Threat-Based Vulnerability Management: Mapping CVEs to the MITRE ATT&CK Framework

Logan McMahon
School of Electronics, Electrical Engineering and Computer Science
Queen's University Belfast, United Kingdom
e-mail: lmcmahon25@qub.ac.uk

Oluwafemi Olukoya
School of Electronics, Electrical Engineering and Computer Science
Queen's University Belfast,, United Kingdom
e-mail: o.olukoya@qub.ac.uk

***Abstract*—Mapping Common Vulnerabilities and Exposures (CVEs) to the MITRE Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) framework plays a crucial role in cybersecurity, particularly in threat mitigation and risk management. Accurate and automated CVE-to-ATT&CK mapping enables defenders to better assess the risks posed by emerging vulnerabilities. Prior work has relied primarily on CVE descriptions to establish links to relevant tactics and techniques. However, these approaches struggle when descriptions are incomplete or poorly written. This research proposes that enriching CVE descriptions with extended features, such as exploitability scores, software weaknesses, system and software identifiers, attack patterns, and classification data, substantially improves mapping accuracy. In unsupervised evaluations, this enrichment increased correct mappings by 42 % to 66.7% and reduced misclassifications by 6%. In supervised experiments, the proposed SecRoBERTa model significantly outperformed prior work. While baseline models achieved a weighted F1 score of 78.88%, the fully extended and *Optuna*-tuned version reached 93.47%, marking a 14.6% improvement. These results demonstrate the effectiveness of combining structured feature enrichment with hyperparameter optimization to enhance the accuracy and reliability of CVE-to-ATT&CK mappings.**

Keywords-MITRE ATT&CK; CVE; Vulnerability; Machine Learning; Data Augmentation; Threat Intelligence.

I. Introduction

In 2024, 40,077 Common Vulnerabilities and Exposures (CVEs) were published, a 39% increase from 2023, underscoring the growing challenge organizations face in managing vulnerabilities at scale [1]. Studies show that most organizations are only able to remediate 10% to 15% of open vulnerabilities each month, leaving a persistent backlog [2]. While only an estimated 1% to 6% of CVEs are actively exploited, these few can have severe consequences [3][4]. According to the Mandiant M-Trends 2025 Report [5], vulnerability exploitation was the most common initial attack vector observed in incident response investigations, emphasizing the importance of effective vulnerability prioritization.

Since it is neither feasible nor necessary to remediate every vulnerability, organizations are shifting toward risk-based vulnerability prioritization. This approach focuses on addressing vulnerabilities that pose the greatest risk, incorporating threat intelligence to enable a threat-informed defense [6]. Central to this approach is the MITRE ATT&CK framework, a widely adopted knowledge base of adversary tactics and techniques derived from real-world observations [7][8]. Within this framework, tactics represent high-level attacker goals (such as *Initial Access* or *Persistence*), while techniques describe how those goals are achieved (e.g., *exploiting a public-facing application or executing a script*).

Mapping CVEs to MITRE ATT&CK tactics and techniques allows defenders to better understand the potential impact of unpatched vulnerabilities, prioritize them based on adversary behavior, and align vulnerability management with real threat scenarios [9]. For example, prioritizing CVEs linked to tactics like *privilege escalation* or *lateral movement* can help security teams mitigate high-impact risks more effectively. However, given the rapid growth of CVEs, manually labeling each with ATT&CK mappings is infeasible. This highlights the urgent need for automated solutions to support scalable, threat-informed vulnerability management.

Recent research [10]–[16] has increasingly focused on automating the mapping of CVEs to MITRE ATT&CK techniques. Most existing methods rely heavily on CVE descriptions, with some efforts incorporating additional data, such as Common Vulnerability Scoring System (CVSS) vectors and Common Weakness Enumeration (CWE) identifiers [17][18]. However, prior studies have shown that patterns derived from CWE and CVSS can be unreliable, and vulnerability descriptions themselves are often inconsistent, incomplete, outdated, or inaccurate [17][19]–[22].

A major challenge remains in accurately mapping CVEs with poor-quality descriptions, as these often lack sufficient detail about exploitation methods or impact. To address these limitations, recent approaches have explored using Large Language Models (LLMs) to infer missing or unclear information, though gaps in domain knowledge constrain these methods and the complexity of vulnerability language [23].

