# Using NLP to Predict the Severity of Cyber Security Vulnerabilities

Saba Janamian (sjanamia@eng.ucsd.edu), Bryan Cook (b1cook@eng.ucsd.edu), James Logan (jlogan@eng.ucsd.edu)
Teck Lim (twlim@eng.ucsd.edu), Ivan Ulloa-Garcia (iulloaga@eng.ucsd.edu), MAS DSE Students
Faculty Advisors: Dr. Amarnath Gupta, Dr. Ilkay Altintas

University of California, San Diego - Jacobs School of Engineering

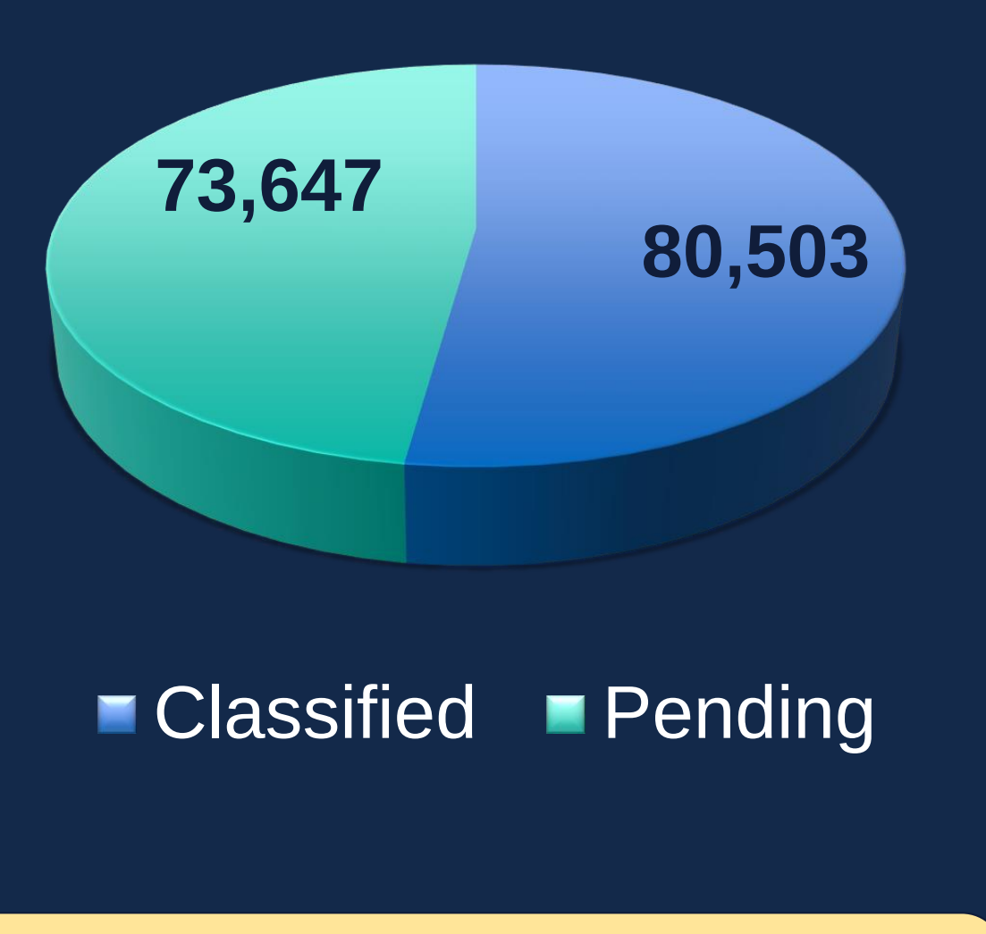
## Problem Statement

Cyber-attacks continue to be one of the world’s foremost safety and economic threats, and, in recent years, have become more numerous and severe. Cybersecurity engineers use industry-standard “Common Vulnerabilities and Exposure” (CVE) records to understand, and address known threats.

