Nombre:

Bryan Roberto Quispe Romero

Materia:

Desarrollo de Software Seguro

NRC:

27894

Tutor:

Angel Geovanny Cudco Pomagualli

Fecha:

24-Nov-2025

Resumen Descriptivo

2.1. Minería de Datos aplicados al Desarrollo de Software Seguro

2.1.1. Conceptos

La minería de datos es una herramienta poderosa en el desarrollo seguro de software, proporcionando mejoras significativas en la detección de vulnerabilidades , optimización del proceso , estimación de esfuerzos y toma de decisiones informadas

2.1.2. Aplicaciones de la Minería de Datos

Las aplicaciones de la minería de datos en el desarrollo de software seguro abarcan diversas áreas críticas:

Detección de Vulnerabilidades y Reparación de Fallos: La minería de datos permite identificar patrones que revelan vulnerabilidades potenciales en el código, facilitando su detección temprana y reparación efectiva.

Optimización del Proceso de Desarrollo: Esta aplicación la asignación y programación de recursos en la pueden optimizar para mejorar la gestión del desarrollo de software

Permiten análisis previos, monitoreo del proyecto y evaluaciones posteriores

Mejora de la Productividad y Calidad: La mineria de datos ayuda a extraer conocimiento valioso que puede mejorar la productividad y calidad del software

Análisis de Datos de Repositorios de Software: La minería de datos en repositorios de software, como los de proyectos de Código abierto, permite entender mejor el proceso de desarrollo y revelar problemas potenciales, contribuyendo.

2.1.3. Técnicas Comunes

Las técnicas de minería de datos aplicadas a la seguridad de software incluyen:

Clasificación y Regresión: La clasificación utiliza algoritmos como árboles de decisión (C4.5, J48), Naive Bayes y Random Forest para categorizar datos en diferentes clases, permitiendo identificar tipos de vulnerabilidades o patrones de código inseguro. Por su parte, la regresión emplea modelos lineales, no lineales, múltiples y logísticos para predecir valores continuos, como el esfuerzo requerido para corregir vulnerabilidades o la probabilidad de ocurrencia de fallos.

Agrupamiento (Clustering): Algoritmos como K-means permiten agrupar elementos similares, facilitando la identificación de patrones comunes en vulnerabilidades o en el comportamiento del código.

Reglas de Asociación: Mediante algoritmos como Apriori, es posible descubrir relaciones entre diferentes elementos del código o del proceso de desarrollo, identificando combinaciones de factores que pueden conducir a vulnerabilidades.

2.1.4. Casos Prácticos

Detección de Vulnerabilidades en Proyectos Open Source: GitHub y otras plataformas utilizan técnicas de minería de datos para analizar millones de repositorios y detectar patrones de código inseguro.

Análisis de Comportamiento de Malware: Empresas de ciberseguridad como Symantec y McAfee aplican clustering (K-means) para agrupar variantes de malware según sus características comportamentales, permitiendo identificar nuevas amenazas basándose en similitudes con familias de malware conocidas.

Optimización de Pruebas de Seguridad: Google utiliza minería de datos para optimizar sus procesos de testing, identificando mediante reglas de asociación (Apriori) qué combinaciones de cambios en el código tienen mayor probabilidad de introducir vulnerabilidades, priorizando así las pruebas de seguridad.

Análisis Forense de Brechas de Seguridad: Organizaciones financieras emplean técnicas de clasificación para analizar logs de sistemas y detectar patrones anómalos que indiquen intentos de intrusión o brechas de seguridad, permitiendo respuestas rápidas ante incidentes.