# **FORMULARIO**

**POSTULACIÓN DE ENTIDADES PÚBLICAS EN LAS PRUEBAS DE CONCEPTO**

## **Descripción**

El presente formulario tiene por objetivo la postulación de ENTIDADES PÚBLICAS al Centro de Excelencia y Apropiación en Big Data y Data Analytics (ALIANZA CAOBA) al que se le asignará uno de los grupos de las Universidades inscritas en la Alianza por proyecto aprobado. Esta información debe ser diligenciada por la persona de contacto y responsable del reto.

## **Datos básicos**

* Nombre de la Entidad Pública Fondo Nacional del Ahorro S.A
* Número de Identificación de la Entidad Pública 899.999.284-4
* Nombre completo del responsable del reto Aura Elisa Guerrero Moreno
* Identificación responsable del reto CC 35510194
* Correo Institucional aguerrero@fna.gov.co
* Teléfono de Contacto 3133685060
* Ciudad principal de la ubicación de la Entidades Públicas Bogotá DC

## **Descripción reto analítico**

El desarrollo de los proyectos se divide en tres niveles:

Nivel 1) Identificación de fuentes de información dentro de la entidad y organización de la información para generar conjuntos de datos abiertos.

Nivel 2) Hacer modelos tipo laboratorio para mejorar la calidad de los conjuntos de datos abiertos dentro del portal y automatizar su cargue.

Nivel 3) Generar modelos de IA/ML con el uso de datos que disponga la entidad (no necesariamente datos abiertos) y disponibilizarlos dentro de la infraestructura de la entidad incluido con tableros de control

Conforme a su nivel de madurez con datos marque el nivel al que desea postular su prueba de concepto según anexo técnico (sólo una opción)

Nivel 1 de complejidad baja:

Nivel 2 de complejidad media:

Nivel 3 de complejidad alta: X

Describa la necesidad y contexto del reto analítico de su organización (Máximo 5000 caracteres)

El Fondo Nacional del Ahorro S.A cuenta con diferentes canales de atención para recibir PQRS, en especial un correo donde aproximadamente llegan entre 5000 a 6000 solicitudes mensuales. Además, existe un equipo el cual realiza la clasificación de las solicitudes en sus diferentes tipologías de forma manual; Por lo tanto, buscamos automatizar esta clasificación a través de un modelo de Machine Learning.

¿Por qué quiere realizar un proyecto de Analítica, Big Data, ¿Machine Learning o Inteligencia Artificial? (Máximo 1000 caracteres)

Este problema requiere de modelos de procesamiento del lenguaje natural, análisis de datos y conocimiento de negocio, por lo que es un reto perfecto para un equipo de analítica.

¿Qué problemática va a resolver? (Máximo 1000 caracteres)

Automatizar procesos manuales y disminuir los tiempos de respuesta de las solicitudes.

¿Quiénes son los directamente beneficiados? (Máximo 500 caracteres)

Tanto el área operativa en la gestión de sus procesos, como el cliente al recibir una respuesta oportuna.

¿Qué información desea lograr para las decisiones de negocio? (Máximo 500 caracteres)

A partir de la clasificación de las solicitudes y de su volumen por tipología, se puede analizar y generar insights, que ayuden entender e identificar las áreas/procesos que se deben mejorar.

¿Qué procesos están o estarían interesados en resolver la necesidad?

(Máximo 500 caracteres)

Gerencia PQRS.

¿Cuál es el objetivo general del reto analítico? (Máximo 200 caracteres)

Clasificación de texto y análisis de datos.

Teniendo en cuenta el objetivo general, ¿cuál es la pregunta analítica que se desea resolver? (Máximo 500 caracteres)

¿De acuerdo con el análisis de datos que se realice y a partir de la clasificación de las solicitudes, cuáles son las posibles acciones que debe tomar el Fondo Nacional del Ahorro S.A. para mejorar y optimizar los procesos en cada una de las áreas?

