



Investigación Sobre El Uso De Aprendizaje Automático Para La Detección Predictiva De Errores Y Pruebas Automatizadas.

Sebastián Cascante Badilla

David Cordero Navarro

Manfred Briceño Álvarez

Universidad CENFOTEC

Introducción a la Ingeniería del software

Rosibel Enríquez Vargas

Fecha: septiembre 2025

Tabla de Contenido

Introducción.....	3
Aprendizaje automático.....	4
Aplicación a la detección predictiva de errores.....	5
Automatización de pruebas.....	6
Estado actual en el ámbito laboral	5
Perspectivas para el futuro.....	5
Tendencias para su evolución	5
Soluciones actuales disponibles en el mercado	5
Plataformas.....	6
Herramientas	6
Conclusiones	7
Referencias.....	8

Introducción

En las últimas décadas, el aprendizaje automático (*machine learning*) ha emergido como una disciplina fundamental dentro del campo de la inteligencia artificial, con aplicaciones que trascienden sectores industriales y académicos. Diversos estudios han documentado cómo esta tecnología permite a los sistemas computacionales identificar patrones complejos, adaptarse a nuevos datos y tomar decisiones autónomas sin intervención humana directa (Jordan & Mitchell, 2015).

Corporaciones globales como Google, Facebook, Apple, Amazon e IBM han destinado recursos significativos al desarrollo de algoritmos de *machine learning*, motivadas por su potencial para optimizar procesos operativos, personalizar servicios y fortalecer la toma de decisiones basada en datos. Estas inversiones no solo reflejan una tendencia de mercado, sino una transformación estructural en la forma en que se concibe la innovación digital.

Más allá de su popularización como término técnico, el *machine learning* se ha integrado de manera profunda en la vida cotidiana. Su presencia es evidente en asistentes virtuales, motores de recomendación, sistemas de detección de fraude, filtrado de contenido no deseado y diagnósticos médicos asistidos por algoritmos. Estas aplicaciones demuestran cómo el aprendizaje automático ha dejado de ser una herramienta experimental para convertirse en un componente esencial de la infraestructura tecnológica contemporánea.

Desde una perspectiva investigativa, el avance del *machine learning* plantea interrogantes sobre la ética algorítmica, la transparencia en la toma de decisiones automatizadas y la necesidad de marcos regulatorios que garanticen su uso responsable. En este contexto, el estudio crítico de sus implicaciones sociales, económicas y epistemológicas se vuelve indispensable para comprender su verdadero impacto en la sociedad global. (Raschka, S. & Mirjalili, V., 2019)

Aprendizaje automático

El aprendizaje automático se puede definir como una rama de la inteligencia artificial encargada de desarrollar métodos capaces de construir modelos analíticos de forma automática partiendo de datos específicos que se usan para su entrenamiento. Estos datos sirven para que el programa aprenda patrones o reglas y resuelva tareas concretas sin que las instrucciones estén programadas.

Se puede decir entonces que el aprendizaje autónomo es cuando solicitamos a un robot o maquina (programas) que se comporten como lo haría un humano aprendiendo. Es decirle les permite a los programas implementar mejoras por sí mismos mediante varios tipos de aprendizaje.

Se dividen en distintas modalidades tales como:

- Supervisada (datos etiquetados, “feedback” directo y predicción de resultados futuros).
- No supervisada (sin etiquetas, sin “feedback” y encuentra estructuras ocultas en los datos).
- Semi-supervisada
- Por refuerzo (Proceso de decisión, sistema de recompensa y aprende serie de acciones).

