

PLIEGO DE PRESCRIPCIONES TÉCNICAS PARA LA EVOLUCIÓN DE UN SISTEMA DE SOPORTE A LA OPERACIÓN A PARTIR DE TÉCNICAS AVANZADAS DE IA

ACLARACIONES A LAS PREGUNTAS PLANTEADAS

Las preguntas recibidas durante el plazo marcado (finalizado el martes 21 de julio de 2026), se listan a continuación con sus respuestas. La numeración corresponde al orden de recepción de las mismas.

Pregunta 1: ¿Se facilitarán al adjudicatario el código fuente, los modelos, los scripts, la documentación técnica, los resultados de validación y las conclusiones correspondientes a la primera fase de PISA-MaT?

Respuesta:

Se facilitará la documentación técnica, los resultados de validación y las conclusiones. No se facilitarán los modelos, código fuente o scripts, aunque se dispondrá de acceso a los mismos si se considerase necesario.

Pregunta 2: Dado que el REQ 1 exige aplicar técnicas alternativas a las utilizadas en el prototipo actual, ¿podrían indicar qué familias de modelos o aproximaciones se emplearon en la primera fase y cuáles fueron sus principales limitaciones?

Respuesta:

La aproximación utilizada se detallará al inicio del proyecto. Como orientación, se utilizaron técnicas de Machine Learning automatizado, así como de Tree Ensemble (Gradient Boosting Machine y Random Forest). Las limitaciones están descritas en el PPT, centrándose en la tasa obtenida, especialmente en condiciones de alta carga. Para un nivel de detalle algo más específico, consultar la respuesta a la pregunta 41.

Pregunta 3: ¿Qué conjuntos de datos proporcionará CRIDA y cuál es aproximadamente su periodo temporal, volumen, formato, frecuencia de muestreo y nivel actual de limpieza o preprocesamiento?

Respuesta:

CRIDA proporcionará datos de vuelo, operacionales y meteorológicos, así como otros potenciales datos ATM de los que disponga y se consideren relevantes. El dato proporcionado es de calidad operacional (el utilizado por ENAIRE en la prestación de servicio). Se trata de fuentes de datos que registran todos los eventos discretos, con la excepción de la vigilancia radar, con posiciones de traza cada 5 segundos. Como orientación, CRIDA dispone de más de cinco años de datos.

Pregunta 4: ¿Los datos, diccionarios de variables, esquemas, muestras y documentación de las diferentes fuentes estarán disponibles desde el inicio del contrato?

Respuesta:

Dependiendo de las fuentes requeridas, la documentación es más o menos completa. Todos los datos y documentación existente serán accesible desde el inicio de las actividades.

Pregunta 5: ¿Cuál es la definición formal de la variable objetivo «Millas a Toma» y cómo se obtiene la etiqueta real a partir de las trazas de vuelo?

Respuesta:



Las Millas a Toma son las millas a recorrer antes de tomar tierra. Su incertidumbre surge porque, al entrar en el TMA, aún no se sabe con certeza qué “camino” volará finalmente cada aeronave.

No se entiende el uso de “etiqueta” en este contexto. Si se refiere al valor real de Millas a Toma, se calculan a partir de la traza real volada por una aeronave.

Pregunta 6: ¿Cómo deberán tratarse las aproximaciones frustradas, los circuitos de espera, los cambios de pista y los vuelos que atraviesen más de una vez los horizontes de 40, 35, 30, 25 o 20 NM?

Respuesta:

Cada nuevo cruce es una nueva toma de decisión por parte del operador. Por tanto, deberá proporcionarse una predicción nueva ante cada posible entrada.

Pregunta 7: ¿Todas las variables enumeradas en el REQ 5 estarán disponibles en tiempo real en el entorno final? En su caso, ¿qué frecuencia de actualización y latencia tendrá cada fuente?

Respuesta:

Sí. Descrito en la respuesta a la Pregunta 3.

Pregunta 8: ¿El tratamiento de los datos y el entrenamiento deberán realizarse íntegramente en infraestructura de CRIDA/ENAIRES o podrá autorizarse un entorno externo específico de desarrollo y entrenamiento?

Respuesta:

Está previsto realizar el tratamiento de datos y entrenamiento en un entorno externo a CRIDA/ENAIRES. El adjudicatario deberá disponer de sus propios recursos de entornos y capacidades hardware.