This research proposes a comprehensive, automated approach to mapping CVEs to MITRE ATT&CK tactics, formulated as a multilabel classification problem that integrates structured data to enhance accuracy, particularly when descriptive fields are limited or ambiguous. The primary contributions of this research are as follows:

- **Extended Unsupervised Mapping Pipeline:** We adapt and expand the SMET framework [13] to operate on a larger, feature-enriched CVE dataset, using mappings from the Centre for Threat-Informed Defence [24]. By

integrating pre-processed CVSS, CWE, and Common Platform Enumeration (CPE) data, the modified pipeline achieves measurable improvements in full and partial mappings without requiring labelled data.

- **Creation of an Enriched Dataset for Supervised Learning:** We compile an extended dataset by incorporating structured data from the National Vulnerability Database (NVD), Common Attack Pattern Enumeration and Classification (CAPEC), and Exploit Prediction Scoring System (EPSS). This dataset enables systematic evaluation of the contribution of each feature to the CVE-to-ATT&CK mapping process.
- **Systematic Feature Evaluation:** We perform a detailed feature importance analysis to assess the impact of each added feature on model performance. This includes tactic-level analysis, particularly focusing on historically difficult-to-predict classes, such as *Initial Access*, *Impact*, *Collection* and *Reconnaissance* [14].
- **Hyperparameter Optimization with Optuna:** We apply *Optuna* [25] to fine-tune model hyperparameters, resulting in notable performance gains on the extended dataset and highlighting the importance of optimization in supervised models.
- **Public Release of Resources:** All datasets, code, and supplementary materials are made publicly available via the project's GitHub repository to support reproducibility. We present and release an extended dataset comprising 7,328 CVE entries, each enriched with CWE, CPE, and CVSS information, and optionally annotated with CAPEC and EPSS data.

The rest of the paper is structured as follows: Section II reviews related work on supervised and unsupervised approaches for mapping CVEs to MITRE ATT&CK tactics and techniques. Section III outlines the methodology, including data collection, preprocessing, model development, and evaluation. Section IV presents the research objectives and hypotheses, and the experimental results. In Section V, we interpret the research findings and the implications. Section VI discusses the limitations of the study, and Section VII concludes with a summary and directions for future work.

II. Related work

A. Supervised Mapping

Existing approaches to mapping CVEs to the MITRE ATT&CK framework predominantly rely on supervised learning. Branescu et al. [14] modelled this as a multi-label classification task using CVE descriptions for ATT&CK tactics mappings. Ampel et al. [15] employed self-distillation to capture long-term textual dependencies. Vulcan Cyber [17] proposed enriching input features with CWE and CVSS Version 3.x data. BERT-based models, such as CVE2ATT&CK [11], showed promise using only CVE descriptions for mapping 31 of 92 ATT&CK techniques. Mendsaikhan et al. [16] expanded coverage to 52 techniques using textual features from CVE descriptions, though performance declined with label expansion due to limited training data. Adam et al. [18] introduced a two-step mapping via CWEs, but their method is constrained by incomplete CWE annotations in CVEs and the lack of comprehensive CWE-to-ATT&CK mappings.

B. Unsupervised Mapping

The SMET framework [13] employs semantic role labelling to rank ATT&CK techniques based on CVE descriptions, without requiring labelled data. Kuppa et al. [12] demonstrated an unsupervised approach that maps CVEs to 37 ATT&CK techniques by extracting relevant phrases from both threat reports and ATT&CK descriptions. However, they observed that many CVE entries contain minimal textual content, resulting in incomplete or failed mappings.

Since SMET does not leverage structured NVD attributes (e.g., CVSS, CWE, CPE) or EPSS probability scores, it struggles with sparse or ambiguous descriptions. Furthermore, as a fully unsupervised pipeline with no learnable parameters, SMET lacks adaptability to evolving CVE patterns or domain-specific requirements. Its logistic regression classifier and embedding model are trained on ATT&CK descriptions, not CVE text, rendering the system insensitive to changes in vulnerability language, emerging exploit types, or shifts in reporting conventions.

The MITRE ATT&CK Enterprise framework comprises 14 tactics, 211 techniques, and 468 sub-techniques, making comprehensive CVE-to-technique mapping a complex task. Due to limited annotated data for many techniques, prior work has focused on a subset of them. In this study, we shift the focus to mapping CVEs to ATT&CK tactics, emphasizing higher-level adversary objectives, such as *Reconnaissance, Initial Access, Collection*, and *Impact*, all of which have been historically difficult for state-of-the-art (SOTA) methods [14].

Mapping at the tactic level offers practical benefits for vulnerability prioritization, attacker path modeling, and risk propagation [26]. To address the challenges of sparse textual descriptions and limited adaptability, we enhance both supervised and unsupervised approaches by incorporating structured NVD data (CVSS, CWE, CAPEC, CPE) and EPSS scores, going beyond CVE descriptions alone.

