

Sistema de Pronóstico Automatizado de Demanda Energética - API Gateway

1. RESUMEN

El presente proyecto tiene como finalidad desarrollar un **API Gateway inteligente** para el pronóstico automatizado de demanda energética del sistema de distribución de EPM en Antioquia, Colombia, cumpliendo con los requisitos establecidos en el Acuerdo CNO 1303 de 2020 y el proyecto de resolución CREG 143 de 2021.

El sistema implementará **modelos predictivos** basados en inteligencia artificial y machine learning que, de forma autónoma, procesarán datos históricos de demanda, variables climáticas y eventos especiales para generar pronósticos de alta precisión con granularidad horaria y de 15 minutos.

La solución se caracteriza por su **capacidad de autoaprendizaje**: el sistema monitoreará continuamente su desempeño y, al detectar que el MAPE mensual supera el umbral del 5%, ejecutará automáticamente un ciclo completo de reentrenamiento, selección del modelo óptimo y promoción a producción, sin intervención manual.

El entregable principal es un **API REST** que expondrá endpoints para:

- Generación de pronósticos en múltiples horizontes temporales (mensual, semanal, diario, intradiario).
- Consulta de métricas de desempeño en tiempo real.
- Monitoreo del estado de salud del sistema.

Este API será consumido por otros componentes del ecosistema de EPM responsables de la visualización, reportería y carga de información hacia XM.

2. ALCANCE DEL PROYECTO

2.1 INCLUIDO EN EL ALCANCE:

1. Pipeline Automatizado de Datos

- Conectores para lectura automática de datos históricos del sistema de almacenamiento de EPM (SCADA, XM, GCE).
- Integración con fuentes de datos meteorológicos, de no tenerlos almacenados en local.
- Sistema automático de limpieza, validación y transformación de datos.
- Feature engineering automático (variables de demanda, calendario y climáticas).

2. Modelos Predictivos

- Desarrollo e implementación de **modelos de machine learning** para pronóstico de demanda.
- Integración de variables climáticas y tipicidad de días por eventos.
- Sistema de entrenamiento automático con datos actualizados hasta el día anterior.
- Versionado y gestión de modelos entrenados.

3. Sistema de Evaluación y Selección Automática

- Motor de cálculo automático de métricas (MAPE mensual, desviaciones diarias y horarias).
- Algoritmo de selección automática del modelo campeón basado en scoring multi-criterio.
- Promoción automática del mejor modelo a producción.
- Estrategia de fallback y rollback.

4. API Gateway

- Endpoints REST para:
 - **/predict**: Generación de pronósticos con múltiples horizontes temporales y granularidades.
 - **/metrics**: Consulta de métricas de desempeño actuales.
 - **/health**: Estado del sistema y modelos en producción.
 - **/retrain**: Trigger manual de reentrenamiento (Si así lo solicitan).

- Autenticación, rate limiting y logging.
- Documentación completa.

5. Sistema de Monitoreo y Reentrenamiento Automático

- Monitoreo continuo de MAPE mensual (rolling).
- Detección automática de degradación del modelo ($\text{MAPE} > 5\%$).
- Trigger automático de reentrenamiento completo al superar umbral.
- Logging de eventos, anomalías y desviaciones.
- Alertas de degradación del modelo.

6. Cumplimiento de Métricas Regulatorias

- MAPE mensual $< 5\%$.
- Desviaciones diarias $< 5\%$.
- Desviaciones horarias < 60 conteos/mes (con desviación $> 5\%$).
- Pronósticos para 60 barras (número variable durante el contrato).

7. Horizontes y Granularidades de Pronóstico

- **Mensual:** Actualización mensual, con un mes de antelación.
- **Semanal:** Entrega jueves antes de 12pm para semana siguiente (lunes-domingo).
- **Diario:** Entrega diaria 6am para día siguiente.
- **Intradiario:** Actualización 3 veces al día.
- **Granularidades:** Horaria y cada 15 minutos.

2.2 EXCLUIDO DEL ALCANCE:

1. Almacenamiento de Datos Históricos

- El sistema **NO** almacenará ni gestionará bases de datos históricas.
- **Responsabilidad de EPM:** Mantener y exponer datos históricos a través de sus sistemas existentes.

2. Plataforma de Visualización

- El sistema **NO** incluye interfaces gráficas, dashboards ni herramientas de visualización para usuarios finales.
- **Responsabilidad del consumidor del API:** Desarrollar interfaces para analistas de EPM.

3. Reportería y Formatos para XM

- El sistema **NO** generará reportes en formatos específicos de XM.
- **Responsabilidad del consumidor del API:** Transformar respuestas del API a formatos requeridos por XM.

4. Ingesta y Persistencia de Datos en Tiempo Real

- El sistema **NO** capturará ni almacenará datos directamente desde SCADA, XM o GCE.
- **Responsabilidad de EPM:** Proveer acceso a datos ya almacenados en sus sistemas.