Actualmente se incluyen técnicas de Deep Learning (aprendizaje profundo) para ayudar con el manejo de los grandes volúmenes de datos. (Raschka, S. & Mirjalili, V., 2019)

Aplicación a la detección predictiva de errores

El sesgo es la diferencia entre el valor medio predicho por el modelo y el valor real. Según Pedro Domingos, mide la tendencia a aprender de forma sistemática cosas equivocadas y de la misma manera. Un alto sesgo da lugar a modelos subajustados (underfitting), por lo que el sesgo debe mantenerse bajo para asegurar que el modelo se ajusta a la realidad. Por otro lado, la varianza mide la tendencia a aprender cosas irrelevantes o muy peculiares para el objetivo. Una varianza alta indica que el modelo está sobreajustado (overfitting), lo cual afecta negativamente su capacidad de generalización. Por tanto, tanto el sesgo como la varianza deben controlarse para lograr un modelo de aprendizaje automático eficaz y equilibrado. (Soria Olivas, E. Sánchez-Montañés Isla, M. A. & Gamero Cruz, R., 2023)

En esta área se presenta un estudio que compara distintos algoritmos de clasificación como Random Forest y SVM, estos mostraron mejores resultados en términos de precisión a la hora de predecir, por lo que se concluye que este tipo de modelos identifica más fácilmente los módulos con riesgo alto, lo que permite enfocar los esfuerzos de aseguramiento a las áreas más vulnerables del sistema. Lo que incrementa la calidad del software y reduce los costos de mantenimiento en comparación a cuando las fallas del sistema se detectan tarde. (Khalid, Ahmad W., Anwar W., Arif A. & Khan, M. A., 2023)

Automatización de pruebas

En el ámbito de la evaluación de modelos de aprendizaje automático, uno de los enfoques más tradicionales consiste en el uso de segmentos de prueba estáticos para realizar análisis fuera de línea. Estos conjuntos de prueba permiten establecer un punto de referencia confiable que facilita la comparación objetiva entre distintos modelos. La naturaleza estática de estos segmentos es fundamental, ya que garantiza la consistencia metodológica en los procesos de validación. De lo contrario, si los modelos se evalúan sobre conjuntos de prueba distintos, se

compromete la validez comparativa de los resultados, dificultando la interpretación rigurosa del rendimiento relativo entre ellos. Esta problemática subraya la importancia de diseñar protocolos de evaluación estandarizados que aseguren la reproducibilidad y la equidad en los estudios comparativos de modelos predictivos.

En esta área el aprendizaje automático demuestra ser una herramienta valiosa para optimizar procesos como la selección de casos de prueba, en este análisis se destaca que los árboles de decisión, los SVM y las redes neuronales permiten el incremento de la eficiencia a la hora de detectar defectos, lo que reduce el costo de ejecución de pruebas haciendo enfoque a los casos relevantes luego de los cambios al software.

Estado actual en el ámbito laboral

Hoy en día hay muchas organizaciones que ya incorporan técnicas de aprendizaje automático e inteligencia artificial en los procesos de pruebas automatizadas, entre las principales razones de su uso están:

- La predicción de errores
- La decisión de que casos de prueba tienen más posibilidad de revelar defectos
- Generación de pruebas automáticas adaptables al comportamiento del software • Análisis de fallos

Esto permite detectar anomalías con más facilidad, optimizar los procesos, reducir tiempos y mejorar la confiabilidad.

Perspectivas para el futuro

Se tienen expectativas altas con respecto al aprendizaje automático, buscando su continua escalada y maduración. Entre lo que se espera sea capaz de hacer se encuentran las pruebas auto sanables para evitar roturas tan constantes, la integración y entrega más fuerte y continúa permitiendo la corrección más rápida de errores, que los modelos predictivos se expliquen de mejor manera para entender cómo y por qué detecta los errores, un mayor uso del Deep Learning para reducir la cantidad de datos necesarios, las pruebas en tiempo real mediante el monitoreo continuo y la detección de anomalías en entornos de producción y mayor cantidad de diferentes tipos de pruebas como las de rendimiento, seguridad, usabilidad entre otros.

