Predictive Model for Cyber Attacks to Assess Financial Exposure for IoHT / IoMT

by Apache Sanders

B.S. in Business Administration, May 2007, Tuskegee University
B.S. in Management Science, May 2007, Tuskegee University
M.S. in Information Systems, May 2010, Strayer University
M.S. in Systems Engineering, May 2018, Johns Hopkins University

A Praxis submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Engineering

August 31, 2023

Praxis directed by

Joel Goodman
Professional Lecturer in Engineering and Applied Science

The School of Engineering and Applied Science of The George Washington University certifies that Apache Sanders has passed the Final Examination for the degree of Doctor of Engineering as of August 1, 2023. This is the final and approved form of the Praxis.

Predictive Model for Cyber Attacks to Assess Financial Exposure for IoHT / IoMT

Apache Sanders

Praxis Research Committee:

Joel Goodman, Professional Lecturer in Engineering and Applied Science, Praxis Director

Benjamin Stirgwolt, Professional Lecturer in Engineering and Applied Science, Committee Member

Joseph Blackford, Professional Lecturer in Engineering and Applied Science, Committee Chair

© Copyright 2023 by Apache Sanders
All rights reserved.

Dedication

I dedicate this paper to my tribe that consistently encouraged, supported, and uplifted me through this process. It truly does take a village. There were moments throughout this journey when I was not sure that I would be able to finish this daunting task. I want to thank my husband, Joshua, for always supporting my endeavors. My success in all things has resulted directly from your support and selflessness. To my children and grandchildren, you have been beacons of light to inspire and make me smile. You were my biggest why. To my mother, I owe my ability to persevere even in the most challenging times. To my maternal grandparents, especially my grandfather, thank you for being an early example of why education is so important. Memories of sitting on the porch spelling 'stop' and other words with my granddad are forever embedded in my mind and are why pursuing education is a part of my identity. I hope you are in heaven smiling down on me, proud of whom I have become in your honor. To my siblings, nieces, nephews, in-laws, friends, and extended family… I appreciate you more than words could ever convey. Thank you for your unwavering support that enabled me to persevere through challenging moments. I would not have been able to accomplish this without you.

Acknowledgments

I extend my heartfelt gratitude to my Praxis director, Dr. Joel Goodman. Your guidance, encouragement, and unwavering support throughout this entire process have been invaluable. I am deeply appreciative of your willingness to share your vast knowledge, precious time, and valuable experiences, which have enriched my journey significantly. Thank you for being a pivotal source of inspiration and motivation in this endeavor.

Abstract of Praxis

Predictive Model for Cyber Attacks to Assess Financial Exposure for IoHT / IoMT

Cybersecurity incidents for Internet of Health Things (IoHT) and Internet of Medical Things (IoMT) organizations are rising. The Common Vulnerability Score System (CVSS) provides a system for both the prioritization and assessment of the impact of vulnerabilities on an organization. Vulnerability descriptions are provided to a vulnerability database where the issues are analyzed and assigned a score. Those scores can then assess a vulnerability's severity and priority. The problem is that these scores are subjective depending on the entity/individual who evaluates the exposure. This thesis aims to provide a predictive model to IoHT/IoMT leaders to assess the financial impact of cyber-attacks on their perspective organizations; this impact is determined using a combination of CVSS scores and the number of financial records impacted.

Natural language processing (NLP) algorithms can help automate vulnerability severity score predictions based on textual descriptions provided by end users. NLP is helpful because it allows machines to process, analyze, and understand human language, which is essential for ensuring reproducibility. Automated prediction simplifies the process of assigning severity scores to vulnerabilities. Organizations can save time and resources by using NLP for this purpose. It will allow organizations to address and prioritize vulnerabilities promptly.

Keywords: CVSS, IoHT, IoMT, BERT, Cyber Security, NLP, Financial Impact

Table of Contents

Dedication iv

Acknowledgments v

Abstract of Praxis ..... **vi**

List of Figures ..... **x**

List of Tables ..... **xi**

List of Acronyms ..... **xiii**

Chapter 1—Introduction 1

1.1 Background 1

1.2 Research Motivation 2

1.3 Problem Statement 8

1.4 Thesis Statement 8

1.5 Research Objectives 10

1.6 Research Questions and Hypotheses 10

1.7 Scope of Research 11

1.8 Research Limitations 11

1.9 Organization of Praxis 12

Chapter 2—Literature Review 14

2.1 Introduction 14

2.2 Overview of Vulnerability Management 14

2.2.1 Common Vulnerability Analysis and Planning 15

2.2.2 The Common Vulnerability Scoring System 15

2.2.2.1 Common Vulnerability Scoring System Metric Groups 16
2.2.2.2 CVSS Related Research 22
2.3 Overview of Natural Language Processing 25
2.3.1 BERT & BERT Derivative Models 26
2.3.2 NLP Related Research 29
2.4 Assessing Financial Impact 31
2.5 Summary and Conclusions 34
Chapter 3—Methodology 35
3.1 Introduction 35
3.2 CVSS Methodology 36
3.2.1 Text Pre-Processing and Vocabulary Selection 36
3.2.2 Vulnerability Dataset and Model Details 37
3.2.3 Model Details and Evaluation Metrics 40
3.2.4 Calculate the CVSS score Based on Predicted Label Values 42
3.3 Financial Impact Methodology 45
3.3.1 Cost based on Number of Records Impact vs Total Cost Per Record 45
3.3.2 Calculate Total Cost Per Record 46
Chapter 4—Results 54
4.1 Introduction 54
4.2 CVSS Score Prediction Results 54
4.2.1 Model Comparison 54
4.2.2 CVSS Score Prediction Results 58
4.3 Financial Impact Prediction Results 63

4.4 Financial Impact Prediction Results 69

Chapter 5—Discussion and Conclusions 72

5.1 Discussion 72

5.2 Conclusions 73

5.3 Contributions to Body of Knowledge 75

5.4 Recommendations for Future Research 75

References 78

List of Figures

Figure 2-1. CVSS Metric Groups. 18

Figure 2-2. NLP Process. 26

Figure 3-1. Praxis Methodology. 35

Figure 4-1. Proportion of Class from Baseline Research. 54

Figure 4-2. Code Snippet of CVSS Calculations. 59

Figure 4-3. CPR by Industry. 67

List of Tables

Table 1-1. Summary of RQs and Hypotheses 10

Table 2-1. Metrics Definition 17

Table 2-2. Metrics Group Formula – v3.1 CVSS Scores 20

Table 2-3. Severity Category 21

Table 3-1. Class Proportions 39

Table 3-2. Hyperparameters Used to Fine-tune Models 41

Table 3-3. CVSS Metric Numerical Value 43

Table 3-4. Industry Cost Factor (F_I)Metrics Definition 47

Table 3-5. Sensitive Data Retained Cost Factor (F_D) 48

Table 3-6. [$\boldsymbol{axb-1}$] Parameters – Data in Centralized Location 49

Table 3-7. TABLE: [$\boldsymbol{axb-1}$] Parameters – Data in Centralized Location 49

Table 3-8. Cause of Data Breach Cost Factor (F_{BC}) 50

Table 3-9. Sensitive Data Encryption Cost Factor (F_E) 50

Table 3-10. BCM Involvement Cost Factor 51

Table 3-11. Regulatory and Industry Sanctions Cost Factor 51

Table 3-12. Class Action Lawsuit Cost Factor 52

Table 4-1. Results from Baseline Research 54

Table 4-2. Results from CVSS Prediction Model 56

Table 4-3. Balanced Accuracy Comparison 56

Table 4-3. Sample Output 59

Table 4-4. Sample Dataset 60

Table 4-5. Regression Analysis 62

Table 4-6. Total Cost Per Record Calculation 65

Table 4-7. Regression Results – Total Cost Per Record 68

Table 4-8. Final Findings 70

Table 5-1. Summary of RQ, Hypotheses, and Outcomes 73

List of Acronyms

AC	Attack Complexity
ALBERT	A Lite BERT
AV	Attack Vector
BART	Bidirectional and Auto-Regressive Transformer
BCM	Business Continuity Management
BERT	Bidirectional Encoder Representations from Transformers
CALCPR	Class Action Lawsuit Cost Per Record
CC	Credit Card
CIA	Confidentiality, Integrity, Availability
CMCPR	Crisis Management Cost Per Record
CNN	Convolutional Neural Network
CPR	Cost Per Record
CVAP	Common Vulnerability Analysis and Planning
CVE	Common Vulnerability Enumerators
CVSS	Common Vulnerability Score System
DeBERTa	Decoding-enhanced BERT with Disentangled Attention
DistilBERT	Distilled BERT
EHR	Electronic Health Records
EMR	Electronic Medical Records

FDA	Food and Drug Administration
FIRST	Forum of Incident Response and Security Teams
HIPAA	Health Insurance Portability and Accountability Act of 1996
ICPR	Investigation Cost Per Record
IoHT	Internet of Health Things
IoMT	Internet of Medical Things
IoT	Internet of Things
ISS	Impact Sub-Score
MDDT	Medical Device Development Tool
ML	Machine Learning
MLM	Masked Language Modeling
MSE	Mean Squared Error
NIST	National Institute of Standards and Technology
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
NSP	Next Sentence Prediction
NVD	National Vulnerability Database
PCI	Payment Card Industry
PHI	Public Health Information
PR	Privileges Required
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network

RoBERTa	Robustly Optimized BERT Pre-training Approach
S	Scope
SCPR	Sanctions Cost Per Record
SME	Subject Matter Expert
SSN	Social Security Number
SVM	Support Vector Machine
TS	Time Series
UI	User Interaction

Chapter 1—Introduction

1.1 Background

Cyberattacks against the medical industries are increasing, resulting in more successful breaches per company each year, increased cyber threat actors, and varied attack methods used. Social engineering, unpatched software, phishing attacks, and advanced persistent threat are some issues that allow healthcare organizations to remain a target. The healthcare sector is targeted for the lucrative information retained in its systems; electronic health records (EHR) and personal information are often sold on the black market for a high value (Seh et al., 2020a). Several factors make the healthcare sector an easy target for cybercriminals. Some of those issues are limited IT security funding, outdated systems, medical devices with minimal security, and the lack of flexibility in allowing downtime for system updates (US Food and Drug Administration, 2017). These are some of the causes for cyber-attacks continuing at an increasing rate.

The National Vulnerability Database (NVD) was created by the National Institute of Standards and Technology (NIST); this database contains the entries known as Common Vulnerability Enumerators (CVEs) (National Institute of Standards and Technology, 2021a). End-users create entries; NVD uses that information combined with any reference information available to analyze, evaluate, and publish it so that it is readily available for organizations (National Institute of Standards and Technology, 2021a). The CVEs are assessed and assigned a CVSS score, a quantitative representation of the severity that organizations can use to prioritize and address vulnerabilities (Shi et al., 2022). CVSS is a widely used industry standard that provides a numerical score to aid in

the measurement of the severity of exposure or vulnerability. CVSS v3.1 is comprised of three metric groups; those metric groups are Base, Temporal, and Environmental (Shi et al., 2022). NVD has made CVSS calculators available to the public to allow assessments to occur while official scores are being determined. The score provides the ability overall score can be used to rank the vulnerabilities' severity and help determine mitigation strategies.

1.2 Research Motivation

In 2013, Vice President Cheney had his heart defibrillator disconnected from the internet to prevent possible cyber-attacks. In 2017, the FDA discovered security vulnerabilities and recalled 465,000 pacemakers (Cyber Defense Magazine, 2017). Due to the rapid pace at which technology is advancing and increasing internet connectivity, hospitals, clinics, laboratories, and other medical organizations are at risk because of the large amounts of private information stored.

In recent years, there has been an increase in cybercrimes related to the IoMT and IoHT. This is partly due to the growing number of connected medical devices, wearables, and other healthcare-related equipment vulnerable to cyber-attacks. In research conducted in 2017, Walker determined that there are an estimated 10 – 15 connected IoMT devices per hospital bed contained in most US hospitals (Walker, 2018). As the number of connected devices grows, IoMT will continue to be targeted. According to Coventry and Branley, cybercriminals choose to target IoMT and IoHT devices for several reasons, including:

1. Stealing personal health information: Stealing personal health information: Misplacing paper medical records or stolen laptops used to only impact hundreds or thousands of patients, but in a digital era with internet connectivity, the impact grows to millions of people (Coventry & Branley, 2018). Medical records, health insurance information, and other personal health information in electronic systems can be sold on the black market.
2. Ransomware attacks: Cybercriminals may use ransomware to encrypt patient data, demanding payment for the decryption key. Floyd, Grieco, and Reid analyzed data breaches in the U.S. to determine how records were potentially used for financial benefit. The authors mined data on the black web to determine that the average cost of a single record is approximately $1.25 per record. It was also reported that stolen personally identifiable information paired with actual medical information could increase the value tenfold (US Food and Drug Administration, 2017).
3. Disrupting healthcare operations: Cyber-attacks on IoMT and IoHT devices can disrupt hospital and clinic operations. Disrupting operations can have a catastrophic impact on IoMT (Cartwright AJ, 2023). It can lead to delays in patient care and potentially compromise patient safety. This makes IoMT officials more likely to comply with requests to ensure patients and their information remain safe and services will not be interrupted. (Cartwright AJ, 2023).
4. Malicious tampering: Cybercriminals may want to sabotage medical devices with the goal of harming patients. Ponemon estimated that 51% of IoMT organizations listed criminal attacks as the root cause (Ponemon Institute, 2017).

Healthcare providers and device manufacturers must take measures to improve the cybersecurity of their systems. Best practices for improving cybersecurity posture are implementing better security measures, maintaining a consistent schedule with deploying software and firmware updates, and conducting security assessments to identify and address vulnerabilities consistently (Miller et al., 2023). Another important measure is to ensure that healthcare providers and device manufacturers comply with relevant regulations and standards to ensure that patient data is kept secure and private.

NVD analyzes the information provided by entities with the goal of providing accurate, current, and complete information about vulnerabilities to aid companies in vulnerability management. When conducting an analysis, CVEs are assigned to employees for assessment. NVD assumes a worst-case scenario because it is more palatable to a business to overstate the risk than find out later that enough measures were not taken. The review of CVEs is extensive and thorough, which often results in longer processing times and a more thorough examination of the available information. CVSS scores are beneficial because they provide organizations with a standardized method of measuring the severity of vulnerabilities, but assessing the exposures and determining the values to calculate the CVSS score is subjective. It can be influenced by factors such as the amount of available and conflicting information. Through investigation and research, Chen, H. and Ruohonen, J. found that an assessor's estimate of severities could lag by as much as 130+ days from exposure (Lewis et al., 2020). Remediations could take more than five months from the moment the vulnerability is identified (Ruohonen, 2019). Research by Feutrill et al. further exasperated the issue by noting that the average time to

enter CVSS metrics for an identified CVE into the NVD database increased drastically from a single day before 2017 to 19+ days since 2018 (Furnell et al., 2020). This trend suggests that determining CVSS scores is becoming more complex and time-consuming; it highlights the need for efficient methods of measuring the severity of vulnerabilities.