¿En cuál de las siguientes tareas analíticas se encuentra enmarcada su proyecto? (Marque máximo 3)

* Regresión
* Clasificación
* Agrupamiento
* Segmentación
* Clustering
* Reglas de asociación
* Detección de anomalías
* Sistema de recomendación
* Aprendizaje profundo
* Minería de texto
* Segmentación de imágenes
* Selección de características o variables.

## **Temática de postulación del reto analítico**

Título tentativo del reto (Máximo 200 caracteres)

Algoritmo de clasificación de texto para análisis de PQRS, aplicado en el Fondo Nacional del Ahorro S.A.

Acercamientos previos al problema (Máximo 4000 caracteres)

Actualmente la clasificación se hace de forma manual a partir de unos parámetros establecidos.

¿El problema ya ha sido previamente afrontado por la organización? ¿Qué aprendizaje se ha tenido? ¿Se puede tener acceso a dichos resultados o productos?

De forma general no se ha afrontado.

¿Existen restricciones que se deban tener en cuenta?

1. **Privacidad y Seguridad de los Datos:** Es crucial proteger la privacidad de los datos personales de los clientes, asegurando el cumplimiento de regulaciones como la Ley de Protección de Datos Personales (Ley 1581 de 2012 en Colombia) y otras normativas relevantes.
2. **Acceso a Datos Sensibles:** La disponibilidad y acceso a datos geoespaciales precisos y actualizados pueden estar limitados debido a restricciones legales, políticas internas de la organización o acuerdos de confidencialidad con proveedores de datos.
3. **Capacidad Tecnológica:** Asegurarse de contar con la infraestructura tecnológica adecuada, incluyendo sistemas de información geográfica (GIS) y capacidades analíticas robustas para manejar grandes volúmenes de datos geoespaciales.
4. **Costo y Recursos:** Implementar y mantener sistemas de georreferenciación y análisis puede requerir una inversión significativa en términos de recursos financieros, tecnológicos y humanos.
5. **Interoperabilidad de Sistemas:** Garantizar que los sistemas de georreferenciación y análisis puedan integrarse eficientemente con otros sistemas internos de la entidad para asegurar la coherencia y precisión de los datos.
6. **Aspectos Éticos y Sociales:** Considerar el impacto ético y social de la segmentación geográfica y personalización de servicios, asegurando que no genere discriminación o exclusiones indeseadas.

Información adicional sobre cualquier posible restricción que se deba tener en cuenta para la ejecución del proyecto. (datos, formatos, confidencialidad, tiempos, recursos, etc.)

No actualmente

Inventario de fuentes de datos disponibles (En los casos en los cuáles no existe fuente de información colocar: No aplica)

FUENTE 1: Solicitudes radicadas por el canal de contáctenos

FUENTE 2: Parametrización de las tipologías

*Listado de las fuentes de datos, descripción y características de la fuente: Detallar cualquier situación especial que pueda demorar o dificultar la disponibilidad de la información. (Permisos, trámites ante entidades, anonimización de los datos, preparación previa, etc).*

*Permisos para acceder y proceso de anonimización de la información*

*Exprese el número de registros y variables por fuente de datos. Identifique su variable independiente y variables dependientes.*

*Variable independiente: Texto de la solicitud (PQRS)*

*Variable dependiente: Etiqueta de la tipología*

*Números de registros: 5000 aprox*

Inventario de fuentes de datos disponibles (En los casos en los cuáles no existe fuente de información colocar: No aplica)

No aplica

Se Autoriza el cargue a la Plataforma de Analítica como Servicio CAOBA en capa (ver términos de referencia) \*:

Si

## **Postulación de documentos**

* Adjuntar carta de asignación rol del personal de apoyo

Muchas gracias por compartir el reto. Por favor enviar el documento al correo [lpantoja@mintic.gov.co](mailto:lpantoja@mintic.gov.co) con el Asunto: Postulación reto calidad de datos + Nombre entidad al correo [lpantoja@mintic.gov.co](mailto:lpantoja@mintic.gov.co). Fecha máxima de postulación 30 de junio de 2024.