo

5.9.1 Objetivo del Módulo

Evaluar el riesgo crediticio de nuevos solicitantes utilizando los modelos entrenados y proporcionar explicaciones interpretables.

5.9.2 Predicción Individual

1. Accede al módulo: **Sidebar** → **Predicción**
2. Completa el formulario con los datos del solicitante
3. Selecciona el modelo a utilizar
4. Haz clic en [Evaluar Riesgo]
5. Revisa el resultado y las explicaciones

5.9.3 Formulario de Entrada

- **Datos personales:** Edad, educación, estado civil
- **Información financiera:** Ingresos, gastos, patrimonio
- **Historial crediticio:** Score, historial de pagos
- **Características del crédito:** Monto, plazo, garantías

5.9.4 Resultados de Predicción

- **Nivel de riesgo:** Bajo/Medio/Alto con probabilidad
- **Explicaciones:** Factores que influyeron en la decisión
- **Recomendaciones:** Acciones sugeridas
- **Límites de confianza:** Incertidumbre de la predicción

5.9.5 Predicción Batch

- **Carga múltiple:** Archivo con múltiples solicitantes
- **Procesamiento:** Evaluación en lote eficiente
- **Reporte:** Resumen estadístico de predicciones
- **Exportación:** Resultados en formato CSV/Excel

Importante

Las explicaciones del modelo ayudan a cumplir con requisitos regulatorios y construyen confianza en las decisiones automatizadas.

5.10 10. Reentrenamiento de Modelos

5.10.1 Objetivo del Módulo

Monitorear el rendimiento de los modelos en producción y reentrenarlos cuando sea necesario para mantener su precisión predictiva.

5.10.2 ¿Por Qué Reentrenar?

- **Drift de datos:** Cambios en las distribuciones de entrada
- **Drift de concepto:** Cambios en la relación entrada-salida
- **Degradación:** Disminución gradual del rendimiento
- **Nuevos patrones:** Comportamientos no vistos durante entrenamiento

5.10.3 Detección de Drift

- **Drift de datos:** KS test, PSI (Population Stability Index)
- **Drift de concepto:** Monitoring accuracy, performance metrics
- **Alerta temprana:** Cambios significativos en distribuciones
- **Dashboards:** Visualización de métricas en tiempo real

5.10.4 Estrategias de Reentrenamiento

- **Reentrenamiento completo:** Con todos los datos disponibles
- **Incremental:** Solo con datos nuevos
- **Transfer learning:** Ajuste fino de modelos existentes
- **Ensemble:** Combinación de modelos viejos y nuevos

5.10.5 Versionado y Rollback

- **Control de versiones:** Historial completo de modelos
- **Comparación:** Métricas entre versiones
- **Rollback:** Volver a versión anterior si es necesario

- **A/B testing:** Comparación en producción

Advertencia

Establece protocolos claros para el reentrenamiento y despliegue de modelos, incluyendo validación exhaustiva antes de ponerlos en producción.

5.11 11. Sistema RAG Educativo

5.11.1 ¿Qué es RAG?

Retrieval-Augmented Generation (RAG) es una técnica que combina recuperación de información con generación de texto, permitiendo respuestas basadas en fuentes específicas.

5.11.2 Base de Conocimiento

- **Papers científicos:** Investigaciones sobre RBMs y riesgo crediticio
- **Libros técnicos:** Fundamentos teóricos y aplicaciones
- **Tutoriales:** Guías prácticas de implementación
- **Casos de estudio:** Aplicaciones en la industria financiera

5.11.3 Uso del Chat Educativo

1. Accede al módulo: **Sidebar** → **RAG Educativo**
2. Escribe tu pregunta en el campo de texto
3. Selecciona el nivel de profundidad deseado
4. Haz clic en [Enviar Pregunta]
5. Revisa la respuesta y las fuentes citadas

5.11.4 Preguntas Sugeridas

- **Nivel básico:** "¿Qué es una Máquina de Boltzmann?"
- **Nivel intermedio:** "¿Cómo se entrena una RBM con Contrastive Divergence?"
- **Nivel avanzado:** "¿Cuáles son las aplicaciones de RBMs en riesgo crediticio?"
- **Implementación:** "¿Cómo configurar los hiperparámetros de una RBM?"

5.11.5 Características del Sistema

- **Búsqueda semántica:** Encuentra información relevante
- **Citación de fuentes:** Referencias a documentos originales
- **Multiple niveles:** Respuestas adaptadas al conocimiento del usuario
- **Actualización:** Base de conocimiento en expansión

Consejo

Utiliza el sistema RAG para profundizar en conceptos específicos antes de implementar configuraciones avanzadas en los módulos de modelado.