Tendencias para su evolución

Entre las tendencias de su evolución se encuentran las pruebas generadas por IA basadas en uso real, el testing predictivo para hallar los fallos antes de ejecutar las pruebas, énfasis en la fiabilidad asegurando que los modelos resistan cambios y las aplicaciones híbridas que combinan el aprendizaje automático con las reglas tradicionales para tener enfoques más seguros. (Gülen K. 2025. Aprendizaje automático en pruebas de software)

Soluciones actuales disponibles en el mercado

En la actualidad el aprendizaje automático se ha integrado en plataformas y herramientas de pruebas de software con el objetivo de la mejora de la detección de errores y la automatización de pruebas. Estas soluciones permiten identificar los módulos del

software con mayor probabilidad de fallar según historial, generar y ejecutar casos de pruebas según cambios en la interfaz y la lógica del sistema y el trabajo en conjunto con los sistemas de integración acelerando los ciclos de prueba y reduciendo los errores en producción.

Plataformas

Entre las plataformas más usadas se encuentran:

- H2O.ai: para la creación de modelos predictivos
- Microsoft Azure ML: para un entorno completo en la nube que se puede integrar con entornos de desarrollo profesionales.

Herramientas

En cuanto a las herramientas de automatización se pueden destacar:

- Selenium: para pruebas de aplicaciones web
- Katalon Studio: combinación de pruebas automatizadas con capacidades de ML para adaptarse a los cambios que puedan existir en el software

Estas soluciones son representantes de la actual tendencia de combinar la automatización de las pruebas con inteligencia artificial, buscando bajar los tiempos de desarrollo y aumentar la calidad de los sistemas. (IEBSchool. 2023. Herramientas de Business Intelligence y Big Data).

Conclusiones

El aprendizaje automático representa un elemento transformador en el aseguramiento de software, al facilitar la predicción de errores, la priorización de pruebas y el análisis automatizado de fallos. Su aplicación permite enfocar los esfuerzos en los módulos más vulnerables del sistema, incrementando la eficiencia y reduciendo los costos asociados a fallos tardíos. Además, la aplicación de estas técnicas en los procesos de prueba facilita la optimización de los recursos, la mejora de la cobertura y el incremento de la confiabilidad de los sistemas, con un impacto positivo en la práctica profesional. Las expectativas para el futuro del aprendizaje automático son crear pruebas auto sanables, continuar insertando los modelos predictivos y hacer que sean más explicativos en sus decisiones, lo cual facilitará aún más la comprensión de cómo y por qué se identifican los errores.

En consecuencia, el monitoreo en tiempo real y la detección temprana de errores asegurarán una cobertura más robusta y efectiva en ambientes de producción, mejorando la calidad del software y la eficacia de los procesos. El aprendizaje automático se posiciona como un componente indispensable en la evolución de la ingeniería de software, no solo optimizando los procesos existentes, sino también permitiendo la innovación en técnicas de prueba y aseguramiento. Su implementación contribuye a sistemas más confiables, eficientes y adaptables, representando un avance muy significativo hacia un desarrollo de software más robusto y sostenible en los entornos profesionales.

Referencias

Huyen, C., & Torres, M. Á. (Trad.). (2023). *Diseño de sistemas de machine learning* (1.^a ed.). Marcombo. <https://elibro.net/es/ereader/ucenfotec/281670?page=415>

Khalid, A., Ahmad, W., Anwar, W., Arif, A., & Khan, M. A. (2023). Software defect prediction analysis using machine learning. *Sustainability*, 15(6), 5517. <https://doi.org/10.3390/su15065517>

Raschka, S., & Mirjalili, V. (2019). *Python machine learning: Aprendizaje automático y aprendizaje profundo con Python, scikit-learn y TensorFlow* (2.^a ed.). Marcombo. <https://elibro.net/es/lc/ucenfotec/titulos/280024>

Sarker, I. H. (2021). Machine learning: Algorithms. *SN Computer Science*, 2(3), 1–21. <https://link.springer.com/article/10.1007/s42979-021-00592-x>

Soria Olivas, E., Sánchez-Montañés, M. A., & Gamero Cruz, R. (2023). *Sistemas de aprendizaje automático* (1.^a ed.). RA-MA Editorial. <https://elibro.net/es/ereader/ucenfotec/235049?page=1>