Machine learning algorithms are used in NLP to develop predictive models that can automatically analyze and understand human language (Sarker, 2021). Training machine learning algorithms on historical vulnerability data allows the machines to learn/recognize patterns and relationships between certain words, phrases, and other linguistic features in the vulnerability descriptions. It also allows the machine to assess the corresponding severity scores assigned by human experts. Processing this information allows the machine to predict severity scores for new vulnerabilities based solely on their text descriptions without human interpretation. One way to do this is to use Bidirectional Encoder Representations from Transformers (BERT), a natural language processor model that was developed by Google. BERT is bidirectional, meaning the model processes input tokens in both directions to analyze each input. Tokens are the equivalent to a word, so the analysis process involves studying the words first as a whole and then studying the word before and after each word to determine the context. This feature allows it to have a deep understanding of both the actual context of the word string and the complexity in the human language leading to more accurate NLP predictions (Kenton & Toutanova, 2019). This feature allows the model to capture more complex relationships between the words that are included in a sentence; doing this improves its ability to understand, interpret and generate natural language.

Multiple variations of the BERT models were used to predict CVSS categories via vulnerability descriptions. Descriptions were pre-processed and cleansed using lemmatization, stemming, and stop words, increasing the accuracy of model predictions. Initial research findings support DistilBERT as the optimal model for accurately predicting the components needed to calculate the CVSS score. After scoring vulnerabilities, post-processing is completed to port the model outputs into a mathematical equation to predict CVSS scores (Costa et al., 2022). Access to timely CVSS scores paired with potential financial impact predictions allows organizations to properly evaluate vulnerabilities and assess the risk posed to companies and their stakeholders.

When interpreting text descriptions and determining CVSS scores, NLP provides a solution to reduce processing time and remove subjectivity. BERT has several advantages compared to those traditional machine learning algorithms. The first advantage is that BERT does not require manual feature engineering; not requiring manual feature engineering means that BERT has the ability to understand the data to enable the extraction of meaningful context used to train the model; this process can take a substantial amount of time and is often prone to encountering errors (Balakrishnan et al., 2022). Instead of manual feature engineering, BERT learns feature representations directly from raw text data, which can lead to better performance on tasks occurring downstream. Second, BERT is a pre-trained model; pre-trained models have already been trained on large amounts of text data to understand general language patterns and relationships. This allows BERT to be fine-tuned on smaller datasets. Lastly, BERT

understands the dependencies and relationships between the words in a sentence. Overall, BERT's ability to capture context and relationships between words, its lack of manual feature engineering requirements, and its pre-trained nature make it a powerful tool for various NLP tasks. The growing threat from cyberattacks explains the importance of a tool with BERT's capabilities.

In 2022, the average data breach cost reached a new record high; the data breach cost was reported as USD 4.35 million; this represents a 2.6% increase compared to the previous year, where the average cost stood at USD 4.24 million (IBM Corporation & Ponemon Institute, 2021). The same surge was witnessed in breaches in the healthcare industry totaling USD 10.10 million; this is an increase in total costs of nearly USD 1 million (IBM Corporation & Ponemon Institute, 2021). This increase in cost marks the healthcare sector as the most expensive industry for data breaches for 12 consecutive years (IBM Corporation & Ponemon Institute, 2021). To protect patient information and prevent costly data breaches, healthcare institutions must invest in security measures and stay updated with the latest threats and vulnerabilities. This praxis aims to provide a predictive model to IoHT/IoMT leaders to gauge the risk of cyber-attacks on perspective organizations.

1.3 Problem Statement

"In 2019, 41.2 million healthcare records were exposed, stolen, or illegally disclosed in 505 healthcare data breaches; the average data breach cost is 4.35M" (Seh et al., 2020b).

Allied Market Research predicted that the IoMT market would increase to 136.8 billion dollars; this prediction was based on factors such as increasing demand for remote patient monitoring, growth in the adoption of wearable devices in the healthcare industry, and technological advancements (Jones et al., 2022). The growth potential of the IoMT market emphasizes the importance of vulnerability assessments and the associated concern with the lack of security. The widespread use of IoMT devices has created numerous attack vectors, making it easier for attackers to compromise the security of these devices and gain access to confidential patient data. This highlights the need for robust security measures to be in place to protect against cyber-attacks in the healthcare industry and secure the sensitive data being handled by IoMT devices (Seh et al., 2020a).

1.4 Thesis Statement

A predictive model for cyberattacks is required to estimate the CVSS scores and financial impact for managers to assess the risk and financial exposure level in IoHT/IoMT organizations.

Security measures must be implemented to address the increasing number of healthcare cyber-attacks. It is essential to secure IoMT devices as they handle sensitive patient information and play a crucial role in providing real-time care. The

Confidentiality, Integrity, and Availability (CIA) triad is important in protecting patients and their associated data. The lack of security concerns when implementing IoMT devices increases the risk of breaches and violations of the CIA triad. Healthcare organizations must take the necessary steps to secure their IoMT devices and infrastructure to protect patient data and ensure the delivery of quality care.

The use of predictive models in combination with existing security measures is essential in protecting patients' sensitive information and ensuring the availability of medical devices. By predicting cyber-attacks in advance, proactive measures can be taken to secure IoHT devices and prevent unauthorized access to sensitive patient information. Predictive models can also identify vulnerable areas in the medical industry's network infrastructure and prioritize security measures based on the severity of potential attacks.

As society continues to embrace the Internet of Things (IoT) and relies heavily on critical infrastructure systems, the level of sophistication and frequency of cyber-attacks has increased, exposing us to substantial financial risks. Ponemon Institute estimated an average cost of $380 for every health record that was lost and/or stolen that contained sensitive and confidential information (Ponemon Institute, 2017). Providing a formula that allows companies to assess financial impact provides decision-makers with a means to assess a range of possibilities and the probability of specific events occurring.

This research aims to contribute to the advancement of cybersecurity in the medical industry and provide healthcare professionals with the tools they need to provide safe and efficient care to their patients.

1.5 Research Objectives

The research objectives are as follows:

- Develop a predictive model using NLP to predict the fundamental values necessary to compute CVSS scores.
- Develop a script that automatically computes the CVSS score based on crucial category values predicted by the NLP model.
- Provide a formula that calculates financial impact.

1.6 Research Questions and Hypotheses

The research questions and hypotheses statements presented below will guide this study in achieving its objectives.

Table 1-1. Summary of RQs and Hypotheses

Research Questions	Hypotheses
RQ1: Can regression analysis identify the most critical CVSS metrics for forecasting CVSS scores?	**H1:** Regression analysis can be used to determine which metric (Attack vectors, attack complexity, privileges required, user interaction, and CIA impact) are most important in deciding CVSS scores for IoHT/IoMT cyberattacks.
RQ2: Can regression analysis be used to identify the correlation between CVSS metrics (Attack vectors, attack complexity, privileges required, user interaction, and CIA impact)?	**H2:** Regression analysis can be used to determine the correlation between CVSS metrics.

RQ3: Can a predictive model be developed for IoMT / IoHT to predict CVSS metrics using the CVE test descriptions?	**H3:** A predictive model can be created to predict CVSS metrics with 90% accuracy.
RQ4: Can the accuracy of a predictive model be increased by combining several models into a single model?	**H4:** Several predictive models can be combined into a singular model to increase accuracy and reduce runtime.
RQ5: Can a model be developed for IoMT/IoHT to assess the financial impact?	**H5:** A model can assess the level of financial exposure of IoHT/IoMT cyberattacks.

1.7 Scope of Research

Multiple versions of the CVSS system exist. CVSS v4 is currently in peer review and will be released later this year. The other two variants currently being used are CVSS v2 and v3. This research is focused on CVSS v3.

1.8 Research Limitations

The study uses machine-learning algorithms to analyze the CVSS v3 data available before December 31, 2020. This research is limited to analyzing CVEs and corresponding CVSS metrics data published in the NVD from 2010-2020 and does not include vulnerabilities identified and published in other databases. Only specific metrics data with non-null numerical or categorical values are used, and invalid and rejected CVE entries are excluded.

CVSS and financial impact are two very different concepts that measure different things for organizations that employ them; the two concepts can, however, be related

when assessing the overall risk posed by a security vulnerability. The CVSS score is not directly meant for assessing financial impact. The CVSS score focuses on aspects such as how easy a vulnerability is to exploit, determining the impact on the CIA triad, whether the affected system is still available after a breach and other technical factors. Furthermore, every data breach is not assigned a CVSS score.

Companies usually refrain from publicly disclosing the precise costs incurred due to a data breach. Instead, entities like NetDiligence and Ponemon Institute conduct extensive research and produce reports that offer estimations and insights into potential data breach costs, drawing from their studies and surveys. The actual financial impact of a data breach can significantly differ based on the scenario and type of organization/industry that is involved. Verifying the equations relied heavily on consolidated reports from those entities.

1.9 Organization of Praxis

This praxis is organized into five chapters. Chapter 1 will provide an overview of the research, its objectives, the problem it addresses, background information, research motivation, problem statement, thesis statement, and detailed research limitations. This chapter will structure and set up the rest of the study. Chapter 2 will provide a review of previously conducted studies as they pertain to this research. The chapter will highlight any notable contributions by other authors. Chapter 3 is broken into two sections, one for CVSS methodology and a second for financial impact methodology. Each section describes how the praxis research was conducted. Chapter 4 will be an analysis of the

data collected. Chapter 5 will highlight contributions to the body of knowledge and any other conclusions reached.

Chapter 2—Literature Review

2.1 Introduction

The literature review examines research from academic journals, conference papers, and online databases. It is divided into four sections: vulnerability management, natural language processing, statistical analysis, and methods for assessing the financial impact of cyber breaches. The first section discusses the reporting and analysis of CVEs, including the history of organizations involved, CVSS calculations, and known issues with the current CVSS system. The second section details NLP and regression techniques used to determine the correlation between CVSS metrics and identify the most impactful metrics. The third section reviews methods used to assess financial impact. The final section explores cyber security trends in IoT/IoMT/IoHT, focusing on current research.

2.2 Overview of Vulnerability Management

Vulnerability management refers to identifying, evaluating, prioritizing, and mitigating vulnerabilities in software and systems (Scarfone et al., 2008). The CVE system is a widely used approach to identify and track vulnerabilities (Cyber Defense Magazine, 2017). CVE assigns a unique identifier to each exposure, making tracking and analyzing vulnerabilities across different systems and vendors easier. The NVD, maintained by the NIST, is a vulnerability information repository supporting vulnerability management efforts. The database contains information about vulnerabilities, including their impact, severity, and remediation information (National Institute of Standards and Technology, 2021a).

The CVSS is a framework for the assessment of vulnerabilities that may impact organizations. The CVSS framework provides a scoring system with Base, Temporal, and Environmental metric groups. The Base Metrics group includes three subgroups: Exploitability, Impact, and Scope. Despite the usefulness of the CVSS framework, there are known issues with the system, including a need for granularity and subjectivity in specific metrics (Rasmy et al., 2021). Researchers have explored ways to improve the CVSS system; some of those methods have included incorporating machine learning algorithms to assist with the prediction of the severity of vulnerabilities, which is covered in 2.3. The next sections cover vulnerability management in detail.

2.2.1 Common Vulnerability Analysis and Planning

Common Vulnerability Analysis and Planning (CVAP) is a vulnerability management technique that involves a continuous cycle of activities aimed at providing structure (Joshi & Singh, 2017). CVAP is critical in the IoHT environment. In the IoHT, vulnerabilities can exist in various components, including medical devices, sensors, gateways, and cloud services. The CVE system, maintained by the NVD, is an essential source of vulnerability information, which can be used to identify potential weaknesses in the IoHT environment. The CVSS system calculates vulnerability scores, which can help organizations prioritize vulnerabilities and take appropriate action (Joshi & Singh, 2017).

2.2.2 The Common Vulnerability Scoring System

The Forum of Incident Response and Security Teams (FIRST) is a global organization that provides organizations an opportunity to share information and best

practices. FIRST is responsible for first introducing the CVSS in 2005. In the context of IoHT, CVEs are created for medical devices, methods, and networks at the description of the organizations; this allows CVSS scores to be assigned for IoHT vulnerabilities. There are multiple versions of the tool exist, but this study focuses on v3.1.

2.2.2.1 Common Vulnerability Scoring System Metric Groups

CVSS consists of three metric groups: Base, Temporal, and Environmental. The Base metrics are the bare minimum needed to compute a CVSS score; these vulnerabilities will not change as time goes on or based on the environment where the vulnerability is exploited. The base metric group is further broken into two metric types: exploitability and impact.

- Exploitability metrics measure how simple it is for a hacker to use a vulnerability. The sub-metrics are Attack Vector (AV), Attack Complexity (AC), Privileges Required (PR), User Interaction (UI), and Scope (S).
- Impact Metrics measure the impact of an exploited vulnerability on the target system. The sub-metrics include the CIA triad, which represents Confidentiality (C), Integrity (I), and Availability (A).

Each sub-metric is defined in the chart below:

Table 2-1. Metrics Definition

Metric	Definition	Potential Values
AV	Describes how an attacker gains access to a system or network	Network (N) Adjacent (A) Local (L) Physical (P)
AC	Describes how difficult it is for an attacker to launch an exploit against a target system	Low (L) High (H)
PR	Describes the amount of access a hacker needs to exploit a vulnerability successfully	None (N) Low (L) High (H)
UI	Describes whether a user is required to perform some action to enable the attack	None (N) Required (R)
S	Describes how much impact a vulnerability has on the system it is attempting to target	Changed (C) Unchanged (U)
C	Describes the extent to which unauthorized access or disclosure of information is prevented	None (N) Low (L) High (H)
I	Describes the impact on reliability and accuracy, or the integrity, of information	None (N) Low (L) High (H)
A	Describes the impact of a system going down or being lost; those services include websites, servers, databases, email, etc.)	None (N) Low (L) High (H)

Temporal metrics are characteristics that change as time goes on but are not based on the environment. Environmental metrics are system or environment attributes that are specific to it. These metrics adjust the Base Score further and reflect the risk posed by the vulnerability in a specific environment (FIRST, 2023e).