5. Análisis de Causalidad para Reportes Ejecutivos

- El sistema **detectará y loggeará** desviaciones y anomalías, pero **NO** generará análisis de causalidad en formato de reporte ejecutivo.
- **Responsabilidad del consumidor del API:** Interpretar logs y métricas para reportes regulatorios.

6. Infraestructura de Despliegue

- El proyecto **NO** incluye provisión de servidores, cloud, ni infraestructura de hosting.
- **Responsabilidad de EPM:** Proveer infraestructura para despliegue del API.

7. Integración Directa con Sistemas de XM

- El sistema **NO** realizará carga directa de pronósticos en plataformas de XM.

Responsabilidad del consumidor del API: Consumir endpoints y cargar datos en sistemas de XM.

3. FASES DEL SISTEMA

3.1 Fase 1: Configuración del Pipeline Automático de Datos

Objetivo: Establecer el sistema automatizado de adquisición, limpieza y preparación de datos para el entrenamiento y predicción continua.

Acciones Clave:

- **Integración con Fuentes de Datos:**
 - Definir y configurar conectores para lectura automática de datos históricos del sistema de almacenamiento de EPM.
 - Integrar fuente de datos meteorológicos (históricos y pronósticos).
 - Establecer contratos de datos y validaciones de esquema.
- **Sistema de Limpieza Automática:**
 - Implementar rutinas automáticas para detección y tratamiento de datos faltantes o de mala calidad.
 - Configurar reglas de validación y transformación de datos en tiempo real.
 - Implementar logging de anomalías y alertas de calidad de datos.
- **Feature Engineering Automático:**
 - Desarrollar pipeline de generación automática de variables predictivas:
 - Variables de demanda (históricos, lags, rolling statistics).
 - Variables calendario (festivos, día de la semana, estacionalidad).
 - Variables climáticas (temperatura, humedad, sensación térmica, correlaciones).
 - Implementar selección automática de features más relevantes.

- **Sistema de Desagregación:**

- Desarrollar metodología automática para desagregar datos horarios a granularidad de 15 minutos.

Entregables:

- Pipeline de datos automatizado y operativo.
- Documentación de contratos de integración.
- Sistema de monitoreo de calidad de datos.

3.2 Fase 2: Desarrollo de Modelos y Sistema de Entrenamiento Automático

Objetivo: Desarrollar los modelos predictivos y el sistema que los entrena automáticamente con datos actualizados.

Acciones Clave:

- **Desarrollo de los Modelos:**

- Implementar arquitectura de los modelos predictivos
- Parametrizar modelos para el mercado MC Antioquia.
- Integrar variables climáticas y tipicidad de días por eventos.

- **Sistema de Entrenamiento Automático:**

- Desarrollar orquestador de entrenamiento que:
 - Obtiene automáticamente datos hasta el día anterior.
 - Ejecuta entrenamiento de los 3 modelos en paralelo.
 - Gestiona versionado de modelos entrenados.
 - Registra métricas y metadatos de cada entrenamiento.

- **Backtesting Automático:**

- Implementar sistema de validación cruzada temporal automática.
- Simular operación en períodos históricos para validar robustez.

Entregables:

- 3 modelos predictivos implementados.
- Sistema de entrenamiento automático funcional.
- Repositorio de versiones de modelos.

3.3 Fase 3: Sistema de Validación y Selección Automática del Modelo

Objetivo: Implementar el sistema que evalúa automáticamente los modelos y selecciona el óptimo para producción.

Acciones Clave:

- **Motor de Evaluación Automática:**
 - Implementar sistema de cálculo automático de métricas:
 - $\text{MAPE mensual} < 5\%$.
 - Desviaciones diarias $< 5\%$.
 - Desviaciones horarias < 60 conteos/mes (con desviación $> 5\%$).
 - Calcular métricas por barra y globales.
- **Algoritmo de Selección del Modelo:**
 - Desarrollar lógica de comparación multi-criterio entre los modelos.
 - Implementar sistema de scoring ponderado (accuracy, estabilidad, robustez).
 - Seleccionar automáticamente el modelo con mejor desempeño.
 - Implementar estrategia de fallback (modelo secundario).
- **Sistema de Promoción Automática:**
 - Promover automáticamente el modelo campeón a producción.
 - Mantener modelos anteriores disponibles para rollback.
 - Registrar decisiones y justificaciones de selección.

Entregables:

- Sistema de evaluación y selección automática operativo.
- Dashboard de métricas comparativas.
- Modelo campeón en producción.

3.4 Fase 4: Despliegue del API y Sistema de Monitoreo con Reentrenamiento Automático

Objetivo: Poner en producción el API Gateway y activar el sistema de monitoreo continuo con capacidad de reentrenamiento autónomo.