CVE records can be found as part of the “National Vulnerability Database” (NVD) where they generally contain “Common Vulnerability Scoring System” (CVSS) scores, which indicate a human-determined level of severity. These scores are important to cybersecurity engineers in threat prioritization. Unfortunately, nearly half of all CVE records have not yet been assigned CVSS v3 scores, a critical component of the overall CVSS score.

The **VulnerWatch** product is introduced as a machine learning solution for predicting CVSS v3 scores. Bidirectional Encoder Representation (BERT) is used on CVE record text descriptions to predict eight metrics that, in aggregate, indicate a CVSS v3 score. VulnerWatch provides the user with a prioritized list of CVE records that do not have human-determined CVSS v3 scores, along with a predicted score. It also allows the engineer to manually enter text describing threats and receive a predicted CVSS v3 score in near real-time.

**CVE Severity Classification**



## Data Science Pipeline

### 1. Acquire

- CVE records are acquired directly using the NVD database API in JSON format.

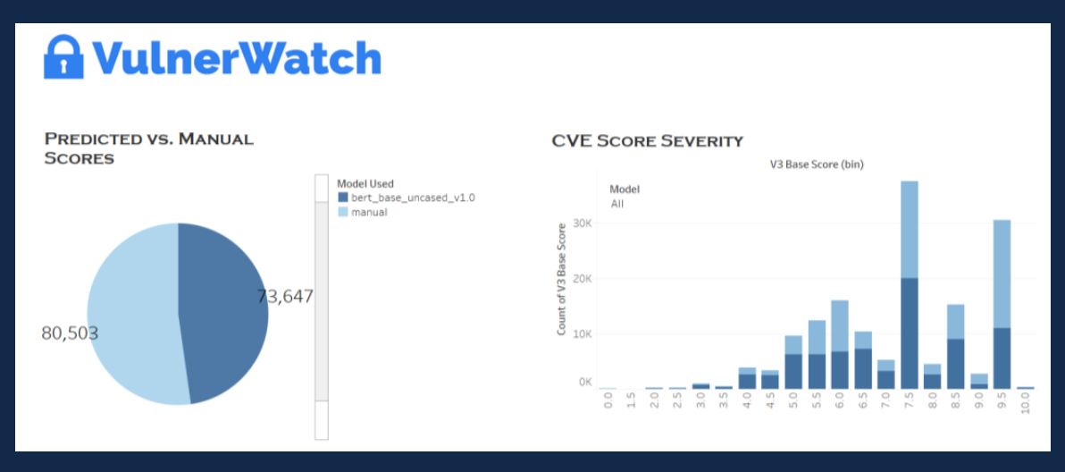
NVD
API
CVE
JSON
PostgreSQL

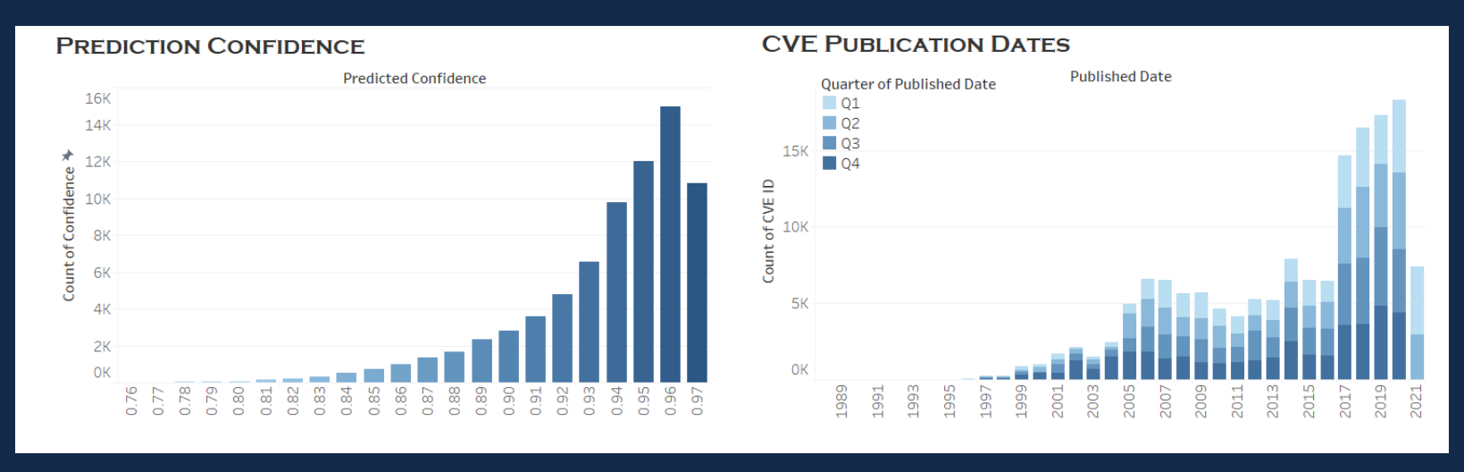
### 2. Prepare

- CVE dataset is pre-processed and added to Postgres database by extracting 4 key features: **1) CVE ID, 2) Date, 3) CVSS score, 4) CVE Text Description.**
- Text descriptions are formatted in preparation for BERT model fine tuning.

### 3. Analyze

- CVE dataset is compared with previous dataset to check for new entries.
- Entries with missing CVSS scores are identified.
- VulnerWatch visualizer is used to extract new insights from the addition of new entries.