While our long-term goal remains mapping to techniques, we argue that tactic-level mapping can be substantially improved with richer input features. This work lays a scalable foundation for future expansion to technique-level mappings with broader dataset coverage.

III. Methodology

The proposed methodology for automatically mapping CVEs to MITRE ATT&CK tactics comprises four main phases: dataset collection, dataset processing, mapping using both unsupervised and supervised approaches, and performance evaluation. The overall architecture is shown in Figure 1, with each phase described in detail below.

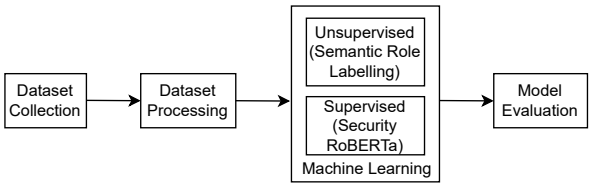

Figure 1. An overview of the proposed framework for automated mapping of CVEs to MITRE ATT&CK Tactics

A. Dataset Collection

The phase begins with the collection of an initial dataset to support both unsupervised and supervised approaches. We use datasets from SMET [13] and Branescu et al.[14], representing state-of-the-art methods in each category. These datasets include CVE IDs and their corresponding descriptions as primary features. To enrich this initial dataset, we incorporated five additional sources from the NVD and related repositories:

- CVE Descriptions (baseline): Textual descriptions of vulnerabilities from CVE entries. For instance, the CVE description for CVE-2025-49163 is "*Arris VIP1113 devices through 2025-05-30 with KreaTV SDK allow booting an arbitrary image via a crafted /usr/bin/gunzip file.*"
- Common Weakness Enumeration (CWE) [27]: Standardized identifiers for software weaknesses (e.g., CWE-79 Improper Neutralization of Input During Web Page Generation ('Cross-site Scripting')).
- Common Vulnerability Scoring System (CVSS) [28]: Quantitative scores (0–10) reflecting exploitability and impact.
- Common Platform Enumeration (CPE) [29]: Machine-readable identifiers for vulnerable software/hardware (e.g., `cpe:2.3:a:microsoft:edge:`)
- Exploit Prediction Scoring System (EPSS)[30]: Probabilities (0–1) estimating the likelihood of a CVE being exploited in the wild.
- Common Attack Pattern Enumeration and Classification (CAPEC) [31]: Titles describing adversary tactics (e.g., CAPEC-209: XSS Using MIME Type Mismatch).

NVD data was retrieved using the official API [32], EPSS scores via the EPSS API [33], and CAPEC titles were obtained through web scraping and HTML parsing of the CAPEC website.

Following enrichment, the unsupervised dataset retains its original size but is enhanced with additional features, including CWE, CPE, and CVSS data. The supervised dataset initially comprises 9,986 CVEs from prior work [14], but is reduced to 7,328 entries after filtering out CVEs lacking sufficient NVD attributes (CWE, CPE, CVSS). Features with missing values across any dataset entries are removed to prevent negative impacts on machine learning performance. This preprocessing step of filtering out rows with null values to improve system effectiveness is consistent with established machine-learning approaches for automated CVE-to-MITRE ATT&CK tactic mapping [34].

For supervised learning, we adopt an 80/20 train-test split as recommended by [14], using 80% of the 7,328 CVEs for training and 20% for testing. The *extended* dataset is enriched with EPSS scores, processed CVSS (v2/v3), CWE, CPE, and later CAPEC features. A comparative summary between the initial dataset [14] and the enriched version used in this work is provided in Table I.

TABLE I. COMPARISON BETWEEN THE INITIAL DATASET INTRODUCED BY BRANESCU ET AL.[14] AND OUR FULLY ENRICHED DATASET

ATT&CK Tactic Class	CVE Record Count	
	Initial Dataset[14]	Our Dataset
Reconnaissance	170	141
Resource Development	170	117
Initial Access	722	573
Execution	2642	1183
Persistence	3016	1591
Privilege Escalation	3218	1731
Defense Evasion	7552	5354
Credential Access	614	534
Discovery	2369	1959
Lateral Movement	1932	620
Collection	663	576
Command & Control	427	382
Exfiltration	171	126
Impact	349	286
Total	**9,986 CVEs**	**7,328 CVEs**

The unsupervised dataset comprises 827 CVEs distributed over 120 ATT&CK techniques, selected from publicly available CVE-to-ATT&CK mappings provided by the Center for Threat-Informed Defense [24], allowing for direct evaluation against verified mappings. This differs from the dataset used in SMET [13], which lacked consistent NVD feature coverage. Accordingly, the SMET baseline was re-evaluated on our 827-entry *description-only* dataset, compared to the original 303 entries used in SMET distributed over 41 techniques from the ATT&CK matrix.