Acciones Clave:

- **Despliegue del API Gateway:**
 - Implementar y exponer endpoints REST:
 - `/predict` - Pronóstico de demanda (múltiples horizontes temporales).
 - `/metrics` - Métricas de desempeño actual.
 - `/health` - Estado del sistema y modelos.
 - `/retrain` - Trigger manual de reentrenamiento (si así lo requiere).
 - Configurar autenticación, rate limiting y logging.
 - Documentar API (OpenAPI/Swagger).
- **Sistema de Predicción en Producción:**
 - Activar generación automática de pronósticos:
 - Mensual (actualización mensual).
 - Semanal (jueves antes de 12pm).
 - Diario (6am día previo).
 - Intradiario (3 veces al día).
 - Configurar granularidades: horaria y cada 15 minutos.
- **Sistema de Monitoreo y Reentrenamiento Automático:**
 - **Monitoreo Continuo:**

- Comparar predicciones vs. demanda real diariamente.
 - Calcular MAPE mensual de forma rolling.
 - Detectar automáticamente cuando $MAPE > 5\%$.
- **Trigger de Reentrenamiento Automático:**
 - Al detectar $MAPE > 5\%$, disparar automáticamente:
 1. Obtención de datos históricos hasta día anterior.
 2. Reentrenamiento de los modelos.
 3. Selección automática del nuevo modelo campeón.
 4. Promoción a producción sin intervención manual.
 - Registrar eventos de reentrenamiento y causas.
- **Detección de Anomalías:**
 - Loggear desviaciones $> 5\%$ por 2 días consecutivos.
 - Identificar patrones anómalos en demanda vs. variables climáticas.
 - Alertar sobre degradación del modelo.
- **Pruebas de Integración:**
 - Validar integración con consumidores del API.
 - Ejecutar pruebas de carga y resiliencia.
 - Simular escenarios de reentrenamiento automático.

Entregables:

- API Gateway en producción.
- Documentación técnica completa.
- Sistema de monitoreo y reentrenamiento automático operativo.
- Logs y trazabilidad de operación.

3.5 SEGUIMIENTO DEL PROYECTO

3.5.1 Reuniones de Avance

Se realizarán reuniones semanales de avance con el equipo designado para:

- Presentar el progreso de cada fase
- Resolver dudas o bloqueos técnicos
- Validar entregables parciales
- Ajustar prioridades si es necesario

Calendario de reuniones:

- Reunión 1: Semana del 10 de noviembre - Avance Fase 1
- Reunión 2: Semana del 18 de noviembre - Avance Fase 2
- Reunión 3: Semana del 25 de noviembre - Avance Fase 3
- Reunión 4: Semana del 2 de diciembre - Avance Fase 4 y entrega final

4. FECHAS Y COTIZACIÓN DEL PROYECTO

4.1 CRONOGRAMA Y FECHA DE ENTREGA

4.1.1 Fecha de Entrega Comprometida

5 de diciembre de 2025

4.1.2 Condiciones de Entrega

La fecha de entrega establecida está sujeta a las siguientes condiciones:

1. Acceso oportuno a datos: se debe proporcionar acceso completo a los sistemas de datos históricos dentro de los primeros 5 días hábiles posteriores al inicio del proyecto.
2. Disponibilidad de infraestructura: se debe garantizar la disponibilidad de la infraestructura de despliegue o confirmar las especificaciones técnicas dentro de los

primeros 5 días hábiles.

3. Validaciones y aprobaciones: Las validaciones y aprobaciones por parte de durante el desarrollo no deberán exceder 2 días hábiles por solicitud.
4. Responsabilidad por demoras: En caso de que el cliente genere demoras, baches, o no proporcione los recursos necesarios en los tiempos establecidos, la fecha de entrega se ajustará proporcionalmente al tiempo de retraso ocasionado.

4.2 INVERSIÓN TOTAL

El costo total del proyecto es de \$14.000.000 COP (CATORCE MILLONES DE PESOS COLOMBIANOS), distribuidos por fase de desarrollo según el siguiente detalle:

4.3 DISTRIBUCIÓN POR FASES

Fase	Descripción	%
Fase 1	Configuración del Pipeline Automático de Datos	30%
Fase 2	Desarrollo de Modelos y Sistema de Entrenamiento	35%
Fase 3	Sistema de Validación y Selección Automática	15%
Fase 4	Despliegue del API y Sistema de Monitoreo	20%
TOTAL PROYECTO		14.000.000

4.4 CONDICIONES DE PAGO

- El pago deberá ser saldado en un plazo máximo de 10 días hábiles contados a partir de la fecha de entrega final (5 de diciembre de 2025).
- Fecha límite: 19 de diciembre de 2025.

4.5 EXCLUSIONES Y CONSIDERACIONES

Las modificaciones de alcance no contempladas en este documento no se encuentran incluidas en el precio.

Supuestos de la cotización:

- EPM cuenta con sistemas de almacenamiento de datos históricos operativos
- Acceso a APIs de datos meteorológicos disponibles
- Sin cambios significativos en los requisitos regulatorios durante el mes de desarrollo