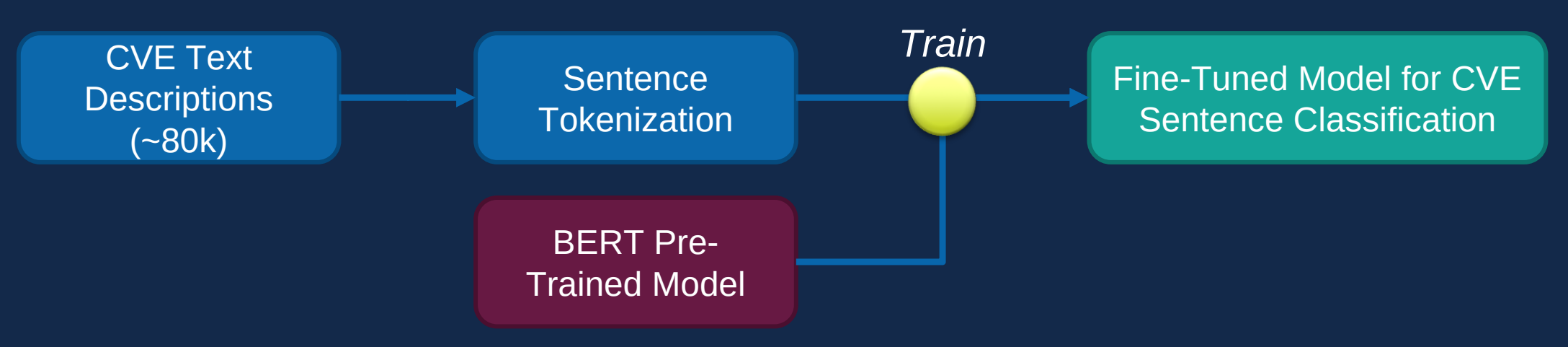




### 4. Scale

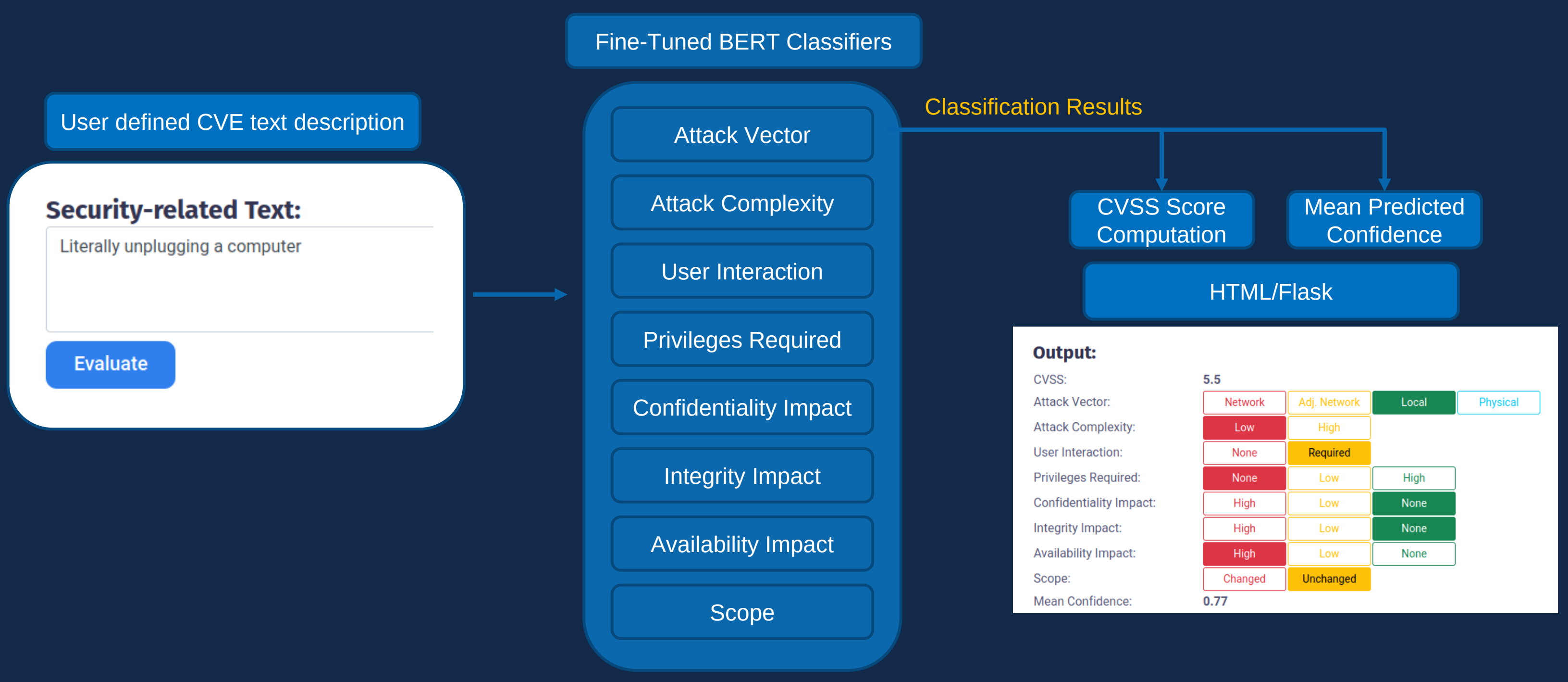
- AWS cloud servers with GPU access used periodically to fine-tune 8 sentence classification models.
- A BERT pre-trained model (bert-uncased) and over 80k CVE text descriptions are utilized to fine-tune all 8 classifiers to predict each of the CVSS metrics.

**BERT for Sentence Classification**


CVE Text Descriptions (~80k)
Sentence Tokenization
Train
Fine-Tuned Model for CVE Sentence Classification
BERT Pre-Trained Model

### 5. Report

- User enters cybersecurity vulnerability related descriptions into the VulnerWatch API.
- A CVSS score is provided as well as each of the metric predicted classifications.
- An average confidence value for the predictions is also shown.
- HTML/Flask and Tableau dashboards provide the user with information about CVEs and predicted scores.



