AN ADVERSARIAL-AWARE EMAIL SPAM DETECTOR MECHANISM

Rasanga Abeykoon

A dissertation submitted in partial fulfilment of the requirement for BEng (Honors) degree in Software Engineering

Department of Computing

Informatics Institute of Technology, Sri Lanka

in collaboration with

University of Westminster, UK

2021

Abstract

Operating in dynamic real world spam domain requires more robust adversarial-aware designs as it is identified as never ending game between learner and attacker. While most of the researchers are focused on either evasion or poison resistant, there are only few researches which are focused on both test and train time attacks. In this research, Ensemble approach has been applied to combine multiple learners which were secured against single attack based on attack influence. It is compared with individual learners and results have shown an attack which is targeted on single learner does not imply the same on Ensemble methods. The proposed method provides motivation for future works of area in ensemble learning under spam detection.

Keywords: Adversarial machine learning, electronic mail, Security