Pregunta 9: ¿Podrían confirmar exactamente las dos alternativas de cumplimiento previstas en el REQ 3 y, en particular, el límite de MAE aplicable a la alternativa del 95 % de predicciones con error inferior a 5 NM?

Respuesta:

Se requiere una tasa de acierto igual o superior al 93% para un $MAE < 4N$, o superior al 95% para un $MAE < 5NM$.

En condiciones de alta carga deberá haber una desviación menor a $\pm 1\%$ de la precisión del modelo general para cualquier condición de tráfico.

Pregunta 10: ¿Las métricas deberán cumplirse globalmente o de manera independiente para cada horizonte de 40, 35, 30, 25 y 20 NM?

Respuesta:

Globalmente.

Pregunta 11: ¿La unidad de evaluación será cada vuelo, cada predicción generada en cada horizonte o una combinación de ambas?

Respuesta:

Cada vuelo.

Pregunta 12: ¿Cómo se contabilizarán los vuelos en los que no sea posible emitir una predicción por ausencia, retraso o corrupción de alguna fuente de datos?

Respuesta:

Como casos no nominales.

Pregunta 13: ¿Las métricas deberán cumplirse también de manera independiente para cada configuración de pista, tipología de aeronave o condición meteorológica?

Respuesta:

Globalmente, con la salvedad de las condiciones de alta carga.

Pregunta 14: ¿Cuál será la definición cuantitativa de una situación de alta carga de tráfico y qué variable o combinación de variables se utilizará para determinarla?

Respuesta:

Aquellas que superen un umbral respecto a la capacidad operativa disponible en el momento. Se considerará el volumen y la complejidad del tráfico, así como la capacidad nominal.

Pregunta 15: En relación con la alta carga, ¿debe entenderse que las métricas no pueden ser inferiores a las globales o que se admite una desviación máxima de ± 1 % respecto de la precisión general?

Respuesta:

Como indica el Requisito 6, en alta carga el objetivo es tener una desviación menor a ± 1 % de la precisión del modelo general para cualquier condición de tráfico.

Pregunta 16: ¿Qué conjunto de datos y qué procedimiento concreto se utilizarán para las pruebas de aceptación final? ¿Se empleará un conjunto temporal reservado que no haya sido utilizado durante el entrenamiento ni la selección de hiperparámetros?

Respuesta:

CRIDA utilizará un dataset reservado, no facilitado antes al adjudicatario ni para el entrenamiento ni para sus pruebas.

Pregunta 17: ¿Se contemplará un periodo de ajuste o subsanación si las métricas no se alcanzan en la primera ejecución de las pruebas de aceptación?

Respuesta:



Sí.

Pregunta 18: ¿La aceptación final y el segundo pago dependerán exclusivamente del cumplimiento estricto de las métricas del REQ 3 o de la valoración conjunta de las métricas, los entregables, la integración y el resto de requisitos del proyecto?

Respuesta:

Si bien el objetivo es alcanzar esas métricas, los responsables de Expediente valorarán los resultados en conjunto. La aceptación final y el segundo pago no estarán necesariamente vinculadas al cumplimiento de las métricas, aunque no se aceptarán soluciones que empeoren los resultados del prototipo inicial.

Pregunta 19: ¿Podrían facilitar información sobre las tecnologías, sistemas operativos, herramientas e interfaces que conforman actualmente el Data Lakehouse de CRIDA/ENAIRE?

Respuesta:

A colación de la respuesta a la pregunta 8, la descripción de los sistemas de CRIDA es innecesaria puesto que el entrenamiento del modelo se llevará a cabo externamente.

Pregunta 20: ¿Qué recursos de CPU, memoria, almacenamiento y GPU estarán disponibles para el entrenamiento, validación e inferencia?

Respuesta:

El entrenamiento e inferencia del modelo se llevará a cabo externamente, sin utilizar recursos de CRIDA.

Pregunta 21: ¿Qué mecanismos deberán utilizarse para la recepción de datos y la devolución de predicciones: API, mensajería, streaming, acceso a bases de datos u otros?