Two primary datasets are created for the unsupervised and supervised mappings: (1) the *Description-Only* dataset, containing CVE IDs and textual descriptions, and (2) the *Extended* dataset, which builds upon the former by incorporating pre-processed CWE, CVSS, CPE, EPSS scores, and optionally CAPEC data.

B. Dataset Processing

A key design decision in this research was to enrich the original CVE dataset with additional structured fields, CWE, CVSS, CPE, and optionally EPSS and CAPEC, to improve the accuracy of CVE-to-ATT&CK tactic mappings. This enrichment supports both the unsupervised (SMET) and supervised (SecRoBERTa) approaches by enhancing the semantic and contextual representation of each CVE. To ensure consistency across data types, all extended features (except EPSS scores, which are numeric) were pre-processed into natural language format. This was necessary due to the inconsistent quality of CVE descriptions and the structured, non-linguistic format of most added fields.

1) CVSS Pre-Processing: CVSS Version 2.0 and CVSS Version 3.x vector strings were transformed into natural language using mappings from the NVD CVSS calculators [35][36]. For example, the CVSS v3.1 vector `CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:U/C:H/I:H/A:H` of CVE-2023-23333 is pre-processed into:

- *"The CVE is Exploited by the Network Attack Vector. The CVE has Low Attack Complexity. The CVE Requires No Privileges. The CVE Does Not Require User Interaction. The CVE scope is Unchanged. The CVE has a High Confidentiality Impact. The CVE has High Integrity Impact. The CVE has High Availability Impact."*

2) CWE Pre-Processing: CWEs were converted into natural language by combining their titles and descriptions. For instance, CWE-78 is rendered as:

- *"The CVE is affected by Improper Neutralization of Special Elements used in an OS Command ('OS Command Injection'): The product constructs all or part of an OS command using externally-influenced input from an upstream component, but it does not neutralize or incorrectly neutralizes special elements that could modify the intended OS command when it is sent to a downstream component."*

For multiple CWEs associated with a single CVE ID, the processed CWE strings will be concatenated to create a longer sentence.

3) CAPEC Pre-processing: CAPEC titles were processed similarly to CWEs. However, due to limited API support and outdated data (last reviewed in 2023), CAPEC enrichment was used selectively. A pilot study indicated that the inclusion of CAPEC features did not improve mapping accuracy.

4) CPE Pre-Processing: CPE strings were parsed to extract three key attributes: component type (application, Operating System, or hardware), vendor, and product. These were reformatted into natural language. For example: `cpe:2.3:o:contec:solarview_compact_firmware:*:*:*:*:*:*:*:*:*` was converted to: *The "CVE affects Contec Solarview_compact_firmware Operating System."*.

To reduce noise from highly variable product names, a standardization step was applied. Generic terms, such as "Product", were substituted when the product name was not essential. In contrast, critical operating system identifiers (e.g., Windows, Linux, Mac_os_x and Linux_kernel) were preserved. For instance: `cpe:2.3:a:microsoft:365_apps:-:*:*:*:enterprise:*:*:*` was converted to: *"The CVE affects Microsoft Product Application."*. This generalization improves model robustness by minimizing irrelevant variance while preserving key distinctions necessary for accurate CVE-to-ATT&CK tactic mapping.

C. Machine Learning

This research addresses the challenge of automatically mapping CVEs to the MITRE ATT&CK framework using both unsupervised and supervised machine learning techniques.

For the unsupervised approach, we employ the SMET framework [13], a state-of-the-art method that does not require labelled data. SMET extracts semantically meaningful attack vectors from CVE textual descriptions by leveraging semantic role labelling and other semantic similarity techniques. We extend the original SMET implementation, designed for *description-only* inputs, to incorporate structured features from the NVD, including CVE ID, CWE, CVSS, and CPE. We hypothesise that even in the absence of labelled data, incorporating this extended feature set enhances the quality of semantic mappings.

