

1. Datos Informativos Módulo: Unidad 2

Nivel: Séptimo

Fecha: 05/05/2025

Tema: Métricas de Defectos – Usabilidad en el software en sistemas de NLP

Integrantes: Aupas Antony, Baraja Cristian, Basantes Geovanny

2. Objetivos

Capacitar a los participantes en el uso e interpretación de métricas de defectos, usabilidad y experiencia de usuario, aplicadas al análisis de calidad del software.

3. Contenido

3.1. Introducción

Las **métricas de defectos** son herramientas esenciales en el desarrollo de software, ya que permiten evaluar su calidad mediante el seguimiento, análisis y control de los errores detectados durante todo el ciclo de vida del producto. Estas métricas ayudan a identificar debilidades en el proceso de desarrollo, facilitando así la mejora continua del software. Su correcta aplicación permite prevenir fallos futuros y garantizar productos más confiables y eficientes para los usuarios finales.

Entre las métricas más comunes se encuentra la **Densidad de Defectos**, que se calcula dividiendo el número de defectos encontrados entre el tamaño del software (por ejemplo, en líneas de código). Esta métrica permite estimar cuán propenso es el sistema a presentar errores. Otra métrica relevante es la **Eficiencia en la Eliminación de Defectos (Defect Removal Efficiency - DRE)**, la cual mide qué porcentaje de defectos fue detectado y corregido antes de que el software llegue a producción. Por otro lado, la **Fuga de Defectos (Defect Leakage)** indica qué porcentaje de errores no fue detectado durante las pruebas y apareció directamente en el entorno de producción, lo cual representa un riesgo considerable para la calidad del producto.

En el ámbito de la interacción con el usuario, las **métricas de usabilidad y experiencia de usuario (UX)** son fundamentales para evaluar la eficiencia, efectividad y satisfacción con que los usuarios interactúan con un sistema. Aunque frecuentemente se confunden, la **usabilidad** se enfoca en la facilidad con la que una persona puede lograr un objetivo usando una herramienta, mientras que la **experiencia de usuario (UX)** abarca aspectos más

amplios como emociones, percepciones, actitudes y satisfacción general del usuario al utilizar el sistema.

Existen diversas métricas para evaluar estos aspectos. Una de las más conocidas es la **SUS (System Usability Scale)**, un cuestionario estándar de 10 ítems que proporciona un puntaje global de usabilidad. El **NPS (Net Promoter Score)** mide la probabilidad de que un usuario recomiende el sistema a otros, mientras que el **CSAT (Customer Satisfaction Score)** evalúa la satisfacción inmediata luego de una experiencia específica. Finalmente, el **CES (Customer Effort Score)** se enfoca en medir el esfuerzo percibido por el usuario al completar una tarea, siendo clave para identificar barreras en el uso del sistema. Estas métricas permiten tomar decisiones fundamentadas para mejorar tanto el diseño como la funcionalidad de las interfaces.

Métricas de defectos y de usabilidad/UX en un caso de estudio.

Caso de Estudio – LinguaMind

Imaginemos el caso de un equipo de desarrollo que ha creado una **plataforma web llamada “LinguaMind”**, una aplicación que permite a usuarios interactuar con un asistente virtual capaz de resumir textos, traducir documentos, responder preguntas técnicas y generar contenido escrito de manera coherente y natural. La plataforma está diseñada para ser usada tanto por estudiantes como por profesionales que necesitan optimizar tareas relacionadas con la lectura, redacción y comprensión de textos.

Durante su fase de pruebas, el equipo detecta varios defectos funcionales, como errores en el guardado de formularios, validación inadecuada de contraseñas y fallas graves como cierres inesperados al ejecutar tareas de análisis semántico en la pantalla de pagos.

Paralelamente, se realiza una evaluación de la experiencia de usuario utilizando una **encuesta SUS**, con el fin de medir la percepción de usabilidad general del sistema. Se complementa con análisis de métricas de uso basadas en interacciones reales con la plataforma.

Este caso representa una oportunidad realista para aplicar **métricas de defectos y métricas de usabilidad y UX**, evaluar el rendimiento del producto, identificar áreas críticas de mejora y proponer soluciones tanto a nivel técnico como en la interfaz de usuario. El

análisis permitirá comprender cómo la calidad de un sistema de NLP no solo depende del rendimiento de sus modelos de lenguaje, sino también de la robustez del software que lo soporta y de la experiencia de quienes lo utilizan.