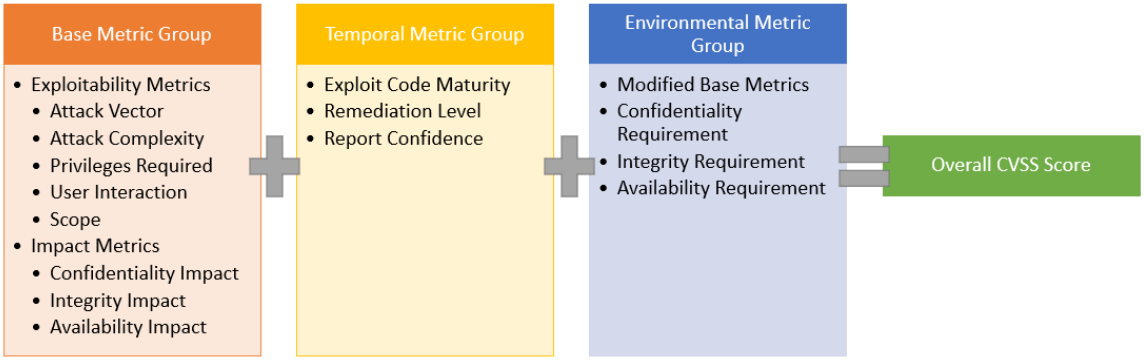

Figure 2-1. CVSS Metric Groups.

While temporal and environmental metrics allow organizations to tailor a score to their environment, these topics are not covered in detail as a part of this research. CVSS scores can be calculated using the base metric group alone without loss of generality.

CVSS assigns a numeric score to represent the severity of a vulnerability using the following steps:

1. Define the attack vector.
2. Determine the attack complexity.
3. Define the privileges required to launch the attack.
4. Determine the user interaction required.
5. Assess the scope of the attack.

6. Assign base scores: Once the attack vector, attack complexity, privileges required, user interaction, scope, and CIA have been determined, the base score for the vulnerability can be calculated using the CVSS formula. (NOTE: For organizations using temporal and environmental metric groups, those scores would be assigned after assessing the temporal and environmental metric groups.)
7. Determine the severity level: Finally, the severity level can be determined based on the overall CVSS score. Severity levels range from low to critical, with higher scores indicating more severe vulnerabilities.

The score is calculated using a mathematical formula founded on the base, temporal, and environmental metrics. This research focuses on version 3.1 using the following formula in the base metric group (the temporal and environmental formulas were only displayed for completeness):

Table 2-2. Metrics Group Formula – v3.1 CVSS Scores

Metric Group	**Formula**	**Score Range**
Base	"If Scope is Unchanged, Roundup (Minimum [(Impact + Exploitability), 10]) If Scope is Changed, Roundup (Minimum [1.08 × (Impact + Exploitability), 10]) Where: If Scope is Unchanged, Impact = 6.42 * ISS If Scope is Changed, Impact = 7.52 * (ISS - 0.029) - 3.25 * (ISS - 0.02)[15] ISS = 1 - [(1 - Confidentiality) × (1 - Integrity) * (1 - Availability)] Exploitability = 8.22 * AttackVector * AttackComplexity * PrivilegesRequired * UserInteraction" (FIRST, 2023d)	0 to 10, with ten being the most severe
Temporal	Base Score * Exploit Code Maturity * Remediation Level * Report Confidence	0 to 10, with 10 being the most severe

Environmental	Round_Up (Adjusted Temporal + Adjusted Impact * Modified Scope) Where: Adjusted Temporal = Temporal Score * Collateral Damage Potential * Target Distribution Adjusted Impact = Impact * Collateral Damage Potential * Target Distribution	0 to 10, with 10 being the most severe

The resulting scores are then mapped to one of the severity categories used to aid organizations in understanding the vulnerability and its associated severity based on its CVSS score.

For CVSS v3, the severity categories are as follows:

Table 2-3. Severity Category

Severity Category	**Score Range**
None	0.0
Low	0.1 - 3.9
Medium	4.0 - 6.9
High	7.0 - 8.9
Critical	9.0 - 10.0

The severity category can help to understand the risk associated with a vulnerability and prioritize remediation efforts in accordance with that risk level. Exposure in the "critical"

variety is the most severe and requires immediate attention. In contrast, a vulnerability in the "low" category may not require close attention but should still be addressed promptly.

2.2.2.2 CVSS Related Research

According to Chase and Coley (2020), the “Rubric for Applying CVSS to Medical Devices” provides a mechanism for conducting the assessment of security vulnerabilities in medical devices. The rubric for assessing the CVSS is created to steer specialists through a series of questions as a flow chart with various decision points to evaluate. Each component of the CVSS vector (AC, AV, S, C, etc.) has its rubric with decision points. It is expected that specialists record their answers and follow the suggested vector options that contribute most significantly to the CVSS score. Some questions may require additional information for risk analysis purposes. Apart of the scoring process should be reaching out to individuals that have been deemed subject matter experts (SMEs) rather than relying solely on a single analyst to achieve better results. The group of SMEs should possess cross-domain knowledge in cybersecurity, patient health and privacy, and information technology, to name a few; if this assessment is for a device or tool, the SME list should include device engineering, tool support as applicable and any integration support (Chase & Coley, 2020). Sharing expertise across these areas ensures a comprehensive evaluation of the CVSS and its implications for patient safety (Chase & Coley, 2020). While this rubric provides clarity, it perpetuates the issue that most SMEs have with CVSS scores; the decision points are subjective and left up to the assessor’s interpretation. Evaluations of the answers will vary from assessor to assessor.

C. Elbaz et al. proposed a method for predicting the CVSS score using base metrics using the vulnerability description. The first step in their approach is to transform the vulnerability description into word representations. This information is presented as a vector with dimensions corresponding to the number of times an individual word appears in the description. 0 was used to indicate the absence of that word. Words deemed unnecessary are removed to increase accuracy and reduce noise in the data. The next step is to train linear regression models to predict a score for each of the sub-metrics of that particular CVSS vector. Explainability, or ease of explanation, is an advantage of simple linear regression models. However, this method has limitations. It does not take into consideration the complexity of the human language, and it assumes a linear relationship between the things input and the respective output. This limitation can limit the performance of the model. Additionally, using the bag of words approach disregards the need to pay attention to context. It discards word ordering; this results in the text data being inadequately represented, again not capturing the nuances of natural language, which sometimes affects the accuracy of the model's predictions. Overall, while the method of C. Elbaz, et al. provides a simple and explainable approach for predicting CVSS metrics; it may have limitations in accurately modeling the complexity of natural language and may not capture the full context of a vulnerability description.

A. Khazaei et al. proposed a method for predicting the CVSS severity scores from the descriptions of vulnerabilities submitted as CVEs. The input data is processed; some of the steps in that process are to remove any stop words from the reports, the words that are left undergo stemming, and finally, a calculation is performed that represents each

word's “Term Frequency–Inverse Document Frequency value” (A. Khazaei et al.). There were three different models trained and studied; those models include SVM, random forest with dimensionality reduction steps, and a fuzzy system. The main issue is that the score that is generated as a result of this model is an approximate value. While the proposed method (A. Khazaei et al.) did not have a mechanism for understanding the decision-making process, one advantage of this approach is that it provides a more fine-grained prediction of the severity score by dividing the continuous interval into ten sub-intervals. However, it is important to note that using an approximate severity score may not always accurately represent the vulnerability's actual severity. The lack of explanation for the model's predictions may also limit its interpretability and trustworthiness.

Under the FDA contract, Chase and Coley (2020) developed a rubric that provides additional considerations of the CVSS to provide a risk assessment that was specific to IoMT and medical devices. In order to produce a usable rubric, several teams were involved in a collaboration effort; that team consisted of security, FDA, medical device manufacturers, healthcare organizations, cyber, and others that would provide valuable insight. The goal of the rubric is to assist users in consistently assessing medical device cybersecurity vulnerabilities using the same criteria because various calculators have been developed to assess CVSS scores, but no unique considerations are given for medical devices. In October 2020, the FDA announced the qualification of the rubric as a Medical Device Development Tool (MDDT). While a specialized tool is helpful in tailoring the CVSS scores, the same issue of subjectivity remains; the inputs of the Excel file are open to interpretation.

2.3 Overview of Natural Language Processing

NLP develops algorithms to understand, analyze, and generate human language (Rocca et al., 2023). NLP integrates computer science, linguistics, and artificial intelligence. NLP techniques and applications include tasks such as:

- Sentiment analysis: determining a given text's sentiment (positive, negative, or neutral).
- Machine translation of the text
- Entity (people, organizations, locations, etc.) recognition.
- Text summarization: generating a summary of a given text.
- Question answering: answering natural language questions based on a text or database.
- Chatbots: creating conversational agents that understand and respond to natural language inputs.

NLP execution is partitioned into four significant steps: lexical analysis, syntactic analysis, semantic analysis, and output transformation. The figure below explains each step.

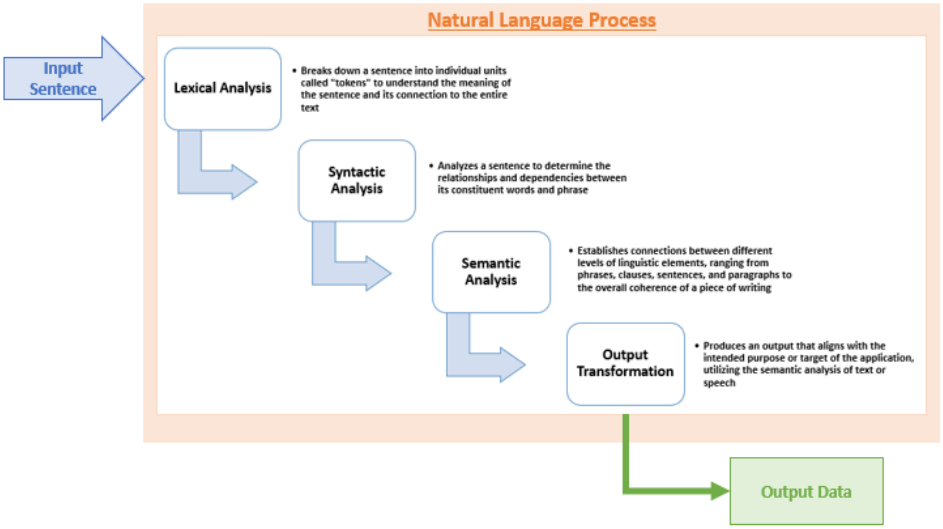

Figure 2-2. NLP Process.

NLP models and algorithms use machine learning, deep learning, and rule-based approaches to complete the steps outlined above.

One continued challenge for NLP languages is the training data needed to yield a learning-based model that performs well. To overcome this, researchers have developed models pre-trained on the human language using unannotated text on the web. BERT is a popular pre-trained language representation model that has changed the field of NLP.

2.3.1 BERT & BERT Derivative Models

BERT is a bidirectionally trained language model based on the Transformer architecture. A Transformer is a neural network that operates by performing a small, fixed number of steps. In each stage, it applies a process that helps it understand the connections and relationships between the words, or tokens, in a sentence, regardless of

their position. This means that the Transformer can learn to recognize context-dependent meanings of words (Ameri et al., 2021). As mentioned before, BERT is pre-trained on a large quantity of unannotated text words; this aids the transformers in handling natural language processing tasks identified earlier. Specifically, language translation, question answering, sentiment analysis, etc., makes handling natural language processing tasks where the context is pivotal successful.

Two tasks are used to train BERT models: Masked Language Modeling (MLM) and Next Sentence Prediction (NSP). In MLM, BERT trains the data by randomly masking various words in the sentence input into the model. It then attempts to predict the missing or previously masked words based on the remaining words left in the sentence, thereby allowing the model to learn about the relationship between the words in the sentence (Koroteev, 2021). In NSP, BERT predicts whether two input sentences are consecutive in the original text (Ameri et al., 2021). After pre-training, BERT can be fine-tuned on various NLP tasks, including the areas identified in Figure 2.2. Fine-tuning is training the BERT parameters on a smaller annotated dataset (Rasmy et al., 2021). The advantage of using BERT is that it can learn contextual relationships between words and capture the meaning of a sentence more complexly than traditional NLP models. BERT has achieved recognition as being effective and producing superior results on many NLP tasks and is often chosen by practitioners for many NLP applications (Koroteev, 2021).

RoBERTa, DistilBERT, ALBERT, and DeBERTa are variations of the original BERT model that have been modified to improve pre-training and fine-tuning processes.

- Robustly Optimized BERT Pre-training Approach (RoBERTa) changes the pre-training process. These changes include removing the NSP task and training on

more data for extended periods. This could lead to performance improvements in natural language inference as well as other various NLP tasks (Singla, 2020).

- Distilled BERT (DistilBERT) is also known as a smaller, faster version of BERT that uses fewer parameters and a simpler architecture. It uses MLM tasks to reduce the number of parameters to allow for speedier training without compromising the accuracy of BERT (Furnell et al., 2020).
- A Lite BERT (ALBERT) is a variation of BERT that introduces a new training strategy called Factorized Embedding Parameterization. This strategy reduces the number of parameters in the model while improving performance. ALBERT also introduced two new concepts, cross-layer parameter sharing, and inter-sentence coherence loss, to enhance training efficiency and generalization (Singla, 2020).
- Decoding-enhanced BERT with Disentangled Attention (DeBERTa) uses disentangled attention to allocate different model parts to work on various aspects of the input sequence. It also uses a new training objective called Masked Entity Prediction; this training helps the model learn to predict named entities even when masked in the input sequence.
- BART (Bidirectional and Auto-Regressive Transformer) removes noise, analyzes it, and then attempts to put the data back together based on the information provided by the nose. It involves two steps: (1) altering the input text using a noise function and (2) training a model to reconstruct the original text. Users suggest BART because it combines the bidirectional encoder from BERT and the left-to-right decoder from GPT, and elements from other pretraining approaches (Lewis et al., 2020).

2.3.2 NLP-Related Research

In a study entitled "XLNet-Based Prediction Model for CVSS Metric Values" (Shi et al., 2022), the authors propose a method for predicting CVSS metric values using the XLNet model. The XLNet model is trained and calibrated using a custom data warehouse developed by the researchers. The research proposed the design of a machine learning model that uses text data to predict more than one vulnerability feature. The study by Shi et al. (2022) validates that the fine-tuned XLNet model improves prediction accuracy and enhances tasks that are performed later in the vulnerability assessment. The findings of this research are that the XLNet model does not produce results that are far superior to logistic regression. The authors acknowledge this and suggest future research directions for combining pre-trained and traditional models to achieve better outcomes (Shi et al., 2022).