Respuesta:

CRIDA facilitará un dataset por un medio a acordar en la reunión de inicio. Una vez los datos estén bajo custodia del adjudicatario, este establecerá sus mecanismos de acceso.

Durante la fase de pruebas, el prototipo se probará en las plataformas de simulación en tiempo real de CRIDA, intercambiándose información mediante mensajería o mecanismo alternativo a acordar.

Pregunta 22: ¿Cuál es la latencia máxima admisible y qué volumen máximo de aeronaves, eventos o predicciones concurrentes deberá soportar la solución?

Respuesta:

Se plantea una latencia máxima de 2 (DOS) segundos. El volumen máximo de predicciones concurrentes se estima en 50 (CINCUENTA)

Pregunta 23: ¿La exigencia de funcionamiento 24/7 implica que la solución deberá incorporar alta disponibilidad, redundancia y recuperación automática? En su caso, ¿qué niveles de disponibilidad, RTO y RPO deberán garantizarse?

Respuesta:

No es objeto de esta licitación la puesta en operación de este sistema. El requisito de funcionamiento 24/7 corresponde a la consideración de esta necesidad en la fase de diseño.

Pregunta 24: ¿El desarrollo o modificación de las interfaces de usuario queda fuera del alcance del contrato, limitándose el adjudicatario a proporcionar el motor de inferencia y sus interfaces de integración?

Respuesta:

Sí.

Pregunta 25: ¿La solución deberá instalarse únicamente en el entorno de preproducción o deberá realizarse también una integración completa en producción?

Respuesta:

Como se indica en la pregunta 23, este sistema no será puesto en producción dentro de este expediente.

Pregunta 26: ¿CRIDA dispone actualmente de herramientas de registro de experimentos, model registry, monitorización y reentrenamiento con las que deba integrarse la solución? ¿El adjudicatario deberá ejecutar futuros reentrenamientos o únicamente dejar preparado el correspondiente pipeline?

Respuesta:

A efectos de esta licitación, el adjudicatario considerará que la integración con las herramientas de gestión de modelos existentes en CRIDA (MLFlow) es opcional, no entendiéndose en ningún caso como obligatoria. Se proporciona más detalle de estas en la respuesta a la pregunta 43.

El adjudicatario deberá preparar el pipeline, siendo futuros reentrenamientos responsabilidad de CRIDA.

Pregunta 27: ¿Existen restricciones específicas sobre lenguajes de programación, frameworks, librerías, imágenes de contenedores o licencias de código abierto que deban tenerse en cuenta?

Respuesta:

Ninguna, aparte de las expresadas en el Pliego de Prescripciones Técnicas.

Pregunta 28: ¿Qué actuaciones concretas comprende la garantía técnica mínima de dos años? En particular, ¿incluye corrección de errores, actualización de dependencias, resolución de vulnerabilidades, reentrenamientos o adaptación a cambios en los datos o en la infraestructura?

Respuesta:



Al no ser un sistema de producción, la garantía se centra en la resolución de incidencias de funcionamiento/instalación.

Pregunta 29: ¿Se establecerán tiempos máximos de respuesta o resolución, un límite de consultas u horas de soporte durante el periodo de garantía? ¿La referencia al funcionamiento 24/7 implica también disponibilidad humana del adjudicatario?

Respuesta:

Al no ser un sistema de producción, no se considera necesario fijar tiempos de respuesta. La mención al funcionamiento 24/7 debe entenderse como un requisito de diseño, no de despliegue.

Pregunta 30: ¿Cuál será el régimen de propiedad intelectual aplicable al código fuente, modelos, pesos, scripts, pipelines, configuraciones y documentación desarrollados durante el proyecto?

Respuesta:

Todos los modelos, códigos, scripts y, en general, cualquier resultado del proyecto, serán propiedad de CRIDA.

Pregunta 31: ¿Podrán mantenerse como propiedad del adjudicatario los componentes genéricos, metodologías, librerías internas o desarrollos preexistentes, concediendo a CRIDA los derechos necesarios sobre la solución específica entregada?

Respuesta:

CRIDA deberá, en todo caso, recibir los derechos necesarios para operar el sistema entregado completo, incluso si ello incluye componentes preexistentes. Cualquier código o modelo desarrollado bajo el marco de este expediente será propiedad de CRIDA.