ID	Título	Severidad	Estado	Asignado a	Fecha creación	Fecha resolución	Tiempo (hrs)
101	Error al guardar formulario de texto	Alta	Cerrado	Juan Pérez	2025-05-14	2025-05-15	5
102	Desbordamiento en lista de usuarios	Media	Cerrado	Carla Ruiz	2025-05-14	2025-05-16	8
103	Fallo en validación de contraseña	Alta	Abierto	—	2025-05-15	—	—
104	Texto no legible en modo oscuro	Baja	Cerrado	Pedro Gómez	2025-05-15	2025-05-15	2
105	Crash al abrir pantalla de pagos	Crítica	Abierto	—	2025-05-16	—	—

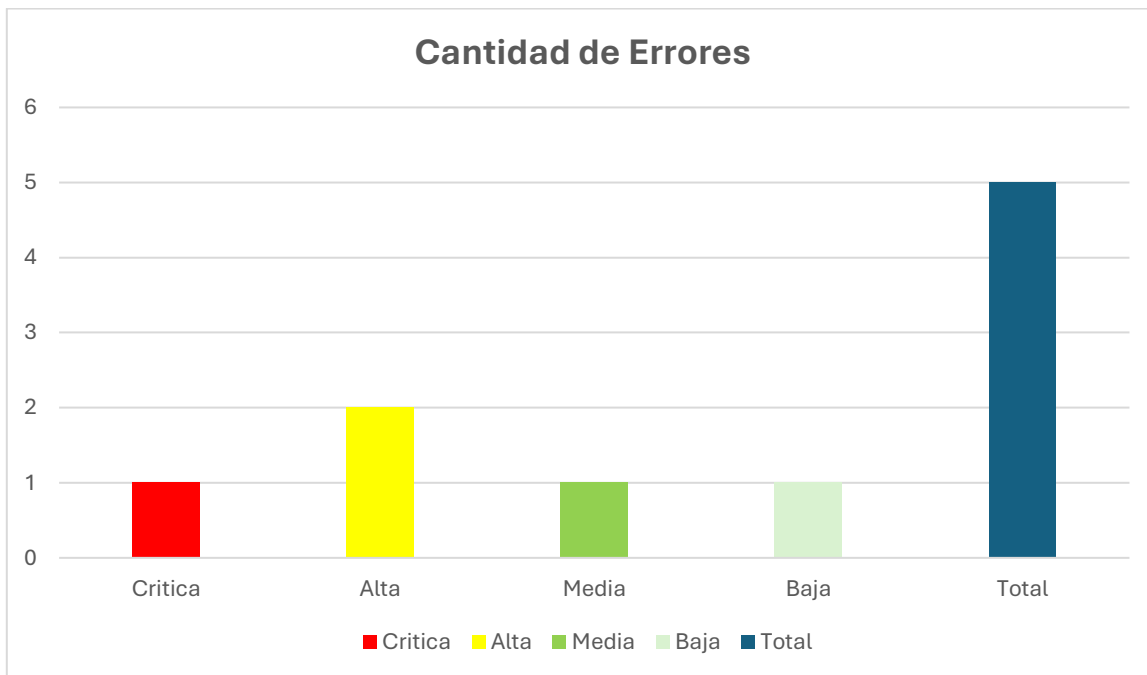
Resultados Encuesta SUS ficticia:

Usuario	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
U1	4	2	4	1	5	2	4	1	5	2
U2	3	1	4	2	4	3	5	2	4	1
U3	4	2	5	1	4	2	4	1	5	2
U4	5	2	3	2	5	2	4	2	4	2
U5	3	3	4	2	4	3	4	2	5	1

1. Métricas de uso (Google Analytics)

Métrica	Valor
Usuarios activos diarios	134
Duración promedio de sesión	3 min 45 s
Tasa de rebote	40 %
Número promedio de pantallas vistas	4.2
Porcentaje de tareas completadas	82 %
Tiempo promedio en completar tarea	2 min

Asumimos un tamaño del software de 1000 líneas de código (LOC).