6 Acceso Rápido

6.1 Primeros Pasos

- **Familiarízate con la interfaz:** 1. Página de Inicio y Navegación
- **Crea tu primer dataset:** 2. Generar Datos Sintéticos

6.2 Análisis de Datos

- **Trabaja con tus datos:** 3. Cargar Datos Externos
- **Explora y visualiza:** 4. Análisis Descriptivo

6.3 Modelado Avanzado

- **Entrena Máquinas de Boltzmann:** 7. Máquina de Boltzmann Restringida (RBM)
- **Construye clasificadores:** 8. Modelos Supervisados

6.4 Producción

- **Predice riesgo en tiempo real:** 9. Predicción de Riesgo
- **Mantén modelos actualizados:** 10. Reentrenamiento de Modelos

6.5 Aprendizaje

- **Aprende sobre RBMs con IA:** 11. Sistema RAG Educativo

7 Soporte y Recursos

7.1 Documentación Adicional

- **Referencia de API:** Documentación técnica detallada
- **Guía de instalación:** Configuración del sistema
- **Ejemplos de uso:** Casos prácticos implementados
- **FAQ:** Preguntas frecuentes y soluciones

7.2 Recursos de Aprendizaje

- **Tutoriales interactivos:** Guías paso a paso
- **Videos demostrativos:** Explicaciones visuales
- **Paper repository:** Investigaciones científicas
- **Community forum:** Discusiones con otros usuarios

7.3 Soporte Técnico

- **Reportar problemas:** Sistema de tickets
- **Solicitar características:** Mejoras futuras
- **Consultas técnicas:** Asistencia especializada
- **Training:** Sesiones de capacitación

Importante

Antes de contactar al soporte, revisa la sección de troubleshooting del módulo correspondiente y la documentación disponible.

8 Próximos Pasos

8.1 Para Nuevos Usuarios

1. **Completa el tutorial inicial** en la página de inicio
2. **Genera un dataset sintético** y explora sus características
3. **Realiza tu primera predicción** con datos de ejemplo
4. **Consulta el sistema RAG** para aclarar conceptos básicos

8.2 Para Usuarios Intermedios

1. **Experimenta con diferentes configuraciones** de RBM
2. **Compara múltiples algoritmos** de clasificación
3. **Analiza el impacto** de diferentes características
4. **Implementa pipelines completos** de principio a fin

8.3 Para Usuarios Avanzados

1. **Integra con sistemas existentes** mediante la API
2. **Desarrolla modelos personalizados** para casos específicos
3. **Automatiza procesos** de reentrenamiento y monitoreo
4. **Contribuye** con mejoras y extensiones

8.4 Recursos de Formación

- **Curso básico:** Fundamentos de riesgo crediticio y ML
- **Curso avanzado:** RBMs y modelos generativos
- **Workshops:** Sesiones prácticas guiadas
- **Certificaciones:** Validación de conocimientos

9 Glosario

Accuracy Medida de precisión general de un modelo de clasificación.

Batch Size Número de muestras procesadas antes de actualizar los parámetros del modelo.

Características Latentes Representaciones aprendidas automáticamente por modelos no supervisados.

Contrastive Divergence Algoritmo utilizado para entrenar Máquinas de Boltzmann.

Drift de Datos Cambio en la distribución de los datos de entrada con el tiempo.

Época Paso completo a través de todo el dataset durante el entrenamiento.

Feature Engineering Proceso de crear nuevas características a partir de datos existentes.

Hiperparámetros Parámetros de configuración de los algoritmos de machine learning.

LTV (Loan-to-Value) Ratio entre el monto del crédito y el valor del inmueble.

Máquina de Boltzmann Modelo generativo basado en física estadística.

Overfitting Cuando un modelo se ajusta demasiado a los datos de entrenamiento.

RAG (Retrieval-Augmented Generation) Sistema que combina recuperación y generación de información.

Score Crediticio Puntuación que representa la probabilidad de incumplimiento.

Validación Cruzada Técnica para evaluar la generalización de modelos.

XGBoost Algoritmo de gradient boosting optimizado.

10 Información de Contacto

10.1 Desarrolladores

- Andrés Fernando Gómez Rojas
- Carlos Andrés Gómez Vasco

10.2 Correo Electrónico

- fernandogomez0621@gmail.com

10.3 Repositorio del Proyecto

- <https://github.com/fernandogomez0621/AppFIsica>

10.4 App desplegada

- <http://157.137.229.69:1111/>

10.5 Documentacion

- <http://157.137.229.69:2222/>

10.6 Licencia

- Este software se distribuye bajo licencia MIT

10.7 Versión

- **Actual:** 1.0.0
- **Fecha de publicación:** 27 de octubre de 2025
- **Próxima actualización:** Enero 2025

Nota

Este manual se actualiza regularmente. Consulta la versión más reciente en el repositorio del proyecto.