

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY



Maestría en Inteligencia Artificial Aplicada (MNA)

Proyecto Integrador (TC5035)

Dra. Grettel Barceló Alonso

Dr. Luis Eduardo Falcón Morales

Avance 6. Conclusiones clave

“Clasificación de ruido en laboratorio de motores eléctricos automotrices a través de métodos de inteligencia artificial”

EQUIPO 14

Andrei García Torres A01793891

Emmanuel González Calitl A01320739

Denisse María Ramírez Colmenero A01561497

Fecha: 9 de junio de 2024

ÍNDICE

Parte A. Viabilidad del Modelo	2
Parte B. Análisis de implementación del modelo en la nube	4
Bibliografía	6

La formulación del problema para este proyecto ha sido identificar y descartar ruido subjetivo durante la fabricación de motores elevadores que únicamente el cliente percibe al final del ensamble y pruebas de módulo final, el cual es difícil de correlacionar o medir de manera objetiva dentro de los laboratorios de prueba de Bosch para valores fuera de especificación.

Como objetivo se tiene que la meta de este proyecto integrador de la Maestría en Inteligencia Artificial Aplicada es la propuesta de un modelo de clasificación de ruido a través de mediciones realizadas y etiquetadas en el laboratorio de ruido de Robert Bosch. Las mediciones han sido realizadas con equipos especializados en la medición, registro y visualización de datos para cada muestra del objeto de estudio. Con este proyecto se busca integrar un modelo de clasificación de ruido al proceso de producción de la empresa que logre la detección temprana de ruidos de motores elevadores mejorando la calidad del producto final aumentando así la satisfacción del cliente, la reputación de la marca y su competitividad en el mercado.

¿El rendimiento del modelo es lo suficientemente bueno para su implementación en producción?

El rendimiento es bueno para su implementación en producción, puede servir como herramienta de apoyo para la clasificación de ruido en los motores elevadores específicos. El modelo de redes neuronales convolucionales con el uso de la librería de Keras y los datos disponibles para el entrenamiento sin el aumento de los mismos es bueno para clasificar ruido aceptable (ok) y poco eficiente para la clasificación de ruido no aceptable (nok).

¿Existe margen para mejorar aún más el rendimiento?

Si existe margen para mejorar el rendimiento del modelo de clasificación con redes neuronales convolucionales para ruido aceptable y ruido no aceptable. El rendimiento puede mejorar aumentando un porcentaje los datos de entrenamiento de manera sintética y validados nuevamente con el ajuste respectivo de los parámetros relacionados para obtener un mejor desempeño de las métricas del entrenamiento para el set de validación y el set de entrenamiento.

¿Cuáles serían las recomendaciones clave para poder implementar la solución?

La solución que ofrece el modelo de clasificación propuesto solo puede ser utilizado para un solo número de parte de motor elevadores. En caso que se desee realizar la clasificación para otro número de parte o tipo de motor se podría utilizar el modelo propuesto solo que se tendrá que realizar el entrenamiento con los datos específicos para un nuevo número de parte y los ajustes a los parámetros para tener el ajuste óptimo para el nuevo número de parte.

Se sugiere aumentar la base de datos de la clasificación de ruido con las nuevas mediciones realizadas en el laboratorio de ruido que se presentan a través de imágenes de espectrogramas. De tal modo se puede reentrenar la red neuronal convolucional y mejorar el rendimiento del modelo de una manera más precisa.

Se recomienda que la aplicación de la clasificación de ruido se pueda implementar en un servicio de nube como MS Azure, AWS o GCP; de este modo el modelo de clasificación pueda ser implementado y utilizado en diferentes bancos de prueba dentro de la locación en Toluca o pueda ser utilizado en otra locación de la compañía en Europa o Asia. Con esta implementación se podría escalar la solución de una manera eficiente.

¿Qué tareas / procedimientos son accionables para las partes interesadas o (stakeholders) ?

Las tareas o procedimientos accionables para las partes interesadas se muestran a continuación.