En total tuvimos 5 defectos de los cuales 101, 102, 105 fueron cerrados y los dos restantes 103 y 105 son defectos abiertos es decir que aún no ha sido solucionado

Cálculo de métricas:

Métrica	Fórmula	Valor
Densidad de defectos	$5 / 1000$	0.005
DRE	$3 / 5 * 100$	60%
Defect Leakage	$2 / 5 * 100$	40%

Evaluación de Métricas de Usabilidad /UX

SUS (System Usability Scale)

Con base en la tabla, el puntaje promedio SUS de los 5 usuarios es **76** (en una escala de 0 a 100). Un puntaje mayor a 68 indica buena usabilidad. LinguaMind cumple con estándares, pero aún puede mejorar en tareas específicas.

CSAT (Customer Satisfaction Score)

Supongamos que 4 de los 5 usuarios calificaron su experiencia con 4 o 5 estrellas (satisfacción positiva). Buena satisfacción general, aunque un 20 % de usuarios no quedaron completamente conformes.

$$CSAT = (4/5) \times 100 = 80 \%$$

CES (Customer Effort Score)

Según los datos, el **tiempo promedio para completar tareas es 2 minutos**, lo cual sugiere bajo esfuerzo. Además, la tasa de tareas completadas es **82 %**, los usuarios encuentran que las tareas se realizan sin mucho esfuerzo, lo cual mejora la fidelidad al producto.

Tasa de Rebote (Bounce Rate)

El 40 % de los usuarios abandonan la plataforma tras ver una sola pantalla. Esta cifra es moderada, pero puede indicar problemas de diseño inicial o falta de contenido atractivo en la primera interacción.

Mejores Técnicas

Optimización del rendimiento de carga inicial: Reducir el tiempo de carga de la aplicación inicial (pantalla principal), ya que puede influir en la tasa de rebote.

Gestión avanzada de errores y logs: Implementar un sistema robusto de captura de errores y fallas (con trazabilidad en tiempo real tipo Sentry o LogRocket).

Refactorización del módulo de autenticación: Ya que hubo fallos con la validación de contraseñas, reestructurar la lógica de entrada de usuarios con pruebas unitarias.

Mejoras de Experiencia de Usuario (UX)

Rediseño del modo oscuro: Solucionar el problema de texto ilegible en modo oscuro, ajustando el contraste y revisando la paleta de colores.

Asistente de ayuda contextual: Incluir pequeños tips emergentes o un chatbot guía que ayude a nuevos usuarios a navegar el sistema con menos esfuerzo.

Optimizar la pantalla de bienvenida y tareas iniciales: Rediseñar la pantalla inicial con contenido claro y visualmente atractivo que motive a los usuarios a explorar más, reduciendo la tasa de rebote.

Conclusión:

Las métricas de defectos y de usabilidad son herramientas fundamentales en la evaluación de la calidad del software. Las primeras permiten identificar la cantidad y severidad de errores durante el ciclo de desarrollo, facilitando la mejora continua y la prevención de fallos en producción. A través de indicadores como la densidad de defectos, la eficiencia de eliminación (DRE) y las fugas de defectos (Leakage), los equipos pueden tomar decisiones informadas sobre el proceso de pruebas y control de calidad.

Por otro lado, las métricas de usabilidad y experiencia del usuario (UX) proporcionan información valiosa sobre cómo los usuarios perciben e interactúan con un sistema.

Instrumentos como SUS, NPS, CSAT y CES permiten evaluar tanto la facilidad de uso como la satisfacción y esfuerzo del usuario, factores cruciales para el éxito del producto. En conjunto, estas métricas no solo mejoran la calidad técnica, sino que aseguran que el software sea funcional, confiable y agradable para quienes lo utilizan.

Bibliografía:

- Pressman, R. S., & Maxim, B. R. (2020). *Ingeniería del software: Un enfoque práctico* (8.^a ed.). McGraw-Hill.
- Nielsen, J. (2012). *Usability 101: Introduction to usability*. Nielsen Norman Group.
<https://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research* (2nd ed.). Morgan Kaufmann.
- ISO. (2018). *ISO 25010:2011 – Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models*. International Organization for Standardization.
- Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.