According to the study conducted by McGlade and Scott-Hayward (2019) titled "ML-based cyber incident detection for Electronic Medical Record (EMR) systems," the researchers propose using a combined approach of Machine Learning (ML) and Time Series (TS) anomaly detection techniques to address C and A cyber-attacks on EMR systems. Their results indicate that ML can successfully detect C incidents, with SVMs achieving the best metric values. This research performed an analysis of the following metric values: highest accuracy, precision, and recall. Furthermore, their research presents a prototype that uses an Exponential Moving Average to detect surges within 10 seconds for availability incident detection (McGlade & Scott-Hayward, 2019). This finding presents a significant advancement toward developing an automated surge

defense system, surpassing the manual methods currently employed. The study also researches the option to deploy these detection settings in a clinical setting (McGlade & Scott-Hayward, 2019).

Han et al. proposed a method for predicting qualitative CVSS severity ratings from vulnerability descriptions as opposed to the traditional quantitative method of predicting CVSS scores. Word embeddings are trained using a method for learning unstructured text data, also known as skip-gram modeling, to represent words in a vector space. CNN models are fed vector sets that are created by attempts at word embeddings to encode word meanings. Those encodings are placed in a vector space. The vulnerability description is then transformed into a set of vectors consisting of the concatenation of word embeddings of words present in the description. Because this method only predicts categories and does not produce a precise numerical severity score. One advantage of this approach is that it uses word embeddings to capture the meaning of words, which can better represent the text data than other methods using a bag-of-words representation. There is no way to verify the accuracy of the model because the predictions were not explained. This can limit the interpretability and trustworthiness of the model. In summary, using CNN and word embeddings, Z. Han et al.'s approach provides a way to predict qualitative CVSS severity ratings from vulnerability descriptions. While this approach has some advantages, such as word embeddings, it also has limitations, such as the inability to explain the model's predictions.

2.4 Assessing Financial Impact

Assessing the financial impact of cyber breaches is a complex task that involves both costs that are direct and indirect. Organizations benefit from understanding both types of costs when developing and optimizing data breach response plans. A study conducted by Symantec and the Ponemon Institute suggests that such understanding can potentially result in a significant reduction of as much as 47% with regard to the costs associated with data breaches (Ponemon Institute, 2017). By analyzing and comprehending the various factors contributing to the costs, organizations can make informed decisions and implement effective strategies to mitigate and manage the impact of data breaches more efficiently. Direct costs are split into three categories: incident response, legal and regulatory expenses, credit monitoring, and costs associated with recovery and remediation. While indirect costs are divided into downtime, staffing, training, legal fees, damage to reputation, loss of customer/ customer confidence, etc.

Furnell et al. (2020) researched the need for organizations to understand the costs associated with cybersecurity breaches comprehensively. To address the gaps in the availability of information on comprehensive approaches, a study commissioned by the Department for Digital, Culture, Media & Sport in the UK aimed to develop a more holistic, robust approach to measuring the costs of cybersecurity breaches. The study identified a range of specific costs that organizations should consider, including changes in cybersecurity practices, compensation/discounts to affected customers, consultant fees, credit rating/insurance premiums, data and software loss, financial theft, fines, and many others. A questionnaire-based data collection tool was developed to support organizations

in capturing and measuring these costs in a consistent and structured manner. The tool was divided into five thematic sections: the nature of the breach, immediate losses or damage, legal or regulatory costs, costs related to technical cybersecurity teams or practices, and other costs to the whole organization. Each section contained specific questions to gather relevant information. The evaluation of the tool involved 15 UK businesses and charities that had previously experienced breaches with significant costs. While the tool provided meaningful results for several organizations, there were challenges in estimating and monetizing time-related costs, measuring longer-term costs, and attributing spending to a specific breach. The comprehensive nature of the tool also required a substantial time commitment from respondents (Furnell et al., 2020).

The RAND report adopts a modeling approach to estimate both the present and the future global costs of cybersecurity risks associated with cyber-attacks and data breaches; this approach considers the uncertainty surrounding how often attacks occur and the specific costs associated with these incidents (Dreyer et al., 2018). The report provides an Excel tool that is used as a modeling and simulation platform that allows users to modify assumptions and explore various research questions. The model used in the report calculates direct costs; these costs are determined by considering the share of value-added related to a specific sector in a particular country and the output value of that sector within that country. The model incorporates two unitless values: one represents a particular industry at risk from each type of exposure, and another represents the fraction of exposure at risk for a given country and sector. By multiplying these unitless values by the output of a specific sector in a particular country, the direct cost to sector output is

determined. It is important to note that the model does not rely on specific examples of cybersecurity breaches. Instead, it relies on the model's underlying assumptions to generate reliable cost estimates. However, the report acknowledges that many of these assumptions are not thoroughly explained, making evaluating their estimation errors challenging.

Algarni and Malaiya (2016) proposed an approach to consolidate the factors influencing the cost of security breaches. The authors reviewed existing data breach calculators and eliminated redundant or insignificant factors under the thought process that current models do not accurately calculate the cost of security breaches because they fail to consider important aspects of data breach risk estimation. Power regression models were used to improve cost calculations associated with data breaches, aiming for a better fit to the data with maintain the ability to include characteristics that affect the direct and indirect costs associated with a data breach. The authors criticize the simplicity of the model used by the Ponemon Institute, stating that it is overly simple and potentially misleading for organizations relying on it for risk analyses. While appreciating the simplicity of the Ponemon model, the argument is that using only the number of records lost in a breach is not inclusive of all of the factors that need to be assessed to understand a breach's financial impact.

In 2021, Algarni and Malaiya furthered this research by developing a model that accurately estimates two crucial risks associated with cyber security: The first risk is the cost of a data breach, and the second is the probability of experiencing a data breach within 12 months. This model takes into consideration the concept of economy of scale,

as researchers found that solely relying on the metric of data breach cost per record can be misleading. Incorporating a more comprehensive approach accounts for the varying factors and complexities that impact the overall cost of a data breach, resulting in a more accurate estimation. This research is covered more in Chapter 3 as it is the baseline research for assessing the financial impact.

2.5 Summary and Conclusions

The literature review discusses cyber security trends in the medical field; emphasis was placed on vulnerability management, natural language processing, statistical analysis, and methods for assessing the financial impact of cyber breaches. The review also provides an overview of vulnerability management concepts; those include the CVSS and the CVAP framework. It explains how CVSS is used to assess the severity of vulnerabilities and how the CVAP framework helps organizations identify and remediate vulnerabilities in their systems. The review also touches upon research for improving the CVSS system using machine learning algorithms. Additionally, it highlights other research studies that propose methods for predicting CVSS metrics and severity scores using vulnerability descriptions. The review concludes by presenting a rubric for assessing medical device cybersecurity vulnerabilities and its qualification as a MDDT by the FDA. The research conducted in this paper will expand upon the concepts established by earlier work.

Chapter 3—Methodology

3.1 Introduction

This chapter presents the methodology used to achieve the objectives of this research. Those objectives are to:

- Predict the values of the eight labels required to compute the base CVSS score based on the text descriptions provided.
- Calculate the CVSS score based on predicted label values.
- Calculate the cost of a data breach.

Below is a flowchart that outlines the major steps in this praxis:

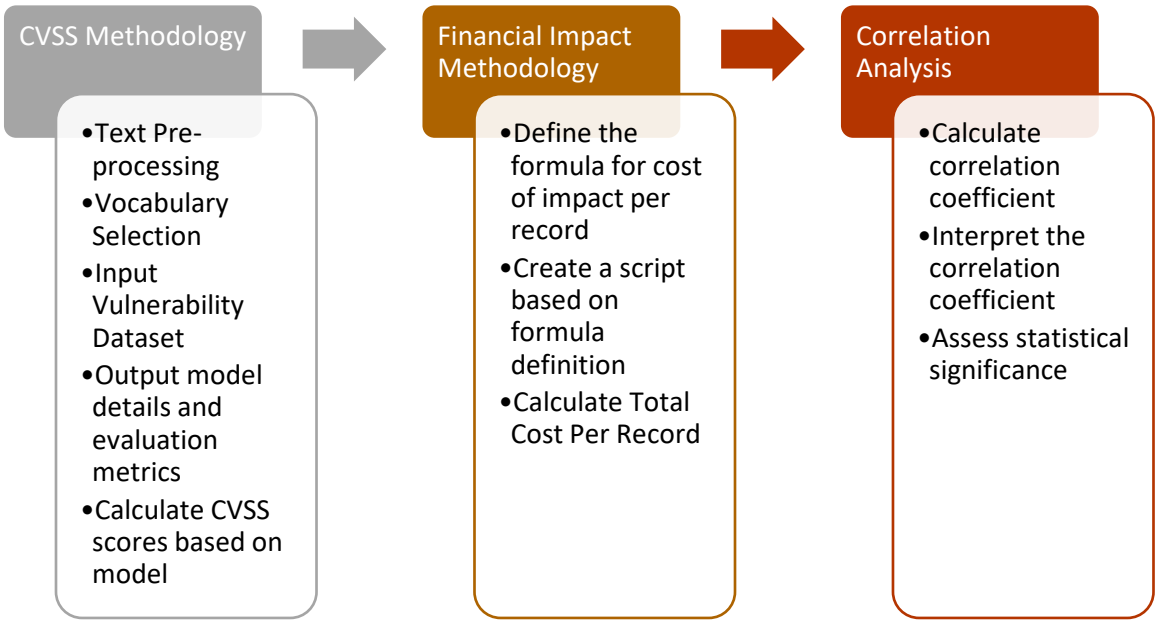

Figure 3-1. Praxis Methodology.

The present research draws upon the work (Costa et al., 2022) in "Predicting CVSS Metric via Description Interpretation." The article will be referred to as "CVSS-baseline research" in this Praxis. The models were made available in the researcher's GitHub repository. The first step was to recreate the results of CVSS-baseline research and ensure that the scripts were functional. After replication of the results obtained from the CVSS-baseline research, the next objective focused on identifying strategies for enhancing the performance of the models. Once the performance was deemed satisfactory, the subsequent step was to leverage the calculation prescribed by FIRST to calculate corresponding CVSS scores. The final objective was to assess the financial impact associated with data breaches. This portion of the research draws upon the work of Algarni, Thayananthan, and Malaiya, as documented in their publication titled "Quantitative Assessment of Cybersecurity Risks for Mitigating Data Breaches in Business Systems." Within the context of this dissertation, the article will be referred to as "financial impact-baseline research." After replication of the results obtained from the financial impact-baseline research, the next objective focused on determining any necessary modifications for the equation.

3.2 CVSS Methodology

3.2.1 Text Pre-Processing and Vocabulary Selection

In NLP, one of the first tasks is to understand and assess it is important to assess the impact or contribution of each word to the classification of specific categories. The Natural Language Toolkit (NLTK) methods and the Transformers library from Hugging

Face are used to tokenize the text. Lemmatization and Stemming are commonly used text pre-processing approaches in NLP. These methods maintain relevant data while ignoring noisy data. In the case of Stemming, only the base word is used by removing the last few characters of a word. Lemmatization considers the context of the word before converting the word into a meaningful base form. This pre-processing was completed first.

The subsequent step involved expanding the model's vocabulary while combining it with the most effective text pre-processing approach. The accuracy was assessed by gradually incorporating 5,000, 10,000, and 25,000 words into the tokenizer's default lexicon. The selection of additional words was based on their frequency in the descriptions, with priority given to the most frequently appearing terms. Only words exclusive to the descriptions and not already present in the default vocabulary were considered to ensure uniqueness, thereby avoiding duplication. This approach optimized the model's performance by refining its vocabulary with relevant terms. Software versions, special characters, extra spaces, and code snippets often appear in some of the data descriptions, so the decision was to apply regular expression filters to digits and special characters to reduce noise in vocabulary addition. This approach further reduces the noise in the data. This methodology was used as is from the baseline study.

3.2.2 Vulnerability Dataset and Model Details

NVD is a repository of vulnerability management data maintained by the United States government; the dataset used in this study is sourced from that database. The information was compiled through the NVD API and used data up to April 2021. This is

the same dataset used in the CVSS-baseline-research. After collecting the data, it was processed to extract vulnerability descriptions and classify them into eight categories: AV, AC, PR, UI, S, C, I, and A. Based on the CVSS documentation, the eight categories were grouped into three metrics: Exploitability metrics (AV, AC, PR, UI), Scope, and Impact metrics (C, I, and A).

In the CVSS-baseline-research, the team focused on CVEs related to version 3 of CVSS; this decreased the dataset to 79,810 instances in total. The dataset was split into train and test sets. The train set contains 63,848 instances, while the test set contains 15,962 instances. This split follows a test ratio 0.2, meaning approximately 20% of the data is allocated to the test set. On average, the description length in the train set is 43.85 words, and in the test set, it is slightly higher at 44.55 words. Table 3-1 details the analytical values displayed in the figures above.

Table 3-1. Class Proportions[1]

Category	Class Prior (%)		
	Train	**Test**	**Dataset**
Attack Vector			
Network	73.04	73.57	73.14
Adjacent	2.21	2.14	2.20
Local	23.62	23.18	23.53
Physical	1.13	1.11	1.13
Attack Complexity			
High	8.03	7.74	7.97
Low	91.97	92.26	92.03
Privileges Required			
High	6.70	6.32	6.62
Low	26.79	26.90	26.81
None	66.51	66.78	66.57
User Interaction			
Required	35.80	35.37	35.71
None	64.20	64.63	64.29
Scope			
Changed	16.72	16.36	16.65
Unchanged	83.28	83.64	83.35
Confidentiality			
High	58.94	59.15	58.98

[1] This table is a part of baseline research, but the values are specific to this research. (Costa et al., 2022)

Category	Class Prior (%)		
	Train	Test	Dataset
Low	19.37	18.65	19.23
None	21.69	22.20	21.79
Integrity			
High	51.27	51.36	51.29
Low	17.52	17.32	17.48
None	31.21	31.32	31.23
Availability			
High	58.49	58.85	58.56
Low	2.60	2.71	2.63
None	38.91	38.44	38.81

This table is useful because it clearly shows that the train and test sets maintain a similar distribution of classes. CVSS-baseline-research made the dataset available to ensure repeatability, and it provided the foundation of this research.

