

PREDICTING THE CYBER SECURITY DATA LEAKS

A project report submitted in partial fulfillment of the requirements
for the award of the Degree of

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE & ENGINEERING

Submitted by,

M.MONALI(16811A0550)

D.ANITHA(16811A0521)

J.HEMANTH(16811A0532)

E.UPENDRA KUMAR(16811A0522)

Under the esteemed guidance
of

Ms.D.PRASANNA KUMARI(M.Tech)
Assistant Professor



Department of Computer Science & Engineering

AVANTHI INSTITUTE OF ENGINEERING AND TECHNOLOGY

(Affiliated to JNTU Kakinada & Approved by AICTE)

TAMARAM, MAKAVARAPALEM, NARSIPATNAM-531113

VISAKHAPATNAM (DIST)

(2016-2020)

AVANTHI INSTITUTE OF ENGINEERING AND
TECHNOLOGY

(Affiliated to JNTU Kakinada & Approved by AICTE)

TAMARAM, MAKAVARAPALEM, NARSIPATNAM-
531113VISAKHAPATNAM (DIST)



CERTIFICATE

This is to certify that the Project Report entitled “Predicting the cyber security data leaks” being submitted, M.MONALI(16811A0550), D.ANITHA(16811A0521), J.HEMANTH(16811A053), E.UPENDRA KUMAR(16811A0522), in partial fulfilment of the requirements for the degree of B.Tech (C.S.E) in Department of Computer Science & Engineering, at AVANTHI INSTITUTE OF ENGINEERING AND TECHNOLOGY affiliated by Jawaharlal Nehru Technological University Kakinada, is a record of bonafide work carried out by them under my guidance and supervision.

The results embodied in this thesis have not been submitted to any university or institute for the award or any degree of diploma.

D. Prasanna Kumari

Ms.D.Prasanna Kumari
Project Guide

P.M. Manohar

P.M.Manohar
Head of the Department

External Examiner

ABSTRACT

This study offers a first step toward understanding the extent to which we may be able to predict cyber security incidents (which can be of one of many types) by applying machine learning techniques and using externally observed malicious activities associated with network entities, including spamming, phishing, and scanning, each of which may or may not have direct bearing on a specific attack mechanism or incident type. Our hypothesis is that when viewed collectively, malicious activities originating from a network are indicative of the general cleanness of a network and how well it is run, and that furthermore; collectively they exhibit fairly stable and thus predictive behavior over time. To test this hypothesis, we utilize two datasets in this study