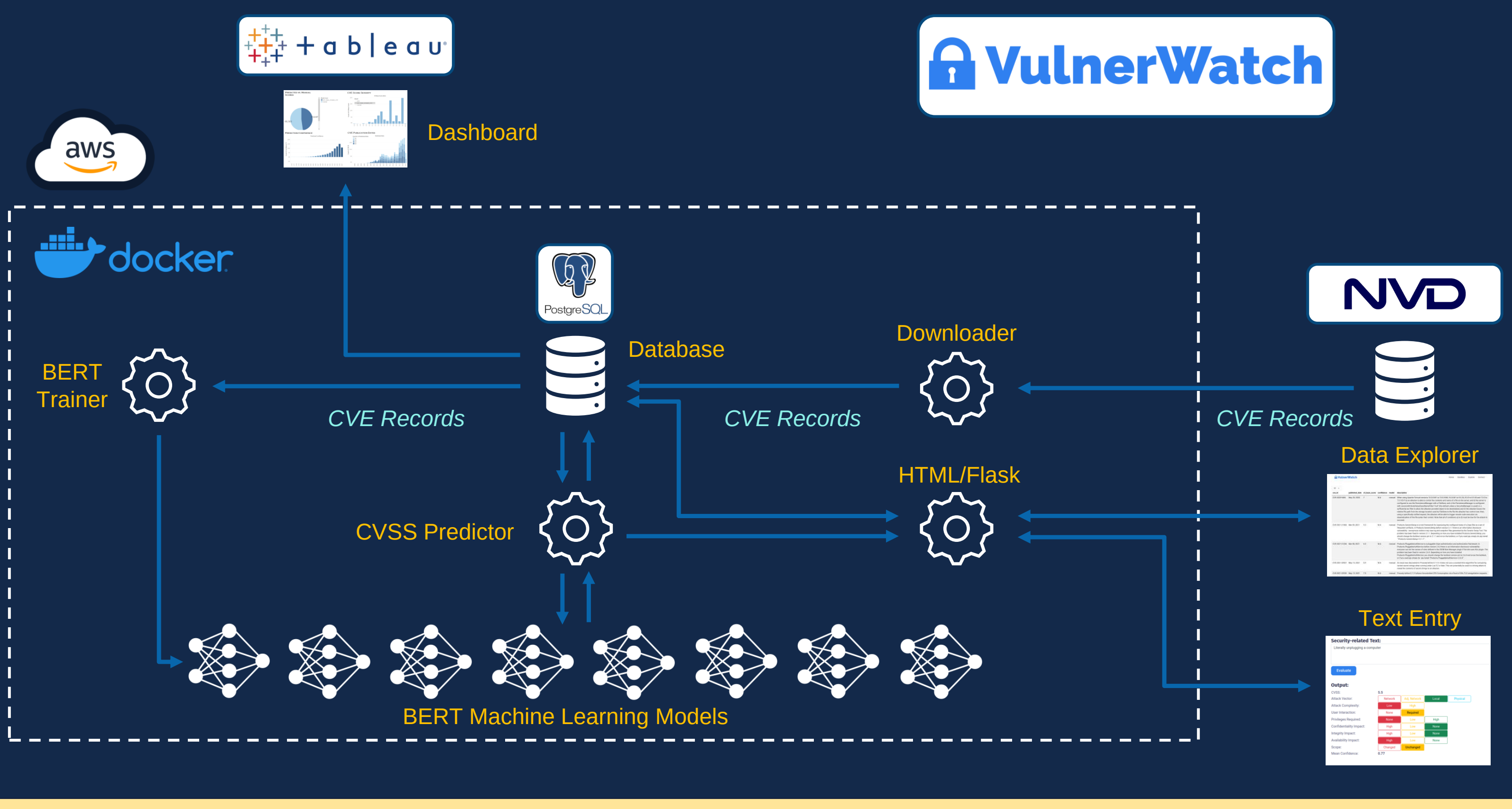
Security-related Text:
Literally unplugging a computer