3.2.3 Model Details and Evaluation Metrics

The CVSS-baseline research used BERT, DistilBERT, RoBERTa, ALBERT, and DeBERTa NLPs to determine the appropriate model for CVSS metrics. BERT is known for its high-quality results and has seen wide adoption for NLP tasks. Later results in Table 4-2 will show that BERT emerged as an obvious choice for this research. Table 3-2 lists the parameters used when building the models for each NLP. An analysis of these

BERT-based models, accounting for variations in parameter sizes, was used to gauge their efficacy in predicting CVSS metrics. Below are the hyperparameters employed by CVSS-baseline research:

Table 3-2. Hyperparameters Used to Fine-tune Models

	Learning Rate	**# of Training Epochs**	**Batch Size**	**Weight Decay**
BERT	3×10^{-5}	3	4	0
DistilBERT	5×10^{-5}	3	8	0
RoBERTa	1.5×10^{-5}	2	4	0.01
ALBERT	3×10^{-5}	3	8	0
DeBERTa	3×10^{-5}	10	4	0

To extend CVSS-baseline research, a strategic determination was made to introduce BART as an additional viable option. The CVSS-baseline research developed eight distinct models based on the categories for each available NLP option (i.e., BERT produced a separate model for AC, AV, PR, UI, S, C, I, and A). Each of these eight models underwent independent training. One of the hypotheses suggested in this research is that enhanced performance can be attained through the implementation of shared learning methodologies; this leads to the consolidation of diverse models into a singular representation for each NLP alternative (i.e., BERT is comprised of a single model that encompasses all eight categories). A deliberate decision was made not to alter the hyperparameters to establish a standardized framework for comparing the factors at hand.

Keeping the hyperparameters constant ensured an unbiased comparative analysis (Costa et al., 2022).

3.2.4 Calculate the CVSS score Based on Predicted Label Values

Section 3.4 yielded prediction labels that correlate to the eight components required to compute the CVSS score. This research focuses on the base formula only. This score relies heavily on the Impact Sub-Score (ISS), Impact, and exploitability. To recap, the formula is as follows:

$$I^2SS = 1 - [(1 - Confidentiality) * (1 - Integrity) * (1 - Availability)$$

The ISS is used to calculate Impact and Exploitability. Impact is determined by the scope.

- If the scope is unchanged,

$$Impact = 6.42 * ISS$$

- If the scope is changed,

$$Impact = 7.52 * (ISS - 0.029) - 3.25 * (ISS - 0.02)^{15}$$

Exploitability is determined by Attack Complexity, Attack Vector, Privileges Required and User Interaction.

$$Exploitability = 8.22 * \text{AttackVector} * \text{AttackComplexity} *$$

$$\text{PrivilegesRequired} * \text{UserInteraction}''$$

[2] These formulas are provided by FIRST. (FIRST, 2023e)

The values are determined by the metric values provided by FIRST. The applicable metrics for the base score are:

Table 3-3. CVSS Metric Numerical Value

Metric	Metric Value	Numerical Value
Attack Vector	Network (N)	0.85
	Adjacent (A)	0.62
	Local (L)	0.55
	Physical (P)	0.2
Attack Complexity	Low (L)	0.77
	High (H)	0.44
Privileges Required	None (N)	0.85
	Low (L)	0.62 (or 0.68 if Scope is Changed)
	High (H)	0.27 (or 0.5 if Scope is Changed)
User Interaction	None (N)	0.85
	Required (R)	0.62
Confidentiality / Integrity / Availability	None (N)	0
	Low (L)	0.22
	High (H)	0.56

The CVSS-baseline research used BERT, DistilBERT, RoBERTa, ALBERT, and DeBERTa NLPs to determine the appropriate model for CVSS metrics. BERT is known for its high-quality results and has seen wide adoption for NLP tasks. Later results in Table 4-2 will show that BERT emerged as an obvious choice for this research. ALBERT

and DistilBERT require fewer parameters than BERT, while RoBERTa and DeBERTa exhibited a greater parameter count relative to BERT. An analysis of these BERT-based models, accounting for variations in parameter sizes, was used to gauge their efficacy in predicting CVSS metrics. Below are the hyperparameters employed by CVSS-baseline research:

CVSS v3.1 formula ranges from 0.0 to 10.0 with an allowed deviation of 0.5 (FIRST, 2023d). To calculate the CVSS scores, a script was generated that performed the following:

- Imported the prediction labels determined in Section 3.4
- Calculated the CVSS score based on prediction labels imported
- Performed a comparison between the CVSS scores generated based on predictions and the CVSS scores provided by FIRST
- Exported the calculated CVSS scores into an Excel file.

As a last step, multiple linear regression was performed to model the relationship between the three-impact metrics (C, I, and A) and the CVSS scores. The R-squared value was calculated; it analyzes the dependent variable and assesses the amount of variability caused by the regression model. The Mean Squared Error (MSE) is a the squared average value of the difference between the predicted values and the values that are the target. It quantifies the accuracy of the regression model's predictions. MSE was calculated. The Root Mean Squared Error (RMSE) is the square root of the MSE.

3.3 Financial Impact Methodology

3.3.1 Cost based on Number of Records Impact vs Total Cost Per Record

Calculating financial impact is nuanced because of the various methods used to estimate cost. A quantitative model provides a systematic method of evaluating the total cost of a data breach, considering factors such as the number of records and other relevant variables. This model should be built upon actual field data and realistic assumptions to ensure accuracy. As previously discussed, the model will incorporate the theory of economies of scale, as the traditional metric of data breach cost per record has been deemed deceptive in understanding the cost of a data breach (Algarni & Malaiya, 2016). In the 2016 conference paper, Algarni and Malaiya analyzed current models and determined that the cost of security breaches is not solely based on the number of records lost, stolen, or otherwise compromised; instead, multiple factors have to be considered in assessing the cost of data breaches. The researchers evaluated data breach calculators and various factors that impact data-breach costs. Their research involved a thorough examination of those factors, which included removing redundant factors where possible, and their effect on data breach risks.

Financial impact-baseline research deduced that the Ponemon Institute method is overly simplistic and misleading; the primary criticism lies in its heavy reliance on the number of impacted records as the sole determining factor. This approach fails to consider the complexities in accurately estimating the true costs associated with a data breach leading to inadequate preparations, misallocation of resources, and an overall

underestimation of the potential impact on the organization. The model proposed by Algarni, Thayananthan, and Malaiya offers a comprehensive approach to determining the cost per record by considering all relevant factors. Furthermore, their model also addresses the likelihood of security breaches, which is crucial for assessing data breach risks. In this praxis, the focus will be on determining the cost per record.

3.3.2 Calculate Total Cost Per Record

Financial impact-baseline research was able to establish a more consistent and complete approach that did not rely on the number of records lost, stolen, or otherwise compromised. This was the result of the application of economies of. By conducting regression analysis on three datasets obtained from Ponemon in 2013, Ponemon in 2014, and NetDiligence, along with using existing data breach cost calculators, the following formula was derived to estimate the total security costs: EB+RB=ESC, where EB is the expected costs due to breaches, RB is the cost of any breaches and ESC is the expected annual security cost. This praxis focuses on EB.

To calculate the cost of a data breach, EB is derived into the following formula:

$$CPR_1 = F_c \times F_I \times F_D \times \sum_{i=1}^{4} c$$

where F_C is the country, F_I is the industry, and F_D is the amount of time a business maintains sensitive information.

The cost factor based on F_C is determined by the country of the organization. The values used in this research were a by-product of research by the Ponemon Institute. By

default, the USA is selected as the baseline country. The cost in other countries may be lower, and this adjustment can be made by using the values in a similar report to Ponemon. The cost factor based on F_I represents the industry where the cyber-attack took place. The values are displayed in the table below:

Table 3-4. Industry [3]Cost Factor (F_I)Metrics Definition

Industry Classification	**Multiplier**
Communications	1.01
Consumer Products	0.88
Education	0.85
Financial Services	1.26
Government Services	0.78
Healthcare and Pharmaceuticals	1.33
Industrial	0.80
Retail	0.84
Services: professional and general services	1.12
Technology and software	1.23
Transportation	0.90
All Others	1 (default)

The cost factor based on F_D represents the amount of time that sensitive information is retained by an industry.

[3] Table 3-4 was used in the financial impact baseline research (Algarni & Malaiya, 2016).

Table[4] 3-5. Sensitive Data Retained Cost Factor (F_D)

Duration	Multiplier
3 months	0.125
1 year	0.5
4 years	1 (default)
6 years	1.5

The factors in the summation formula will be multiplied by 4 partial costs, which are:

- Incident Investigation Cost Per Record:

$$ICPR = [ax^{b-1}] \times F_{BC} \times F_E \times F_P$$

where x is the number of impacted records, a and b are cost parameters defined by in the financial impact-baseline research study; the a and b parameter decision points were broken into two separate tables for ease of viewing:

[4] Table 3-5 was used in the financial impact baseline research (Algarni & Malaiya, 2016).

Table 3-6. $[ax^{b-1}]$ Parameters – Data in Centralized Location [5]

Decision Points:		If the data is in a centralized location, fraud is expected, and federal lawsuit filed:	If the data is in a centralized location, fraud is expected, but no federal lawsuit filed:	If the data is in a centralized location, no fraud is expected, but federal lawsuit filed:	If the data is in a centralized location, no fraud is expected, and no federal lawsuit filed:
PHI and SSN	a	1532.8	1473.2	1452.4	1473.2
	b	0.57	0.50	0.56	0.50
CC	a	1108.1	1093.5	1054.2	1093.5
	b	0.57	0.49	0.56	0.49

Table 3-7. TABLE: $[ax^{b-1}]$ Parameters – Data in Centralized Location

Decision Points:		If the data is not in a centralized location, fraud is expected, and federal lawsuit filed:	If the data is not in a centralized location, fraud is expected, but no federal lawsuit filed:	If the data is not in a centralized location, no fraud is expected, but federal lawsuit filed:	If the data is not in a centralized location, no fraud is expected, and no federal lawsuit filed:
PHI and SSN	a	901.37	459.15	789.08	459.15
	b	0.64	0.66	0.65	0.66
CC	a	650.41	338.74	338.74	1.4
	b	0.64	0.66	0.64	0.66

[5] Tables 3-5 and 3-6 were used in the financial impact baseline research (Algarni & Malaiya, 2016).

The cost factor based on F_{BC} is data breach costs:

Table 3-8. Cause of Data Breach Cost Factor (F_{BC}) [6]

Cause of Data Breach	Multiplier
Malicious Attack	1.19
Negligence	0.67
System glitch	0.69
Unsure	1 (default)

The cost factor based on F_E is sensitive data encryption:

Table 3-9. Sensitive Data Encryption Cost Factor (F_E)

Encryption Sensitive Data	Multiplier
Yes	0.51
No	1.05
Not Sure	1 (default)

The cost factor based on F_P is organization privacy:

- Crisis Management Cost Per Record:

$$CMCPR = [ax^{b-1}] \times F_{BCM}$$

where x is the number of impacted records, and F_{BCM} is business continuity management, which translates into whether an organization has an emergency response plan.

[6] Tables 3-8 and 3-9 were used in the financial impact baseline research (Algarni & Malaiya, 2016).

[7]Table 3-10. BCM Involvement Cost Factor

BCM Involvement	Multiplier
Yes	0.82
No	1.08
Unsure	1 (default)

- Regulatory and Industry Sanctions Cost Per Record:

$$SCPR = [ax^{b-1}]$$

where x is the number of impacted records, PCI is Payment Card Industry, PHI is Private Health Information, SSN is Social Security Number, CC is Credit Card, and the values for a and b are defined in the table below:

Table 3-11. Regulatory and Industry Sanctions Cost Factor

		Is PCI compliance an issue?	
		Yes	No
PHI and SSN	A	19,145	865,764
	B	0.43	0.02
CC	A	11,308	610,611
	B	0.47	0.03

[7] Tables 3-10 and 3-11 were used in the financial impact baseline research (Algarni & Malaiya, 2016).

- Class Action Lawsuit Cost Per Record:

$$CALCPR = [ax^{b-1}]$$

where x is the number of impacted records and values for a and b are defined in the table below:

[8]Table 3-12. Class Action Lawsuit Cost Factor

Decision Points:		Fraud Expected and Federal Class Action Lawsuit Filed?	Fraud expected, but no Federal Class Action Lawsuit Filed?	No Fraud expected, but Federal Class Action Lawsuit Filed?	Fraud Expected and Federal Class Action Lawsuit Filed?
PHI and SSN	a	0.36	0	0.09	0
	b	1.16	0	1.16	0
CC	a	5.12	6	5.68	6
	b	1.04	1	1.01	1

Now that the values are defined, the next step of the process involves using a separate dataset containing information about the cyber-attack. That dataset includes a description of the attack and the number of files affected. This dataset is then inputted into the model designed to predict CVSS scores. Once the CVSS scores are generated, the same dataset is passed through a script that employs the equations mentioned earlier. This script calculates the cost associated with the data breaches caused by cyber-attacks.

The absence of CVE entries can make it challenging to directly obtain the severity assessment of an attack from official vulnerability databases. Organizations can estimate the severity of cyber-attacks by utilizing a separate dataset and employing a model to

[8] Table 3-12 was used in the financial impact baseline research (Algarni & Malaiya, 2016).

predict CVSS scores even when CVE entries are not available. This allows them to quantify the potential impact of the attacks and calculate the associated costs of data breaches. Correlation analysis is done to compare the CVSS score with the cost of the data breach. Next, a correlation analysis between CVSS scores and the cost of data breaches was performed to provide insights into the relationship between the severity of a cyber-attack and its financial impact.

Chapter 4—Results

4.1 Introduction

This chapter presents the results based on the methodology proposed in Chapter 3. Chapter 4 assists in the evaluation of whether the research objectives were accomplished.