Equipo de Desarrollo:

1. Entregar modelo de clasificación para su implementación.
2. Mantenimiento al modelo de clasificación de ruido.
3. Soporte a los técnicos especialistas de ruido para la implementación del modelo de clasificación de ruido.
4. Mantener el compromiso de la privacidad del manejo de datos de las mediciones de ruido para el motor elevacristales específico.

Laboratorio de Ruido Bosch Toluca:

1. Generar más cantidad de imágenes para mejorar la eficiencia del entrenamiento del modelo y agregarlas a la base de datos definida.
2. Utilizar y probar el modelo solo para el tipo de motor elevacristales especificado.
3. Retroalimentar al equipo de desarrollo la calidad del funcionamiento y la implementación del modelo de clasificación de ruido para el motor elevacristales.

Gerente de procesos de desarrollo:

1. Asegurarse de la disponibilidad de los recursos informáticos y administrativos para la implementación del modelo de clasificación para el departamento del laboratorio de ruido.
2. Asegurarse de la implementación de seguridad de la información para el manejo de datos y uso del modelo de clasificación de manera interna en la organización.

3. Divulgar dentro de la organización la habilidad para el desarrollo de nuevos modelos de Inteligencia Artificial en Robert Bosch Toluca.

B. Realizar un análisis detallado para determinar la plataforma de servicios en la nube más adecuada para implementar la solución de ML.

Implementar un modelo de CNN (Red Neuronal Convolucional) para la clasificación binaria de imágenes en la nube es un proceso que puede ser altamente beneficioso debido a la escalabilidad, accesibilidad y capacidad de manejo de grandes volúmenes de datos que ofrecen los servicios en la nube. A continuación se muestran aspectos a considerar:

Manejo de Recursos:

- **Tipo de Instancia:** Selecciona instancias adecuadas (CPU, GPU, TPU). En este caso será necesario usar una instancia de tipo GPU ya que se estará usando imágenes.
- **Escalabilidad:** En dado caso que el cliente quisiera usar este modelo en más bancos de pruebas, un sistema de nube permitiría el escalado automático para manejar la carga variable de predicciones.
- **Monitoreo y Optimización:** Estos servicios tecnológicos permiten utilizar herramientas de monitoreo para optimizar el uso de recursos y reducir costos por cómputo.

Seguridad y Mantenimiento

- **Autenticación y Autorización:** La nube asegura que solo usuarios autorizados puedan acceder al modelo y datos. La marca Azure de Microsoft permitiría vincular directamente la autenticación de Active Directory propio con los servicios de cómputo.
- **Actualizaciones y Versionamiento:** También la nube permitiría mantener al modelo y su entorno actualizados. Así como implementar un versionamiento para gestionar actualizaciones y cambios en el modelo.

Tabla 1. Comparativa de servicios en la nube para montar el modelo de aprendizaje automático.

	Amazon Web Services	Microsoft Azure	Google Cloud Platform	IBM Watson
Autenticación:	Servicio de autenticación propio.	Sí, Se puede integrar el Microsoft Active Directory de la compañía de una manera orgánica. Cada usuario podría autenticarse de manera segura con su usuario de Windows.	Servicio de autenticación propio.	Servicio de autenticación propio.
Despliegue y gestión:	Permite el despliegue de modelos con un solo clic, escalado automático y endpoints HTTP/HTTPS para realizar inferencias.	Permite el despliegue de modelos en contenedores, con escalado automático y endpoints RESTful.	Proporciona despliegue fácil de modelos con escalado automático y endpoints HTTP/HTTPS.	Permite el despliegue de modelos con escalado automático y endpoints API para inferencias.
Desventajas:	Curva de aprendizaje larga para nuevos usuarios. Costos pueden escalar rápidamente con el uso intensivo de recursos.	Puede ser costoso en comparación con otros proveedores. La interfaz y las herramientas pueden ser complejas para nuevos usuarios.	Menos servicios complementarios en comparación con AWS. Soporte puede ser limitado en comparación con otros proveedores para clientes empresariales grandes.	Puede ser más caro en comparación con otros proveedores. Menos flexible en comparación con AWS o GCP para ciertos casos de uso.
Facilidad de uso:	Moderada, ya que permite trabajar en ambientes gráficos	Moderada, ya que permite trabajar en ambientes gráficos	Alta, se trabaja en un entorno de consola.	Moderada, ya que permite trabajar en ambientes gráficos
Costo de tiempo de computo:	Request pricing \$0.0000002 per request. Duration pricing \$0.00001667 for every GB-second used thereafter.	\$ 2 USD/ 1 Token (request)	First active Multiple-target delivery pipeline (per billing account) No charge Each additional active multiple-target delivery pipeline \$5 per billing month	\$ 1050 USD/ mes - Nivel estándar

