









**CARACTERIZACIÓN DE ADICIONES CONTRACTUALES PARA PROCESOS DE CONTRATACIÓN
PÚBLICA EN INFRAESTRUCTURA**

TABLA DE REQUERIMIENTOS				
ASPECTO	REQUERIMIENTO	PRUEBA PREVISTA	CRITERIO O MÉTRICA DE EVALUACIÓN Y RANGOS DESEADOS	CUMPLIMIENTO
Desempeño	Selección del modelo supervisado para predecir probabilidad de adiciones en contratos públicos.	Evaluación comparativa de modelos utilizando datos de entrenamiento y validación y GridSearchCV.	El modelo seleccionado debe tener el desempeño más alto en las métricas de Accuracy, F1 y AUC. El valor de AUC debe ser mayor a 60%. El Accuracy, F1 y AUC se medirán en una escala del 0 al 1.	
Interpretabilidad	El modelo debe ser comprensible y explicar claramente cómo se toman las decisiones.	Evaluación de la capacidad del modelo para proporcionar explicaciones claras y comprensibles utilizando técnicas como Reglas de Asociación, Regresión LIME, o Diagramas de Influencia.	Porcentaje de explicaciones entendibles por personas sin conocimientos técnicos en el área, evaluado mediante pruebas con usuarios finales y expertos en el dominio del problema. El rango deseado es del 80% o más de explicaciones consideradas entendibles por los evaluadores.	
Funcional	Implementación de API para cálculo de probabilidad de adiciones en nuevos contratos	Pruebas de integración y funcionamiento de la API con datos de prueba	API funcionando correctamente y dando resultados precisos de probabilidad de adiciones. CUMPLE/NO CUMPLE.	
Escalabilidad	Evaluación de la escalabilidad del modelo para manejar grandes volúmenes de datos.	Pruebas con diferentes tamaños de datos para evaluar la escalabilidad del modelo.	El modelo es capaz de manejar grandes volúmenes de datos de manera eficiente, sin degradación significativa en el rendimiento. Se puede medir mediante pruebas con diferentes tamaños de datos y evaluación del tiempo de respuesta y consumo de recursos del sistema en función del tamaño de los datos. Criterio: CUMPLE/NO CUMPLE.	 
Desempeño	Generación de reporte en PDF con resultados de predicciones y	Validación de descarga del reporte en PDF con resultados	Reporte en PDF completo y visualmente comprensible con los resultados de predicciones y	

	visualización de características.	de predicciones y visualización de datos	visualización de características. CUMPLE/NO CUMPLE.	
Seguridad	Establecer políticas y procedimientos de seguridad para el manejo de datos sensibles.	Verificación de las políticas y procedimientos de seguridad establecidos.	Las políticas y procedimientos de seguridad deben estar documentados, implementados y seguidos correctamente. Seguimiento de incidentes de seguridad: Se puede registrar y evaluar cualquier incidente de seguridad que ocurra en el proceso y crear un porcentaje de políticas documentadas e implementadas correctamente en relación con el total de políticas. Errores < 10%.	✓
Funcional	Acceso a datos: Incluir datos de Colombia Compra Eficiente de la base de datos de SECOP I.	Conexión directa desde Jupyter Notebook a través de Socrata Open Data API.	Verificar que los datos sean obtenidos de forma segura y completa, cumpliendo con las políticas de acceso y privacidad establecidas. CUMPLE/ NO CUMPLE.	✓ ⚠
Funcional	Transformación de datos: Verificar balanceo, generar dummies y transformaciones numéricas.	Aplicar procesos de generación de variables dummy y normalización de datos a partir de códigos.	Criterio cualitativo: Verificar el balanceo y que las variables sean creadas correctamente y que la normalización de datos sea exitosa.	✓
Temporalidad	Optimizar el tiempo de respuesta del sistema de predicción de adiciones en contratos públicos.	Medición del tiempo de ejecución del proceso de asignación de probabilidad.	El tiempo de ejecución del proceso de extracción debe estar dentro de los límites establecidos y ser eficiente en términos de rendimiento (menos de 30 segundos por contrato).	✓
Mantenimiento	Actualización periódica de los datos de contratos públicos	Verificación de la actualización periódica de los datos de contratos públicos en el sistema.	Los datos de contratos públicos deben ser actualizados periódicamente cada semana, con un margen de error de +/- 3 días.	✓

Calidad	Limpieza de datos	Verificación de la corrección de errores de reconocimiento de tildes y la unicidad en la forma de llamar a las observaciones.	Los errores de reconocimiento de tildes deben corregirse y la unicidad en la forma de llamar a las observaciones debe mantenerse. Máximo el 0,5% de los datos podrá contener errores alfanuméricos.	
Mantenimiento	Diseñar un sistema de predicción de adiciones en contratos públicos que sea fácil de mantener y actualizar.	Evaluación de la estructura del código fuente, la documentación y la facilidad de actualización del sistema.	El sistema debe tener una estructura de código fuente clara y bien organizada, con comentarios y documentación adecuada para facilitar su mantenimiento. Debe ser fácil de actualizar con nuevas funcionalidades o mejoras sin causar interrupciones en el funcionamiento del sistema. Este criterio se evaluará cualitativamente, por ejemplo, mediante una revisión por parte del equipo desarrollador (los contratistas).	
Calidad	Garantizar la calidad de los datos utilizados para entrenar y evaluar el modelo de predicción de adiciones en contratos públicos.	Análisis de la calidad de los datos utilizados, como la completitud y consistencia de los mismos.	Los datos utilizados para entrenar y evaluar el modelo deben cumplir con criterios de calidad establecidos previamente, como tener una completitud y consistencia adecuada. Al menos el 95% de los datos utilizados deben ser completos y sin errores.	