Output:

| | | | | |
|---|---|---|---|---|
| CVSS: | 5.5 | | | |
| Attack Vector: | Network | Adj. Network | Local | Physical |
| Attack Complexity: | Low | High | | |
| User Interaction: | None | Required | | |
| Privileges Required: | None | Low | High | |
| Confidentiality Impact: | High | Low | None | |
| Integrity Impact: | High | Low | None | |
| Availability Impact: | High | Low | None | |
| Scope: | Changed | Unchanged | | |
| Mean Confidence: | 0.77 | | | |

## Solution Architecture

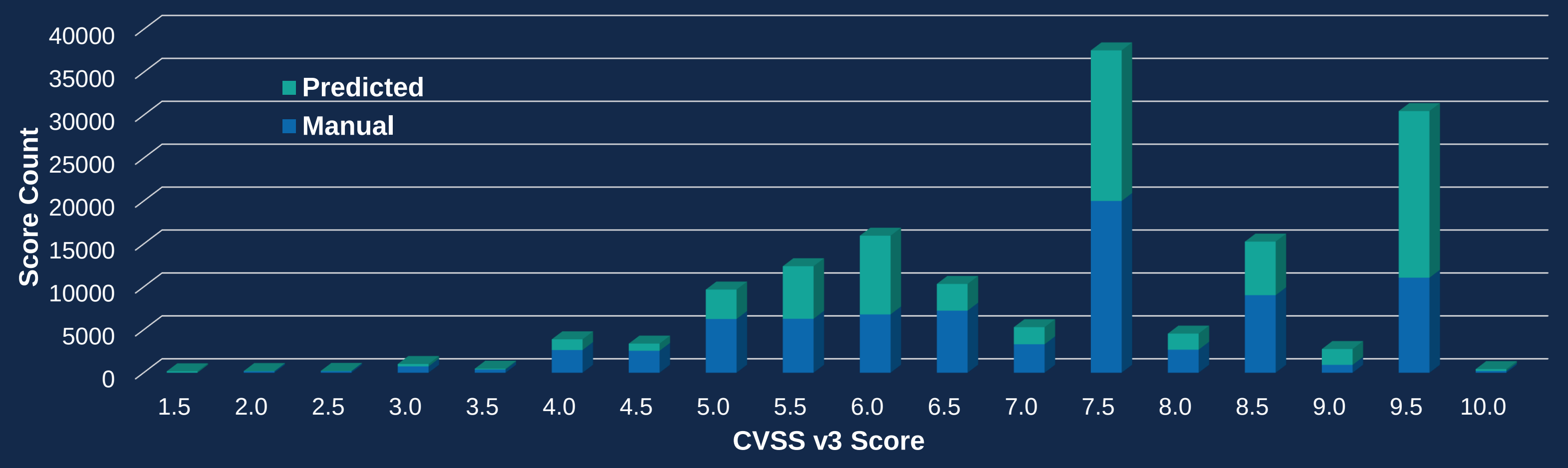
1. **System Design:** The VulnerWatch product components have been implemented to maximize resource efficiency by switching on/off Docker containers when needed.
2. **Portability:** Compartmentalized containers using Docker for BERT training, acquisition and predictions, user interface, and database management allows for reduced operating costs as well as fast deployment in multiple environments.
3. **ETL Pipeline:** 1) New CVE Dataset is downloaded using the NVD API and stored in a PostgreSQL database. 2) User defined text descriptions are captured by the HTML/Flask UI where a CVSS predictor executes 8 separate classification models to obtain a CVSS score. 3) Information may be queried and analyzed using the HTML/Flask user interface dashboard and database viewer.
4. **Training Pipeline:** Periodic training events are employed using the latest CVE datasets.