Tabla 2. Comparativa resumida de los servicios en la nube.

Característica	AWS SageMaker	Azure ML	GCP AI Platform	IBM Watson Studio
Facilidad de Uso	Moderada	Moderada	Alta	Moderada
Escalabilidad	Alta	Alta	Alta	Moderada
Costos	Competitivos	Alto	Competitivos	Alto
Integración	Amplia	Amplia	Amplia	Moderada
Soporte de Frameworks	Amplio	Amplio	Amplio	Amplio
Innovación	Alta	Alta	Muy alta	Moderada
Seguridad y Cumplimiento	Alta	Alta	Alta	Alta

Ver excel completo en:

<https://docs.google.com/spreadsheets/d/1b2oggFht68biFDKw3o8mR39YTv-VIePQAAG4Jv88EM4/edit#gid=0>

Implementar un modelo de CNN para clasificación binaria de las pruebas de motores eléctricos en la nube implica una serie de pasos desde la selección del proveedor de servicios en la nube hasta el despliegue y mantenimiento del modelo. La nube ofrece ventajas significativas en términos de escalabilidad, accesibilidad y gestión de recursos, haciendo que este enfoque sea ideal para proyectos de machine learning en producción.

En caso de que la empresa decida utilizar un servicio en la nube externo, el equipo sugiere utilizar los servicios de nube de Microsoft Azure ya que la parte de seguridad es primordial para el cliente, al usar Azure se permitirá la fácil integración de la autenticación de Windows con este servicio de AI, además de que el costo es favorable para la compañía.

Addendum:

*La carpeta "Modelos 8 de junio" del repositorio de Github tiene los modelos más actuales del equipo con datos de entrenamiento sin aumentación:

https://github.com/emm-gl/project_mna/tree/main/Modelos%2008%20de%20junio

Bibliografía

[1] Chojecki, P. (2020). *Artificial Intelligence Business: How you can profit from IA*. Pack Publishing.

[2] *Aprendizaje supervisado frente a aprendizaje no supervisado: diferencia entre los algoritmos de machine learning* - AWS. (s. f.). Amazon Web Services, Inc.

<https://aws.amazon.com/es/compare/the-difference-between-machine-learning-supervised-and-unsupervised/>

[3] *Productos y servicios*. (s. f.). Bosch En México.

<https://www.bosch.com.mx/productos-y-servicios/>

[4] Efe. (2024, 7 marzo). *Para una conducción más segura: Bosch y Microsoft exploran nuevas posibilidades con la IA generativa*. EFE Comunica.

<https://efecomunica.efe.com/para-una-conduccion-mas-segura-bosch-y-micros oft-exploran-nuevas-posibilidades-con-la-ia-generativa/>

[5] *Calculadora de precios*. (s. f.). Microsoft Azure. Recuperado 8 de junio de 2024, de <https://azure.microsoft.com/es-mx/pricing/calculator/>

[6] *IBM watsonx | Precios*. (s. f.). Recuperado 8 de junio de 2024, de <https://www.ibm.com/mx-es/watsonx/pricing>

[7] *Precios | Cloud Deploy | Google Cloud*. (s. f.). Google Cloud. https://cloud.google.com/deploy/pricing?hl=es#pricing_table