4.2 CVSS Score Prediction Results

4.2.1 Model Comparison

The first step taken was to recreate the CVSS baseline-research results. The results of that baseline research are in the table below:

Table 4-1. Results from Baseline Research

	BERT			RoBERTa			ALBERT			DeBERTa			DistilBERT			BART		
Category	ACC	F1	BA	ACC	F1	BA	ACC	F1	BA	ACC	F1	BA	ACC	F1	BA	ACC	F1	BA
Attack Vector	73.57	62.37	25.00	90.55	62.37	62.03	90.87	90.6	67.76	80.86	78.99	37.17	91.04	90.85	70.60	BART was not included in the CVSS-baseline research.		
Attack Complexity	92.26	88.55	50.00	92.26	88.55	50.00	92.26	88.55	50.00	92.26	88.55	50.00	95.16	94.32	70.74			
Privileges Required	66.78	53.48	33.33	66.86	83.49	67.61	66.78	53.48	33.33	66.86	63.65	41.15	86.21	85.91	75.09			
User Interaction	64.63	50.75	50.00	82.73	92.57	91.26	64.63	50.75	50.00	82.73	82.85	81.90	93.01	92.98	91.95			
Scope	83.64	76.19	50.00	96.02	95.91	90.20	83.64	76.19	50.00	83.64	76.19	50.00	96.28	96.19	90.98			
Confidentiality	59.15	43.97	33.33	86.20	85.90	81.08	86.18	85.98	81.74	86.20	86.05	82.37	86.29	86.16	82.67			
Integrity	61.52	50.10	55.84	86.93	86.85	84.41	51.36	34.85	33.33	51.37	34.87	33.34	87.46	87.42	85.52			
Availability	58.85	43.61	33.33	83.50	87.89	67.49	58.85	43.61	33.33	83.50	82.47	57.54	88.55	88.00	67.51			

The CVSS baseline-research results indicate that DistilBERT outperforms all other models across all categories and metrics. The model with the lowest performance is ALBERT, despite having the fewest parameters (11M) (Costa et al., 2022).

With the goal of assessing class inference based on vulnerability descriptions, the number of parameters seems to impact performance variability. Insufficient parameters

(ALBERT) and excessive parameters have the same impact of hindering the model's ability to learn. Upon initial inspection, BERT, ALBERT, and DeBERTa all appear to have similar performance. Digging into the dataset shows that this performance similarity is due to a dataset imbalance. In these cases, the models may prioritize higher accuracy. This prioritization is misleading because what is happening behind the scenes is that the model is producing the same output for every instance. When dealing with imbalanced datasets, accuracy cannot be used as an accessor for determining which model is the best fit; it is more appropriate to consider other metrics like balanced accuracy. Below are the dataset proportions:

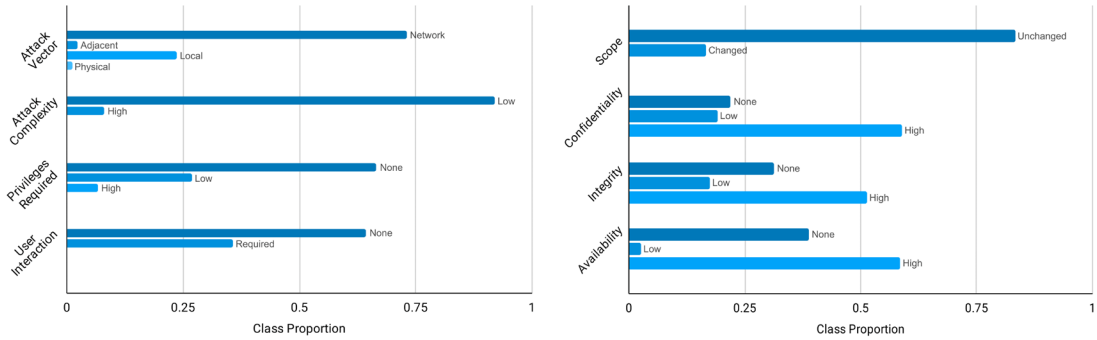

Figure 4-1. Dataset Imbalance.

Balanced accuracy considers both sensitivity and specificity, providing a more balanced assessment of the model's performance across different classes, especially when one class significantly outweighs the other in terms of sample size.

Following the analysis of the CVSS baseline research results, the performance of the models developed in this study was evaluated. Just as a recap, the CVSS-baseline

research developed eight distinct models based on the categories for each available NLP option (i.e., BERT produced a separate model for AC, AV, PR, UI, S, C, I, and A). Each of these eight models underwent independent training. One of the hypotheses suggested in this research is that enhanced performance can be attained through the implementation of shared learning methodologies; this leads to the consolidation of those models into a singular representation for each NLP alternative (i.e., BERT is comprised of a single model that encompasses all eight categories). The results of the models are below:

Table 4-2. Results from CVSS Prediction Model

	BERT			RoBERTa			ALBERT			DeBERTa			DistilBERT			BART		
Category	**ACC**	**F1**	**BA**	**ACC**	**F1**	**BA**	**ACC**	**F1**	**BA**	**ACC**	**F1**	**BA**	**ACC**	**F1**	**BA**	**ACC**	**F1**	**BA**
Attack Vector	83.35	80.86	41.29	89.26	65.12	62.15	83.62	81.52	44.11	79.35	78.27	38.47	92.59	92.22	71.29	45.00	48.49	32.66
Attack Complexity	92.28	91.51	64.88	92.65	89.76	53.59	93.17	91.41	59.95	83.32	85.13	56.50	95.70	95.11	75.08	78.00	78.88	43.33
Privileges Required	73.55	68.77	45.50	72.08	84.21	67.90	71.15	64.02	41.29	70.42	66.10	42.61	86.30	86.71	75.61	42.00	44.50	40.32
User Interaction	80.19	78.17	72.82	83.56	91.52	91.27	73.37	68.48	63.07	80.32	84.34	81.93	94.11	94.07	92.98	53.00	54.23	53.70
Scope	89.44	87.47	68.61	96.72	93.16	90.59	85.68	80.78	56.44	85.17	80.38	56.04	96.32	**96.16**	**89.63**	88.00	83.31	49.44
Confidentiality	74.68	73.90	66.22	87.74	87.26	82.14	70.04	64.83	52.26	87.98	87.35	83.03	88.52	88.21	82.86	50.00	49.36	36.08
Integrity	69.13	69.16	65.42	87.33	85.13	85.57	65.42	61.35	52.65	58.50	56.40	51.88	89.35	89.26	86.46	47.00	42.40	33.59
Availability	73.62	73.42	58.13	81.39	81.02	63.59	73.17	71.64	48.34	83.39	83.91	57.94	91.08	90.52	69.25	40.00	35.98	61.36

The green color represents the top-performing model. The red-highlighted numbers indicate instances where the new model's performance was inferior to the CVSS baseline-research study. A more straightforward comparison of the balanced accuracy study is below:

Table 4-3. Balanced Accuracy Comparison

Balanced Accuracy												
	BERT		RoBERTa		AlBERT		DeBERTa		DistilBERT		BART	
Categories	Baseline	Current	Baseline	Current	Baseline	Current	Baseline	Current	Baseline	Current	Baseline	Current
Attack Vector	25.00	41.29	62.03	62.15	67.76	44.11	37.17	38.47	70.60	71.29	BART was not included in the CVSS baseline research.	
Attack Complexity	50.00	64.88	50.00	53.59	50.00	59.95	50.00	56.50	70.74	75.08		
Privileges Required	33.33	45.50	67.61	67.90	33.33	41.29	41.15	42.61	75.09	75.61		
User Interaction	50.00	72.82	91.26	91.27	50.00	63.07	81.90	81.93	91.95	92.98		
Scope	50.00	68.61	90.20	90.59	50.00	56.44	50.00	56.04	90.98	**89.63**		
Confidentiality	33.33	66.22	81.08	82.14	81.74	52.26	82.37	83.03	82.67	82.86		
Integrity	55.84	65.42	84.41	85.57	33.33	52.65	33.34	51.88	85.52	86.46		
Availability	33.33	58.13	67.49	63.59	33.33	48.34	57.54	57.94	67.51	69.25		

The findings have brought to light several observations:

- In 90% of all models, shared learning led to an improvement in overall balanced accuracy.
- DistilBERT consistently outperformed other models, establishing its superiority throughout the evaluation process.
- When the accuracy was lower, but the F1 score was higher, it resulted in higher balanced accuracy. The F1 score considers both precision (also called positive predictive value) and recall (also called sensitivity or true positive rate)
- Throughout the evaluation, BART consistently demonstrated inferior performance compared to other models. While powerful, BART is known to be a more complex model, which can make it more challenging to fine-tune and optimize. For specific NLP tasks, simpler models might suffice and perform more efficiently. During the CVSS baseline research, additional vocabulary was introduced to the models to aid in fine-tuning. The variations included vocabulary additions of 5k, 10k, and 25k, respectively. Adding 5,000 words yields the best overall results among the different vocabulary variations tested. This suggests that adding more words to the vocabulary positively contributes to enhancing model performance, but additions more than that amount do not lead to incremental performance gains. Increasing the vocabulary size may cause the model's attention to be dispersed among less relevant words, ultimately hindering its performance. This effect is particularly noticeable in cases where the 25,000-word

addition performs worse than the baseline, such as in the Attack Vector and Attack Complexity categories (Costa et al., 2022).

Once the models predicted the values for each of the classes, the CVSS score predictions were calculated.

4.2.2 CVSS Score Prediction Results

Section 3.4 yielded prediction labels that correlate to the eight components required to compute the base CVSS score, which is the focus of this research. This score depends on the ISS, Impact, and exploitability. To recap, the formula is as follows:

$$ISS = 1 - [(1 - Confidentiality) * (1 - Integrity) * (1 - Availability)$$

The ISS is used to calculate Impact and Exploitability. Impact is determined by the scope.

- If the scope is unchanged,

$$Impact = 6.42 * ISS$$

- If the scope is changed,

$$Impact = 7.52 * (ISS - 0.029) - 3.25 * (ISS - 0.02)^{15}$$

Exploitability is determined by AC, AV, PR and UI.

$$Exploitability = 8.22 * \text{AV} * \text{AC} *$$

$$\text{PR} * \text{UI}$$

The values are derived from the metric values supplied by FIRST and are detailed in Section 3.2.4. To ensure accuracy and streamline calculations, a script was created to automatically compute the CVSS scores, thus avoiding errors and expediting the process.

```
# Loop through each row of the test data frame
for i in range(len(df_test_final)):
    scope = df_test_final['Scope(Predicted)'][i]
    attack_vector = df_test_final['Attack Vector(Predicted)'][i]
    attack_complexity = df_test_final['Attack Complexity(Predicted)'][i]
    privileges_required = df_test_final['Privileges Required(Predicted)'][i]
    user_interaction = df_test_final['User Interaction(Predicted)'][i]
    confidentiality = df_test_final['Confidentiality(Predicted)'][i]
    integrity = df_test_final['Integrity(Predicted)'][i]
    availability = df_test_final['Availability(Predicted)'][i]
    equation_dict['Attack Vector Value(Predicted)'].append(attack_vector_dict[attack_vector])
    equation_dict['Attack Complexity Value(Predicted)'].append(attack_complexity_dict[attack_complexity])
    equation_dict['Privileges Required Value(Predicted)'].append(privileges_required_dict[privileges_required][scope])
    equation_dict['User Interaction Value(Predicted)'].append(user_interaction_dict[user_interaction])
    equation_dict['Confidentiality Value(Predicted)'].append(confidentiality_integrity_availability_dict[confidentiality])
    equation_dict['Integrity Value(Predicted)'].append(confidentiality_integrity_availability_dict[integrity])
    equation_dict['Availability Value(Predicted)'].append(confidentiality_integrity_availability_dict[availability])

    iss = get_iss(confidentiality, integrity, availability)
    equation_dict['ISS(Predicted)'].append(iss)
    impact = get_impact(scope, iss)
    equation_dict['Impact(Predicted)'].append(impact)
    exploitability = get_exploitability(scope, attack_vector, attack_complexity, privileges_required, user_interaction)
    equation_dict['Exploitability(Predicted)'].append(exploitability)
    base_score = get_base_score(scope, impact, exploitability)
    equation_dict['Base Score(Predicted)'].append(base_score)
```

Figure 4-2. Code Snippet of CVSS Calculations.

A snippet of the output of that formula is below:

Table 4-3. Sample Output

	Attack Vector Value(Predicted)	Attack Complexity Value(Predicted)	Privileges Required Value(Predicted)	User Interaction Value(Predicted	Confidentiality Value(Predicted)	Integrity Value(Predicted)	Availability Value(Predicted)	ISS(Predicted)	Impact(Predicted)	Exploitability (Predicted)	Base Score(Pre
PHP source code, via the admin/downl	0.85	0.77	0.85	0.62	0.22	0.22	0	0.3916	2.726750844	2.83525473	6.1
(Contract Name: MyAdvancedToken7)	0.55	0.77	0.85	0.62	0.56	0.56	0.56	0.914816	5.87311872	1.83457659	7.8
tions on upload of .js files.	0.55	0.77	0.62	0.85	0.56	0.56	0.56	0.914816	5.87311872	1.83457659	7.8
yzer G2 allow remote attackers to byp	0.85	0.77	0.85	0.62	0.56	0.56	0.56	0.914816	5.87311872	2.83525473	8.8

The output columns consist of the text data description, individual predicted values for each metric, and the computed base score.

In the CVSS 3.1 scoring system, the category that has the greatest impact on the overall CVSS score is the "Impact" category. The "Impact" category considers the worst of the three impact values (C, I, or A) and plays a crucial role in determining the overall

CVSS score. If any of the impact values (C, I, or A) has high severity, it will significantly influence the overall score, making "Impact" the most impactful category in the CVSS 3.1 scoring system. A dummy dataset was created to demonstrate the influence of the "Impact" category in the CVSS scoring system, where the only values that changed were C, I, and A. The dataset is below:

Table 4-4. Sample Dataset

CVE Dummy ID	C	I	A	CVSS Score		CVE Dummy ID	C	I	A	CVSS Score
1	0	0	0	1		27	0.56	0.56	0.56	1.5123
2	0	0	0.22	1.0484		18	0.22	0.56	0.56	1.47544
3	0	0	0.56	1.3136		24	0.56	0.22	0.56	1.47544
4	0	0.22	0	1.0484		26	0.56	0.56	0.22	1.47544
5	0	0.22	0.22	1.08615		9	0	0.56	0.56	1.45158
6	0	0.22	0.56	1.36781		21	0.56	0	0.56	1.45158
7	0	0.56	0	1.3136		25	0.56	0.56	0	1.45158
8	0	0.56	0.22	1.36781		15	0.22	0.22	0.56	1.41009
9	0	0.56	0.56	1.45158		17	0.22	0.56	0.22	1.41009
10	0.22	0	0	1.0484		23	0.56	0.22	0.22	1.41009
11	0.22	0	0.22	1.08615		6	0	0.22	0.56	1.36781
12	0.22	0	0.56	1.36781		8	0	0.56	0.22	1.36781
13	0.22	0.22	0	1.08615		12	0.22	0	0.56	1.36781
14	0.22	0.22	0.22	1.1156		16	0.22	0.56	0	1.36781
15	0.22	0.22	0.56	1.41009		20	0.56	0	0.22	1.36781
16	0.22	0.56	0	1.36781		22	0.56	0.22	0	1.36781

17	0.22	0.56	0.22	1.41009		3	0	0	0.56	1.3136
18	0.22	0.56	0.56	1.47544		7	0	0.56	0	1.3136
19	0.56	0	0	1.3136		19	0.56	0	0	1.3136
20	0.56	0	0.22	1.36781		14	0.22	0.22	0.22	1.1156
21	0.56	0	0.56	1.45158		5	0	0.22	0.22	1.08615
22	0.56	0.22	0	1.36781		11	0.22	0	0.22	1.08615
23	0.56	0.22	0.22	1.41009		13	0.22	0.22	0	1.08615
24	0.56	0.22	0.56	1.47544		2	0	0	0.22	1.0484
25	0.56	0.56	0	1.45158		4	0	0.22	0	1.0484
26	0.56	0.56	0.22	1.47544		10	0.22	0	0	1.0484
27	0.56	0.56	0.56	1.5123		1	0	0	0	1

The left table displays the inputs into the script, representing the C, I, and A metrics, along with the output, which is the CVSS score calculated using the 3.1 CVSS formula. All other variables in the CVSS scoring formula were maintained consistent with a value of one. The same information is presented on the right side but sorted by the highest calculated CVSS score. A visual inspection of the data indicates that as the CIA values increase, the corresponding CVSS score also increases.