Pregunta 32: ¿Qué niveles mínimos y qué medios concretos se exigirán para acreditar la solvencia económica y financiera y la solvencia técnica o profesional?

Respuesta:

Como indica la sección 7 del Pliego de Prescripciones Técnicas, a través del certificado propio que garantice que se dispone de la documentación administrativa necesaria para contratar con el Sector Público, en los epígrafes allí listados.

Pregunta 33: En el caso de empresas de nueva creación, ¿podrán acreditarse la solvencia técnica y profesional mediante las titulaciones, experiencia, proyectos y medios de los profesionales integrantes del equipo?

Respuesta:

Las empresas licitadoras deberán entregar un certificado propio que garantice los epígrafes listados en la sección 7. La documentación administrativa efectiva sólo deberá ser presentada por la empresa que resulte adjudicataria, una vez se haya notificado su condición de adjudicataria del presente Pliego y siempre antes de la firma del contrato.

Pregunta 34: ¿Se permitirá complementar la solvencia y los medios propios mediante otras entidades o profesionales externos? Asimismo, ¿se admite la presentación mediante UTE y cómo se acumularán o valorarán la solvencia, la experiencia y los medios aportados por sus integrantes?

Respuesta:

Ver respuestas anteriores. Sí se admite la presentación mediante UTE.

Pregunta 35: ¿Se permite la subcontratación parcial de perfiles o actividades especializadas? En su caso, ¿deberán identificarse en la oferta las actividades y entidades previstas?

Respuesta:

No se consideran exclusiones en este sentido.

Pregunta 36: ¿Existe un modelo oficial para el certificado relativo a la documentación administrativa o será suficiente una declaración responsable firmada por el representante? ¿Qué documentación concreta deberá aportar posteriormente el adjudicatario dentro del plazo de cinco días hábiles?

Respuesta:

Ver respuesta a la pregunta 33. La documentación concreta es la listada en la sección 7 del PPT (certificados propios en todos los casos).

Pregunta 37: ¿Se exigirá garantía definitiva, seguro de responsabilidad civil o profesional, o alguna certificación concreta en materia de ciberseguridad, seguridad de la información, calidad o protección de datos?

Respuesta:

Se exigirá, y en su caso valorará, exclusivamente lo indicado en el Pliego de Prescripciones Técnicas.

Pregunta 38: ¿Qué nivel de presencialidad se prevé durante los nueve meses de ejecución y cuántas reuniones, jornadas de trabajo, instalaciones o pruebas presenciales se estiman aproximadamente?

Respuesta:

Las actividades de desarrollo no se llevarán a cabo en las instalaciones de CRIDA, que únicamente se utilizarán para reuniones y pruebas de validación. La cantidad de reuniones, instalaciones u otros aspectos serán propuestas por los licitadores en sus planificaciones y, eventualmente, acordadas con CRIDA tras el arranque del expediente.

Pregunta 39: ¿Cuáles son las principales limitaciones o carencias del modelo predictivo actual (fase inicial de PISA-MaT) que han motivado el lanzamiento de esta evolución? Conocer qué aspectos concretos se busca mejorar (*precisión, robustez ante ciertos eventos, latencia, escalabilidad, etc.*) nos permitiría orientar la aproximación técnica con mayor foco.

Respuesta:

Las principales carencias del modelo son las descritas en el PPT: una tasa de precisión inferior a la deseada y, principalmente, el decaimiento de esta en condiciones de alta carga.

Pregunta 40: En relación con los casos no nominales descritos en REQ 4 (*frustradas, esperas prolongadas en el TMA, vectores ATC imprevistos, discrepancias plan de vuelo / traza radar, cambios de configuración de pistas o meteorología*), ¿se abordó ya algún tipo de estrategia específica para estos eventos en la primera fase? En caso afirmativo, agradeceríamos una descripción, aunque sea a alto nivel, del enfoque seguido y de los resultados obtenidos.

Respuesta:

Cada anomalía identificada en la fase inicial implicó una decisión independiente. Por ejemplo, el caso de las frustradas y esperas se descartaron del dataset de entrenamiento. El caso de los cambios de configuración asumimos los posibles errores. En el caso de las discrepancias entre traza y plan de vuelo intentamos caracterizarlas con la información más fiable. Los vectores/directos o recortes de trayectorias son algo que no se puede predecir con exactitud, pero se incluyen variables en el modelo que permiten inferir cuando estos pueden ser más probables (ej. Ocupación en el entorno aeroportuario).

