



INSTITUTO TECNOLÓGICO DE ESTUDIOS SUPERIORES DE MONTERREY

Maestría en Inteligencia Artificial Aplicada

Proyecto Integrador

Avance 6: Conclusiones Clave

Datos Generales del Proyecto

Estudiantes:

A01793188 - Carlos Alberto Reynoso González

A01334976 - Israel Campos Báez

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

El modelo Ensemble U-Net tiene un MAE de 1.4444, RMSE de 3.1302 y una correlación R^2 alta de 0.7504 entre las predicciones del modelo y los valores reales de tensión. Estos resultados indican una mejora frente a modelos anteriores. Es posible liberarlo a producción como un modelo que permite tener resultados preliminares muy rápidos. El modelo ha sido evaluado bajo los criterios de éxito previamente establecidos, considerando tanto el desempeño cuantitativo como cualitativo en la etapa de modelado. Los resultados obtenidos sugieren que el enfoque basado en redes neuronales profundas, con U-Net y U-Net++ combinadas en un ensamble heterogéneo, permite obtener predicciones precisas en la estimación de tensiones en materiales compuestos. La integración de Machine Learning y Análisis por Elementos Finitos (FEA) ha demostrado ser una estrategia viable para optimizar el diseño de estos materiales en aplicaciones automotrices.

¿Existe margen para mejorar aún más el rendimiento?

Si; a pesar del buen desempeño, aún existen ciertos casos límite donde la precisión podría mejorarse, especialmente en regiones con distribuciones de estrés complejas. Aunque el modelo actual cumple con los criterios de éxito, existen varias áreas donde se podría mejorar aún más su rendimiento. La principal herramienta de mejora es incrementar el tamaño de la base de datos de entrenamiento que, como ya es bien sabido, en este proyecto represento un cuello de botella por la complejidad y lentitud de generar los datos mediante el simulador ANSYS, así como incorporar más simulaciones de FEA con una mayor diversidad de geometrías y condiciones de carga. De igual forma se podrían explorar estrategias de meta-aprendizaje o modelos adicionales para mejorar la combinación de predicciones, mejorar la calibración de hiperparámetros y aplicar el modelo en datos experimentales para evaluar su generalización fuera del entorno simulado.

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

Para asegurar una implementación exitosa, se recomienda:

- **Desplegar en un entorno de pruebas:** Validar el modelo en un entorno de pruebas con datos reales antes de la implementación completa.
- **Automatización del pipeline:** Integrar el modelo dentro de un flujo de trabajo automatizado, incluyendo preprocesamiento de datos y validación de resultados.
- **Monitoreo continuo:** Implementar un sistema de monitoreo para detectar posibles degradaciones en el desempeño con el tiempo

¿Qué tareas / procedimientos son accionables para las partes interesadas?

Para garantizar el éxito del proyecto, cada grupo de interés deberá considerar los siguientes procedimientos accionables:

Equipo de Ingeniería:

- Probar el modelo con datos reales para validar su rendimiento fuera del entorno simulado.
- Integrar el modelo en el proceso de diseño de materiales, asegurando interpretabilidad y usabilidad.
- Optimizar tiempos de inferencia en la infraestructura elegida para garantizar eficiencia en la ejecución del modelo.

Equipo de Datos/IA:

- Continuar con la optimización del modelo y generación de nuevos datos de entrenamiento.
- Establecer necesidades futuras de optimización, identificando mejoras en arquitectura y entrenamiento.
- Implementar monitoreo y actualización automática.

Equipo de Infraestructura/IT:

- Asegurar seguridad y disponibilidad del modelo, protegiéndolo contra accesos no autorizados y fallos en producción.
- Decidir si la herramienta se distribuirá como un software de escritorio o basado en la web, evaluando ventajas y limitaciones de cada opción.

Gerencia/Dirección:

- Evaluar costos de despliegue, considerando recursos de hardware, software y mantenimiento.
- Definir una estrategia de despliegue escalonada, priorizando casos de uso de alto impacto.

Clientes y Usuarios Finales:

- Validar que el modelo aporta valor en términos de reducción de tiempos de diseño y mejora en la selección de materiales

Evaluar las características y ventajas de los principales proveedores como: AWS, Azure, GCP e IBM Watson.

Se analizaron cuatro plataformas principales: AWS, Azure, Google Cloud (GCP) e IBM Watson, considerando facilidad de uso, escalabilidad, costos y servicios ofrecidos.

Criterio	AWS (SageMaker)	Azure	GCP (Vertex AI)	IBM Watson ML
Facilidad de uso	Media (requiere configuración inicial)	Alta (integra con herramientas de Microsoft)	Alta (interfaz simplificada)	Media (requiere adaptación)
Escalabilidad	Alta (infraestructura flexible)	Alta (ajuste automático de recursos)	Alta (entrenamiento distribuido)	Media (menos opciones de escalado)
Servicios específicos de ML	Amplia variedad (AutoML, Reinforcement Learning, etc.)	AutoML, soporte para modelos en contenedores	AutoML, modelos preentrenados y personalizados	AutoAI, modelos preentrenados
Costos	Competitivo (pago por uso, instancias reservadas)	Competitivo (uso por consumo, beneficios con licencias MS)	Económico (mejor costo en AutoML y TPU)	Más costoso en grandes volúmenes
Seguridad y cumplimiento	Alta (certificaciones, IAM avanzado)	Alta (cumple estándares empresariales)	Alta (cumple con GDPR, HIPAA, etc.)	Alta (soporte a nivel corporativo)
Compatibilidad e integración	Amplia (TensorFlow, PyTorch, Kubernetes, etc.)	Excelente para entornos Microsoft (Power BI, SQL Server)	Buena (BigQuery, TensorFlow, Jupyter)	Integración con IBM Cloud y SAP

Proporcionar una justificación fundamentada para la elección del proveedor que consideren más apropiado.

Tras analizar los criterios mencionados, la mejor opción con base en la flexibilidad, escalabilidad y variedad de servicios de ML, **AWS (Amazon SageMaker)** y **GCP (Vertex AI)** destacan como las opciones más sólidas.

- **AWS SageMaker** es ideal para proyectos que requieren personalización, control avanzado sobre el entrenamiento y despliegue del modelo, y una infraestructura escalable con integración de múltiples herramientas.
- **GCP Vertex AI** ofrece un enfoque más automatizado con herramientas optimizadas para ML y costos más bajos en entrenamiento de modelos con TPUs. Es una excelente opción si el equipo busca una solución más accesible y optimizada para ML sin tanta configuración manual.

Si se prioriza flexibilidad y control, AWS SageMaker es la mejor opción. Si se busca simplicidad y costos más bajos, GCP Vertex AI es más conveniente.

Conclusión general

El modelo desarrollado representa un avance significativo en la predicción de tensiones en materiales compuestos, combinando técnicas avanzadas de Deep Learning con simulaciones FEA. Si bien los resultados son prometedores y permiten considerar su implementación en producción, es recomendable seguir iterando la optimización y validación del modelo para maximizar su impacto y asegurar una adopción en la industria automotriz.