A multi-linear regression model was created to prove this statistically, and analysis was performed. The given information is the result of a multiple linear regression model.

Table 4-5. Regression Analysis

Feature	**Regression Coefficient**
Confidentiality	0.355381
Integrity	0.359329
Availability	0.342429

The coefficients represent the estimated effect of each feature (CIA) on the target variable (CVSS Score) based on the multiple linear regression model. These coefficients indicate the direction and magnitude of the relationship between each feature and the target variable. In this case, all three features (CIA) have positive coefficients, suggesting that an increase in the value of each feature will lead to an increase in the predicted target variable, the CVSS score.

Other metrics collected were r-squared, MSE, and RMSE.

- The R-squared value is 0.8096082982501055. In this case, the R-squared value of approximately 0.81 indicates that the three features included in the model can explain around 81% of the variance in the target variable. A higher R-squared value also indicates a better fit of the model to the data.
- The MSE value is 0.005899269679522155. A lower MSE indicates better model performance, as it represents more minor prediction errors.
- The RMSE value is 0.07680670335017742. Similar to the MSE, a lower RMSE indicates better model performance and a better fit for the data.

The given coefficients represent the estimated impact of each feature on the target variable, and the R-squared, MSE, and RMSE provide measures of the model's performance and how well it fits the data. A higher R-squared and lower MSE and RMSE indicate that the model is performing well in predicting the target variable based on the provided features.

4.3 Financial Impact Prediction Results

Algarni and Malaiya evaluated multiple data breach calculators using power regression models to derive a formula for assessing the financial impact of data breaches. The research was based on data from the Ponemon Institute; the statistics were based on the Ponemon Institute's "Cost of a Data Breach Report 2022":

- The cost of a data breach reached an all-time high of $4.35 million (Ponemon Institute, 2017).
- Critical infrastructure organizations (including healthcare) data breach costs averaged approximately USD 4.82 million (Ponemon Institute, 2017).
- Healthcare organizations faced a significant surge in breach costs that reached a record high of USD 10.10 million, an increase of nearly USD 1 million compared to previous years (Ponemon Institute, 2017).
- The global per-record cost of a data breach in healthcare in 2022 was USD 289; The probability of that cost was 12.7%.

$$\$289 * 12.7 = \$36.70$$

$36.70 value was used to calculate the cost for FC, which will be used in the next paragraph (Ponemon Institute, 2017).

- **NOTE**: Baseline research used 217 based on the current information at the time of the study. This praxis opted to use the current amount of 289.

The first step was understanding the formulas used by financial impact baseline research and creating a script based on the formulas identified in Section 3.3.2. The multipliers are used to determine the total cost per record. Consider the following example:

> A healthcare organization in the U.S. experienced a data breach due to employee negligence that compromised private health information and social security numbers. The data is encrypted on centralized servers. Fraud has been detected, and customers have filed a lawsuit requesting pay for damages. Federal class action lawsuit expected due to PCI compliance issues. The company has an engaged BCM team, and an enterprise-wide business response plan is being followed. Approximately 100,000 records were impacted.

Based on the description, the following values would be chosen based on the values provided in Chapter 3:

Table 4-6. Total Cost Per Record Calculation

Total Cost Per Record														
F_C *	F_I *	F_D	$\sum_{(i=1)}^{4}c$											
			ICPR					CMCPR			SCPR		CALCPR	
			a	b	F_{BC}	F_E	F_P	a	b	F_{BCM}	a	b	a	b
36.703	1.33	0.125	1532.8	0.57	0.67	0.51	1	184	0.105	0.82	19,145	0.43	0.36	1.16

Using the formulas, this equates to $32.69 per record. The calculations for the Cost Per Record are below:

$$CPR_1 = F_c \times F_I \times F_D \times \sum_{i=1}^{4} c$$

$$CPR_1 = 36.703 \times 1.33 \times 0.125 \times \sum_{i=1}^{4} c$$

$\sum_{i=1}^{4} c$ s comprised of the following formulas calculated as such:

$$ICPR = [ax^{b-1}] \times F_{BC} \times F_E \times F_P$$

$$ICPR = \left[(1532.8 \times 100000)^{(.57-1)}\right] \times 0.67 \times 0.51 \times 1 = 0.00010325$$

$$CMCPR = [ax^{b-1}] \times F_{BCM}$$

$$CMCPR = \left[(184 \times 100000)^{(0.105-1)}\right] \times 0.82 = 0.00000026$$

$$SCPR = [ax^{b-1}]$$

$$SCPR = (19{,}145 \times 100000)^{(0.43-1)} = 0.0000051195105$$

$$CALCPR = [ax^{b-1}]$$

$$CALCPR = (0.36 \times 100000)^{(1.16-1)} = 5.35807021$$

At this point, the CPR_1 formula can be resumed:

$$CPR_1 = 164 \times 1.33 \times 1 \times \sum_{i=1}^{4} c$$

$CPR_1 = 36.703 \times 1.33 \times 0.125 \times (0.00010325 + 0.00000026$

$+0.0000051195105 + 5.35807021$

$CPR_1 = 32.69$

To calculate the breach's total cost, the cost per record is multiplied by the amount of impacted files. For example, if the cost per record is 32.69, and the number of records impacted is 100,000:

Cost per record * Number of records impacted = Total cost

$32.69 * 100,000 = $3,269,000

Based on the statistics presented by the Ponemon Institute, the total cost of $3,269,000 falls within the acceptable range for a breach involving 100,000 records. The list of acceptable ranges is below:

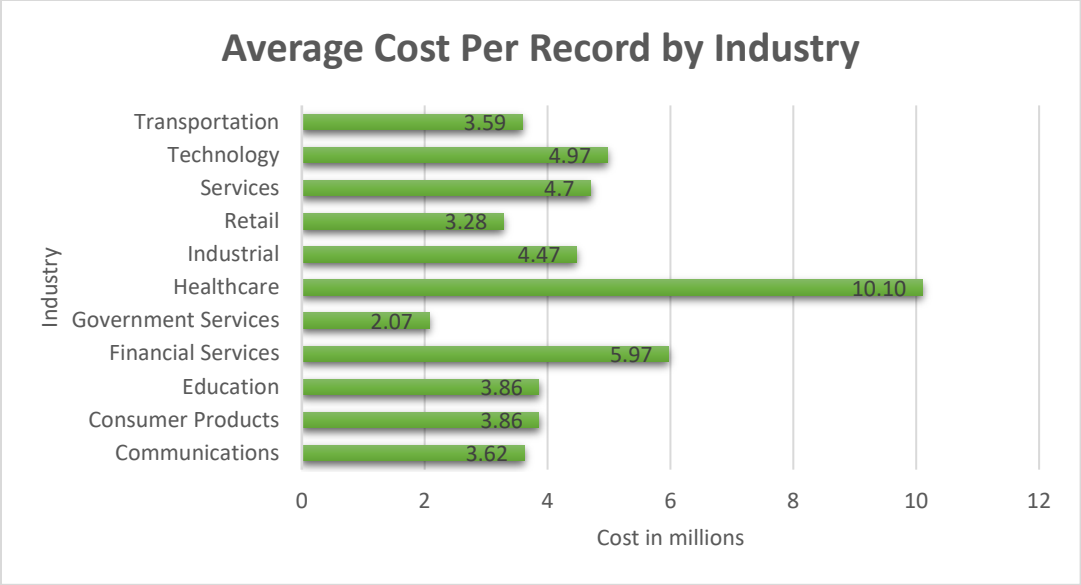

Figure 4-3. CPR by Industry.

After comprehending the formula and confirming that it accurately predicted the cost within the acceptable range, the subsequent task involved constructing a comprehensive data set to test all possible combinations of values. The following steps were followed:

- A script was developed to process a data file containing the relevant columns used in the formulas.
- The script was executed, and the dataset was imported.
- Necessary calculations were performed by the script.
- Outputs were generated that showed the total cost of a breach.

It was observed that each calculated value fell within the acceptable range for the corresponding industry, as specified by the Ponemon Institute and detailed in Figure 4-3 of this praxis.

Using the output of the dataset and calculated values, the following are the regression results:

Table 4-7. Regression Results – Total Cost Per Record

Regression Statistics	
Multiple R	0.985618
R Square	0.971444
Adjusted R Square	0.96175
Standard Error	8.999476

The regression analysis confirms that the regression model is a good fit.

- The multiple R-value of 0.985618394 indicates a highly significant positive linear relationship between the variables. F
- The R-square value of 0.971443618 reveals that approximately 97.14% of the variance is explained by the model.
- Although the adjusted R-square value of 0.961750143 is slightly lower than the regular R-squared, this is expected in models with multiple independent variables.
- The standard error value of 8.999476384 indicates that the predicted values are close to the actual values on average.

4.4 Financial Impact Prediction Results

CVSS and financial impact are two different concepts that measure two different scenarios, but combined, they can assist with assessing the overall risk of a security vulnerability. As detailed in this Praxis, the CVSS score can indicate the potential severity of a vulnerability, and this information helps organizations prioritize security efforts and allocate resources more effectively. CVSS scores assist with assessing financial impact by:

- If a vulnerability has known exploited vulnerabilities, it increases the likelihood of severe financial impact. If a vulnerability is spreading rapidly, the financial loss can be almost immediate.
- A vulnerability in a widely used software component might have a more significant financial impact due to its widespread deployment.
- Organizations with low-risk tolerance due to not having a continuity business management plan or adequate IT staffing and support can make wise decisions on which vulnerabilities to exploit to prevent potential financial losses.
- A high-severity vulnerability may have a lower financial impact if it is patched or mitigated; this decision cannot be made if organizations are unaware of CVSS scores and vulnerabilities.

It is important to note that while the CVSS score can provide helpful information, it is just one aspect of a comprehensive risk assessment. Organizations should conduct risk assessments, considering various factors, including the specific assets at risk, the business

environment, and potential financial consequences. Factors like insurance coverage and incident response capabilities can also play a role in managing financial impacts resulting from security vulnerabilities. Another nuance is that all data breaches are not reported as provided a CVE ID.

A third dataset was generated through manual efforts and utilized to establish a connection between the cost of CVSS scores. This dataset provides concise information about the vulnerability or breach and the corresponding number of impacted records. The dataset was then employed alongside the predictive model developed during this research to derive CVSS scores, and the financial impact was calculated using the formulas in 4.3. the financial impact.

The analysis revealed a correspondingly high risk of financial impact in 84.67% of cases with a high CVSS score. The findings are summarized in the chart below:

Table 4-8. Final Findings

TCPR	CVSS	TB
33.32617	6.1	$37,558.60
33.57656	3.2	$130,478.51
36.50298	5.9	$59,937.90
105.7202	7.5	$528,600.90
108.1236	7.2	$1,881,350.21
111.2415	7.7	$12,790,441.22
105.1954	8.1	$1,199,122.86
116.5163	7.5	$4,455,467.74
112.3405	8.4	$3,358,080.75

108.9971	8.7	$1,403,555.10
112.8909	8.3	$35,221,953.74
112.6539	8	$7,564,369.73
115.5908	8	$6,068,516.86
210.2845	9.3	$946,280.31
233.834	9.1	$2,447,540.57
215.8072	9.5	$7,322,986.79
226.9885	8.9	$794,459.84
231.9312	9.1	$1,296,727.18

Table 4-8 is a snippet of the data. As the CPR record increases, so does the CVSS score. The CVSS score was calculated using the predictive model proposed in this study. The CPR was calculated using the formulas in section 4.3. Of the 231 assessed in Table 4-8, 196 of them fell within the correct parameters identified in Figure 4-3.

Chapter 5—Discussion and Conclusions

5.1 Discussion

There were three research objectives identified during this study:

- Develop a predictive model using NLP to predict the fundamental values necessary to compute CVSS scores.
- Develop a script that automatically computes the CVSS score based on crucial category values predicted by the NLP model.
- Provide a formula that calculates financial impact.

The research objectives support the primary goal of this study to develop a predictive model that leaders in IoHT/IoMT domains can employ to evaluate the financial consequences of cyber-attacks on their respective organizations. The evaluation would consider a combination of CVSS scores and the number of financial records affected to determine the impact.

The research was divided into two parts. The first part focused on using NLP techniques to forecast CVSS scores based on textual descriptions. Numerous studies have previously attempted to make similar predictions using various machine-learning approaches. The second part aimed to estimate the financial impact resulting from data breaches. Determining the financial cost of a cybersecurity incident can be a daunting and complex task due to the multitude of factors involved, many of which may be challenging to quantify in monetary terms (i.e., damage to an organization's reputation, future sales loss, etc.). Numerous researchers have extensively discussed the financial aspects of

calculating the impact of a cybersecurity breach. However, finding research that derived an actual formula for this purpose proved to be a challenging task. This research incorporated findings from both components to create a comprehensive and robust method for assessing the financial impact of cyber-attacks.

5.2 Conclusions

As part of this study, three research questions were examined, and corresponding hypotheses were formulated to support or refute them.