Pregunta 41: Con el fin de proponer técnicas alternativas coherentes con REQ 1, ¿podrían indicar qué familia de algoritmos y qué estrategia de entrenamiento se emplearon en el prototipo actual? En particular también nos interesa saber: (a) si se utilizó un único modelo o modelos separados por horizonte de distancia (40, 35, 30, 25 y 20 NM), y (b) si se realizó optimización de hiperparámetros y bajo qué criterio.

Respuesta:

Se entrenaron distintos modelos, uno por cada horizonte. Se realizaron pruebas con varios algoritmos, se utilizaron técnicas de Tree Ensembles, como Random Forest y modelos de GBM; el ajuste de hiperparámetros en este caso se utilizó un grid search optimizando el RMSE. También se utilizó la librería de H2O para realizar AutoML para la elección del mejor modelo y el ajuste de hiperparámetros optimizando el RMSE.

Pregunta 42: Entendemos que el objeto de esta contratación se circunscribe al aeropuerto Josep Tarradellas Barcelona-El Prat (LEBL), y que no se persigue un modelo genérico ni multi-aeropuerto en el marco de este expediente. ¿Es correcta esta interpretación?

Respuesta:

Sí.

Pregunta 43: En línea con REQ 5 y REQ 17, ¿podrían precisar el número aproximado de fuentes de datos que se pondrían a disposición del adjudicatario, su tipología (traza radar, plan de vuelo, meteorología, datos contextuales del aeropuerto, etc.) y una estimación del volumen de cada una? Esta información es determinante para dimensionar el pipeline de ingesta y evaluar si el tratamiento requiere un enfoque de Big Data.

Respuesta:

Se recomienda consultar la respuesta a la Pregunta 3. No es posible responder a esta pregunta de manera unívoca, ya que el tamaño de los datos puede variar en las distintas

iteraciones para el entrenamiento del modelo en función de las variables que se consideren. Como referencia, en el dataset con todas las posibles variables (datos limpios y procesados, antes de la selección de variables) utilizado para el entrenamiento del modelo para cada modelo (x5) fueron 150.000 registros y 400 MB (para un año de datos).

Los datos en bruto pueden tener un peso mayor ya que contienen fuentes más pesadas como traza y meteorología en el entorno de LEBL. En cualquier caso, todos estos datos se consideran necesarios para hacer los procesamientos previos y obtener las variables relevantes para el entrenamiento.

Los licitadores tendrán en cuenta que pueden considerar fuentes de datos distintas a las consideradas en el prototipo inicial de CRIDA. En cada caso, la disponibilidad de datos en CRIDA supera los tres años.

Pregunta 44: Respecto al entorno de preproducción y al Data Lakehouse mencionados en REQ 7: (a) ¿la infraestructura actual de CRIDA está ya preparada para realizar despliegues contenerizados (Docker/Kubernetes)? (b) ¿está dimensionada para el procesamiento de volúmenes de datos de gran escala? (c) ¿disponen actualmente de alguna herramienta de gestión del ciclo de vida del modelo (por ejemplo MLflow o servicio análogo) sobre la que integrarse, o el marco de MLOps requerido en REQ 9 debe proveerse íntegramente por el adjudicatario?

Respuesta:

Se recomienda consultar la respuesta a la pregunta 26. Los licitadores tienen la opción de utilizar la infraestructura de CRIDA descrita abajo, pero no la obligación, excepto para el despliegue cara a la validación final.

- a) CRIDA dispone de una infraestructura para el despliegue de modelos dockerizados.
- b) La infraestructura de CRIDA permite la gestión y procesamiento de datos a gran escala. En el prototipo inicial, ya realizado, se consumían y procesaban los datos en tiempo real para la entrega de las predicciones también en tiempo real.
- c) CRIDA dispone de un servidor MLflow. Los licitadores tendrán la posibilidad de integrar los desarrollos y el modelo en este para la gestión del ciclo de vida si lo consideran oportuno, no siendo obligatorio.

Madrid, 22 de julio de 2026

Autor: José Manuel Cordero García