SCReeD Dataset Declaration Form

Dataset name:*   Vulnerable Functions

Dataset version:*   V1

Dataset URL:*   https://figshare.com/s/c1b8fe7dca01b698afe4

Creation date:*   2024-01-22

Last update:*   2024-01-22

Author(s):*   Triet Huynh Minh Le, M. Ali Babar

Author(s) affiliation(s):   The Centre for Research on Engineering Software Technologies, The University of Adelaide, Cyber Security Cooperative Research Centre.

Author contact(s):*   triet.h.le@adelaide.edu.au, ali.babar@adelaide.edu.au

Keywords:*   Software engineering, software vulnerability

Description/background:*   Investigates ML for function-level SV assessment, emphasising the importance of vulnerable statements in performance on dataset with 1,782 functions, 429 SVs, 200 projects.

Dataset funding: N/A

Attribute details:*   Vulnerability Fixing Commits

Intended target (if specified): Click or tap here to enter text.

Format:*   parquet

License:*   N/A

Standard compliance: Follow the compliance of the origin source – NVD , GitHub Advisory Database, and VulasDB

Type:*   Vulnerability Fixing Commits

Size:*   23.43 MB

Availability: Open

Data status: Raw

Data provenance:*   The dataset is collected from NVD, GitHub Advisory Database, and VulasDB

Source computing infrastructure:*   Simulated environment

Accompanying program(s)/script(s): https://figshare.com/s/c1b8fe7dca01b698afe4

Software installer or VM for replication: N/A

Generated or captured via:*   Other, please specify: The dataset is collected from NVD, GitHub Advisory Database, and VulasDB

Category/categories:*   Software engineering, Vulnerability Assessment, Software Security, Machine Learning, Software Vulnerability, Mining software Repositories

Published in: Figshare

Open research question(s) (if any):   N/A

Potential use case(s) or application area(s): Vulnerability Assessment, Software Security, Machine Learning, Software Vulnerability

Data access control:*   Organizational, please specify:

Data retention period:*   N/A

Data validation/checksum:* N/A

GDPR compliance:*   No
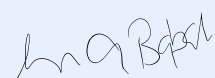
Consent: N/A

Ethics approval: N/A

Ethics considerations: N/A

       * denotes mandatory field

☒   I confirm that I have read, understood, and agreed to the submission guidelines, policies, and submission declaration.

☒   I confirm that the contributors of the dataset have no conflict of interest to declare.

☒   I agree to take public responsibility for my dataset's contents.

Signature of Corresponding Author

(signed on behalf of all contributors)

Date: Click or tap here to enter text.