## Key Insights

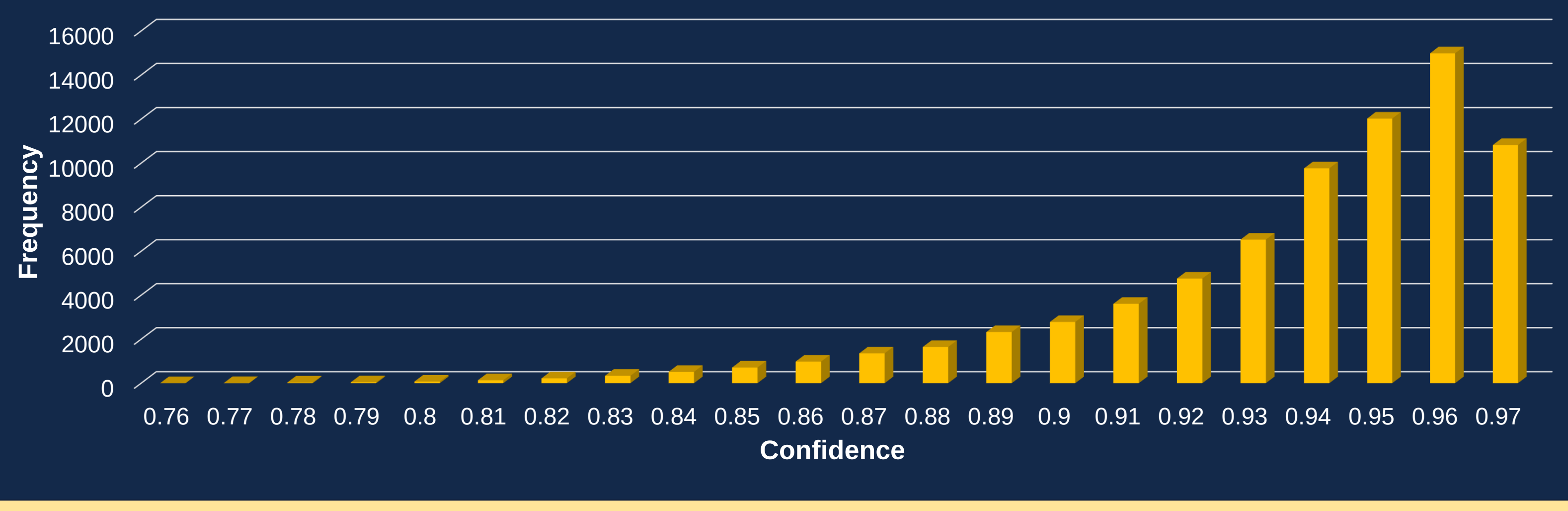
1. **Significant number of CVEs remain without assigned CVSS severity scores.** As of April 2020, 48% (~73k) of all CVE records do not yet have human-ascribed CVSS severity scores.
2. **Large number of high severity scores predicted.** About 29% (~21k) of VulnerWatch predicted CVSS scores have a severity of 9.0 or higher.

**CVE Score Severity**



3. **High prediction accuracy and confidence using BERT for NLP.** It was determined that VulnerWatch can provide a list of CVE issues with machine-predicted scores with an accuracy that should provide good utility for engineers. The accuracy of predictions for metrics determining CVSS v3 scores is favorable, averaging close to **0.9**, with similar levels of precision and recall. Resultant CVSS v3 score predictions are also favorably accurate **(MSE = 1.27, MAE = 0.5, R2= 0.51).** Similarly, over 90% of CVSS predictions have a mean predictive confidence score of **0.9** or higher.

**Confidence of Predictions**



## Acknowledgement

- Dr. Amarnath Gupta – Faculty Advisor
- Dr. Ilkay Altintas – Course Instructor
- Special thanks to all administrative members of the Masters of Advanced Studies program, the Data Science staff and Jacobs School of Engineering at UCSD for their unconditional support.

## References

1. McCormick, Chris. BERT Introduction and Tutorials - https://mccormickml.com/
2. Wu, Zhengxuan, and Desmond C. Ong. “On Explaining Your Explanations of BERT: An Empirical Study with Sequence Classification.” 2021, doi:arXiv:2101.00196.