



US 20240214422A1

(19) **United States**

(12) **Patent Application Publication**
SONG

(10) **Pub. No.: US 2024/0214422 A1**

(43) **Pub. Date: Jun. 27, 2024**

(54) **MACHINE LEARNING-BASED
HARMFUL-WEBSITE CLASSIFICATION
METHOD**

Publication Classification

(51) **Int. Cl.**
H04L 9/40 (2006.01)
H04L 41/16 (2006.01)
(52) **U.S. Cl.**
CPC **H04L 63/1483** (2013.01); **H04L 41/16**
(2013.01)

(71) Applicant: **DATAKOBOLD CO., LTD.**,
Seongnam-si Gyeonggi-do (KR)

(72) Inventor: **Nam Goo SONG**, Suwon-si
Gyeonggi-do (KR)

(73) Assignee: **DATAKOBOLD CO., LTD.**,
Seongnam-si Gyeonggi-do (KR)

(21) Appl. No.: **18/537,884**

(22) Filed: **Dec. 13, 2023**

(30) **Foreign Application Priority Data**

Dec. 27, 2022 (KR) 10-2022-0185197

(57) **ABSTRACT**

A machine learning-based harmful-website classification method performed by a main server includes performing tokenization, by the main server, by extracting and pre-processing an HTML source code of a website by accessing to the website, vectorizing each token according to a preset algorithm, inputting each vector value is input into a machine learning model, and determining whether the website is a harmful website.