Table 5-1. Summary of RQ, Hypotheses, and Outcomes

Research Questions	Hypotheses	Outcome / Comments
RQ1: Can regression analysis identify the most critical CVSS metrics for forecasting CVSS scores?	**H1:** Regression analysis can be used to determine which metrics (AV, AC, PR, UI, and CIA impact) are most important in deciding CVSS scores for IoHT/IoMT cyberattacks.	Regression analysis was conducted in 4.2.2, which provided that the C, I, and A factors are critical when determining CVSS scores.
RQ2: Can regression analysis be used to identify the correlation between CVSS metrics (AV, AC, PR, UI, and CIA impact)?	**H2:** Regression analysis can be used to determine the correlation between CVSS metrics.	There is no variability between trends/common vulnerability enumerators and CVSS with the dataset being used. Additionally, the regression analysis performed in section 4.2.2 identified correlations among the CVSS metrics.

RQ3: Can a predictive model be developed for IoMT / IoHT to predict CVSS metrics based on CVE test descriptions?	**H3:** A predictive model can be created to predict CVSS metrics with 90% accuracy.	The DistilBERT model consistently demonstrated the average highest accuracy rate of 90.5%, outperforming other NLP models.
RQ4: Can the accuracy of a predictive model be increased by combining several models into a single model?	**H4:** Several predictive models can be combined into a singular model to increase accuracy and reduce runtime.	This study demonstrated that the models developed achieved increased accuracy by enabling collaborative learning. The regression analysis findings indicated a correlation between the factors that influence CVSS.
RQ5: Can a model be developed for IoMT/IoHT to assess the financial impact?	**H5:** A model can assess the level of financial exposure of IoHT/IoMT cyberattacks.	Section 4.3 calculated the financial cost of data breaches and assigned CVSS scores to breaches that were previously unscored. This enabled regression tests to explore the correlation between CVSS scores and the financial impact of the breaches.

CVSS research and the financial impact of data breaches are interconnected and can be used to understand, prevent, and mitigate security risks. Accurately assessing the severity of vulnerabilities by calculating CVSS metrics has several benefits. Organizations can prioritize their security efforts and allocate resources based on the severity of the exposure. CVSS scores also assist with vulnerability management which can include identifying high-risk areas. Studying the financial impact of data breaches can inform CVSS researchers about the real-world implications of vulnerabilities.

Combining this knowledge from CVSS research and understanding the financial impact of data breaches, organizations can reduce the economic impact of data breaches, position their perspective organizations in a better place to make security decisions, and assist with understanding vulnerabilities.

5.3 Contributions to Body of Knowledge

The research presented in this paper provides several contributions to the body of knowledge. Those contributions are:

- Introduces a predictive model that enables collaborative learning, leading to improved model performance and runtime efficiency.
- Combines financial cost calculations and CVSS scores, allowing organizations to understand the monetary implications associated with different vulnerability severity levels.

5.4 Recommendations for Future Research

There are several areas recommended for future research. Those recommendations are:

- Evaluate the difference between CVSS v3.1 and CVSS v4.0. CVSS v4.0 is an updated scoring system version that introduces notable changes and enhancements. One significant change is the inclusion of threat intelligence and environmental metrics. This revision acknowledges that vulnerabilities cannot be assessed solely based on their intrinsic

characteristics but must also consider external factors such as the threat landscape and the specific environment in which the vulnerability exists.

- This research focuses on the base metric group only. While assessing the base metric group alone can provide valuable insights into vulnerability severity and its inherent properties, it is essential to acknowledge that the complete CVSS score, which includes the temporal and environmental metric groups, offers a more comprehensive and contextual assessment of vulnerabilities.
- The formula utilized to calculate financial impact was based on datasets from the 2015 timeframe. Conducting an analysis to investigate whether the metric values used in determining the total cost have undergone any changes would be advantageous. Additionally, future research could reassess existing data breach calculators to explore potential evolutions in their methodologies.
- The existing method for calculating financial impact linked to a data breach excludes maintenance, upgrades, insurance premiums, and insurance settlements expenses. Creating a predictive model that incorporates this information would enhance the accuracy and comprehensiveness of the cost estimation.
- Just like the predictive model developed for forecasting CVSS scores through text descriptions using NLP, a new model should be established to predict the financial consequences of data breaches using the formulas presented in this research. Rather than depending on human judgment for

assessing the formula's values, efforts should be focused on streamlining the process through NLP technology.

References

A. F. Adoma, N. -M. Henry and W. Chen, "Comparative Analyses of Bert, Roberta, Distilbert, and Xlnet for Text-Based Emotion Recognition," *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, Chengdu, China, 2020, pp. 117-121, doi: 10.1109/ICCWAMTIP51612.2020.9317379.

A. Ramos, M. Lazar, R. H. Filho and J. J. P. C. Rodrigues, "Model-Based Quantitative Network Security Metrics: A Survey," in *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2704-2734, Fourthquarter 2017, doi: 10.1109/COMST.2017.2745505.

Algarni, A. M., & Malaiya, Y. K. (2016). A consolidated approach for estimation of data security breach costs. In Proceedings of the 2016 2nd International Conference on Information Management (ICIM), London, UK, 7–8 May 2016

Algarni, A. M., Thayananthan, V., & Malaiya, Y. K. (2021). Quantitative Assessment of Cybersecurity Risks for Mitigating Data Breaches in Business Systems. NATO Advanced Science Institutes Series E: Applied Sciences, 11(8), 3678.

Ameri, K., Hempel, M., Sharif, H., Lopez Jr., J., & Perumalla, K. (2021). CyBERT: Cybersecurity Claim Classification by Fine-Tuning the BERT Language Model. Journal of Cybersecurity and Privacy, 1(4), 615-637. https://doi.org/10.3390/jcp1040031

Attaran M. (2023). The impact of 5G on the evolution of intelligent automation and industry digitization. Journal of ambient intelligence and humanized computing, 14(5), 5977–5993. https://doi.org/10.1007/s12652-020-02521-x

Balakrishnan, V., Shi, Z., Law, C. L., Lim, R., Teh, L. L., & Fan, Y. (2022). A deep learning approach in predicting products' sentiment ratings: a comparative analysis. The Journal of supercomputing, 78(5), 7206–7226. https://doi.org/10.1007/s11227-021-04169-6

Cartwright AJ. (2023). The elephant in the room: cybersecurity in healthcare. J Clin Monit Comput. 2023 Apr 24:1–10. doi: 10.1007/s10877-023-01013-5. Epub ahead of print. PMID: 37088852; PMCID: PMC10123010.

Chase, M., & Coley, S. C. (2020). Rubric for Applying CVSS to Medical Devices. Retrieved from https://www.mitre.org/news-insights/publication/rubric-applying-cvss-medical-devices

Chen, H.; Liu, J.; Liu, R.; Park, N.; Subrahmanian, V. VEST: A System for Vulnerability Exploit Scoring & Timing. In Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19), Macao, China, 10–16 August 2019; pp. 6503–6505.

Costa, J. C., Roxo, T., Sequeiros, J. B. F., Proença, H., & Inácio, P. R. M. (2022). Predicting CVSS Metric via Description Interpretation. IEEE Access, 10.1109/ACCESS.2022.3179692. doi: 10.1109/ACCESS.2022.3179692

Coventry, L., & Branley, D. (2018). Cybersecurity in healthcare: A narrative review of trends, threats and ways forward. *Maturitas*, *113*, 48-52. https://doi.org/10.1016/j.maturitas.2018.04.008

Cyber Defense Magazine. (2017). FDA Recalls 465,000 Pacemakers Open to Cyber Attack. Retrieved from https://www.cyberdefensemagazine.com/fda-recalls-465000-pacemakers-open-to-cyber-attack/ on January 23, 2023.

Dreyer, P., Jones, T., Klima, K., Oberholtzer, J., Strong, A., Welburn, J. W., & Winkelman, Z. (2018). Estimating the Global Cost of Cyber Risk: Methodology and Examples. Santa Monica, CA: RAND Corporation.

FIRST. (2023a). A Complete Guide to the Common Vulnerability Scoring System Version 2.0. Retrieved from https://www.first.org/cvss/v2/guide

FIRST. (2023b). A Complete Guide to the Common Vulnerability Scoring System (CVSS) v1 Archive https://www.first.org/cvss/v1/guide

FIRST. (2023c). Common Vulnerability Scoring System v3.0: User-Guide. Retrieved from https://www.first.org/cvss/v3.0/user-guide

FIRST. (2023d). Common Vulnerability Scoring System v3.1: User-Guide. Retrieved from https://www.first.org/cvss/v3.1/user-guide

FIRST. (2023e). CVSS - Common Vulnerability Scoring System. Retrieved from https://www.first.org/cvss/

Furnell, S., Heyburn, H., Whitehead, A., & Shah, J. N. (2020). Understanding the full cost of cyber security breaches. Computer Fraud & Security, 2020(12), 6-12. https://doi.org/10.1016/S1361-3723(20)30127-5

Hong Leow, Seuk Wai Phoong, & Wai Khuen Cheng. (2022). State of the art: a review of sentiment analysis based on sequential transfer learning. *Artificial Intelligence Review, 42*(2).

IBM Corporation, Ponemon Institute. (2021). Cost of a Data Breach Report 2021.

Ponemon Institute.

Jireh Yi-Le Chan, Khean Thye Bea, Steven Mun

Jones, G., Bryant, A., & Wright, J. (2022). Democratizing Global Health Care Through Scalable Emergent (Beyond the Mobile) Wireless Technologies. JMIR Biomedical Engineering, 7(1), e31079. https://biomedeng.jmir.org/2022/1/e31079 DOI: 10.2196/31079

Joshi, C., & Singh, U.K. (2017). Information Security Risks Management Framework – A Step Towards Mitigating Security Risks in University Network. Journal of Information Security and Applications, 25, 128-137. https://doi.org/10.1016/j.jisa.2017.06.006

Kenton, J. D. M. W. C., & Toutanova, L. K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of naacL-HLT (Vol. 1, p. 2).

Koroteev, M. (2021). BERT: A Review of Applications in Natural Language Processing and Understanding. Retrieved from arXiv preprint server. (Accession No. arXiv.2103.11943)

Lewis, M. T., Liu, Y., Goyal, N., Ghazvininejad, M., Stoyanov, V., & Zettlemoyer, L. (2020). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 703-713. https://aclanthology.org/2020.acl-main.703/

Locatelli M, Seghezzi E, Pellegrini L, Tagliabue LC, Di Giuda GM. Exploring Natural Language Processing in Construction and Integration with Building Information Modeling: A Scientometric Analysis. *Buildings*. 2021; 11(12):583. https://doi.org/10.3390/buildings11120583

M. Cadariu, E. Bouwers, J. Visser and A. van Deursen, "Tracking known security vulnerabilities in proprietary software systems," *2015 IEEE 22nd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, Montreal, QC, Canada, 2015, pp. 516-519, doi: 10.1109/SANER.2015.7081868.

M. Dave, "Internet of Things Security and Forensics: Concern and Challenges for Inspecting Cyber Attacks," *2022 Second International Conference on Next Generation Intelligent Systems (ICNGIS)*, Kottayam, India, 2022, pp. 1-6, doi: 10.1109/ICNGIS54955.2022.10079829.

McGlade, D., & Scott-Hayward, S. (2019). ML-based cyber incident detection for

Electronic Medical Record (EMR) systems. Smart Health, 12, 3-23. ISSN 2352-6483. https://doi.org/10.1016/j.smhl.2018.05.001

Miller, M., Kisiel, A., Cembrowska-Lech, D., Durlik, I., & Miller, T. (2023). IoT in water quality monitoring—Are we really here? Sensors, 23(2), 960. https://doi.org/10.3390/s23020960

National Institute of Standards and Technology. (2021a). National Vulnerability Database (NVD). Retrieved from https://nvd.nist.gov/vuln

National Institute of Standards and Technology. (2021b). Cybersecurity Framework. Retrieved from https://www.nist.gov/cyberframework

Ponemon Institute. (2017). Cost of Data Breach Study: United States. Retrieved on February 19, 2018 from https://www.ponemon.org/news-updates/blog/security/2017-cost-of-data-breach-study-united-states.html

Rasmy, L., Xiang, Y., Xie, Z., Tao, C., & Zhi, D. (2021). Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. npj Digital Medicine, 4(1), 86. https://doi.org/10.1038/s41746-021-00455-y

Rocca, R., Tamagnone, N., Fekih, S., Contla, X., & Rekabsaz, N. (2023). Natural language processing for humanitarian action: Opportunities, challenges, and the path toward humanitarian NLP. Frontiers in big data, 6, 1082787. https://doi.org/10.3389/fdata.2023.1082787

Ruohonen, J. (2019). A look at the time delays in CVSS vulnerability scoring. Applied Computing and Informatics, 15, 129-135.

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications, and Research Directions. SN Computer Science, 2(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Scarfone K., Souppaya M., Cody A., Orebaugh A. (2008). NIST SP 800-115 Technical Guide to Information Security Testing and Assessment. Retrieved on February 27, 2023 from https://doi.org/10.1016/j.cose.2021.102376

Seh, A. H., Zarour, M., Alenezi, M., Sarkar, A. K., Agrawal, A., Kumar, R., & Khan, R. A. (2020a). Healthcare Data Breaches: Insights and Implications. Healthcare (Basel, Switzerland), 8(2), 133. https://doi.org/10.3390/healthcare8020133

Seh, I., Hussain, M., & Salah, K. (2020b). Healthcare Cybersecurity: Current Status, Future Directions, and a Proposed Model. Journal of medical systems, 44(10), 183. doi: 10.1007/s10916-020-01648-4

Shi, F., Kai, S., Zheng, J., & Zhong, Y. (2022). XLNet-Based Prediction Model for CVSS Metric Values. Applied Sciences, 12, 8983. https://doi.org/10.3390/app12188983

Singla, S. (2020). Comparative analysis of transformer based pre-trained NLP Models. International Journal of Computer Science and Engineering, 8, 40–44. https://doi.org/10.5121/csit.2021.110111

T. Floyd, M. Grieco and E. F. Reid, "Mining hospital data breach records: Cyber threats to U.S. hospitals," *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, Tucson, AZ, USA, 2016, pp. 43-48, doi: 10.1109/ISI.2016.7745441.

US Food and Drug Administration. (2017). FDA Safety Communication: Cybersecurity vulnerabilities identified in St. Jude Medical's implantable cardiac devices and Merlin@home Transmitter. https://www.fda.gov/medical-devices/safety-communications/cybersecurity-vulnerabilities-identified-st-jude-medicals-implantable-cardiac-devices-and-merlinhome

Walker, T. (2018). Interoperability a must for hospitals, but it comes with risks. Retrieved on February 28, 2018 from http://managedhealthcareexecutive.modernmedicine.com/managed-healthcare-executive/news/interoperability-must-hospitals-it-comes-risks