For the supervised approach, we utilize SecRoBERTa [37], a transformer-based model derived from RoBERTa [38], which is an optimized version of BERT (Bidirectional Encoder Representations from Transformers), and has been fine-tuned on cybersecurity-specific corpora. Prior work has shown that SecRoBERTa achieves state-of-the-art performance in mapping CVEs to ATT&CK techniques [14]. Castano et al. [39] trained five BERT-based models and found SecRoBERTa to be the most effective at linking CTI sources via external references, resulting in more complete datasets and improved threat intelligence.. We further fine-tune a pre-trained ATT&CK-BERT model from *Hugging Face* [37] using our extended dataset, which includes NVD features, EPSS probability scores, and CAPEC identifiers. Tokenization and model management are performed using the *Hugging Face* Transformers library [40]. While we maintain the default settings for batch size and number of epochs, we adjust the learning rate to `3.884755049077609e-05` and the dropout rate to `0.4864913766068174` to optimize model performance.

The dual-method study aims to investigate whether both unsupervised and supervised models can benefit from enhanced CVE representations, which could improve automated mappings to adversarial tactics in ATT&CK.

D. Model Evaluation

We evaluated the machine learning models using *accuracy, validation loss*, and both *macro* and *weighted F1 scores*. *Validation loss* serves as an indicator of generalization performance, with lower values suggesting reduced overfitting or underfitting. *Accuracy* reflects the overall proportion of correctly predicted tactic labels. The *macro F1 score*, as the unweighted average of per-tactic F1 scores, emphasizes performance on less frequent classes. The *weighted F1 score*, our primary metric, accounts for class imbalance by weighting each tactic's F1 score by its frequency.

IV. EVALUATION

To support a comprehensive evaluation of our system, we define specific hypotheses for validation:

- **H1**: An *extended* dataset improves overall mapping accuracy compared to the commonly used *description-only* datasets [11][13][14][16].
- **H2**: Tactics that are typically harder to classify, such as *Reconnaissance, Initial Access, Collection*, and *Impact*,

as identified by Branescu et al. [14], will show improvements in their F1-scores.
- **H3**: Hyperparameter tuning leads to additional gains in mapping accuracy.

A. Unsupervised Mapping Validation

Given that SMET is an unsupervised methodology that ranks mappings based on semantic similarity. In the proposed solution, rankings greater than 0.1 are considered potentially correct mappings. When an entry is labelled as *Completely Accurate*, it indicates a 1:1 match with the testing data provided for a CVE. If the entry is designated as *Semi-Accurate*, it means that while the accepted mappings included correct ones, they also incorporated some incorrect mappings that exceeded the threshold. Conversely, if the entry is marked as *Inaccurate*, it signifies that no correct mappings were obtained that met the threshold (>0.1). We compared the SMET results from a *Description Only* dataset with an *extended* dataset, verifying an increase in mapping accuracy. As shown in Table II, the enriched unsupervised dataset with CVSS vectors, CWE and CPE summaries outperformed the baseline *description only* in every metric.

TABLE II. COMPARISON OF UNSUPERVISED MAPPING ACCURACY ON 828 CVES, *Description Only* VS. *Enriched* DATASET

	Description Only	**Enriched (+CVE, CVSS & CPE)**	**Impact (%)**
Completely Accurate	6	10	+66.7
Semi Accurate	88	125	+42.0
Inaccurate	733	692	-5.6

Despite notable improvements in mapping accuracy, over 80% of CVEs remain incorrectly mapped without supervised learning, demonstrating that feature enrichment alone is insufficient. While enriched features capture semantically meaningful attack vectors, they do not match the performance of supervised models. Results show that mapping CVEs to ATT&CK Techniques suffers from low accuracy due to limited labelled data.

For example, SMET, a state-of-the-art unsupervised method, uses text similarity between CVEs and ATT&CK technique descriptions, enabling semantic mapping of 303 CVEs to 41 techniques. In contrast, our improved unsupervised dataset includes 828 CVEs mapped to 120 techniques, with both datasets averaging 7 CVEs per technique. Meanwhile, the leading supervised dataset includes 9,985 CVEs across 14 tactics, with each tactic supported by a minimum of 170 samples and an average of 713 entries (see Table I). This data imbalance leads to better performance when mapping to ATT&CK Tactics rather than Techniques.

Given the limited performance of unsupervised methods, this research adopts a supervised approach. Nonetheless, the unsupervised results confirm that *enriched* datasets are more effective than *description-only* inputs for offensive technique mapping.

B. Supervised Mapping Validation

This section presents the results of the supervised learning experiments. First, we analyze performance across dataset variants to assess the impact of added features, including comparisons with and without CAPEC. Second, we report per-tactic F1 scores for MITRE ATT&CK tactics. Finally, we benchmark our approach against the state-of-the-art method by Branescu et al. [14].

Table III presents the overall performance across the supervised dataset variants, enabling a detailed comparison of feature-specific contributions. Incorporating the EPSS feature alone consistently improves all four performance metrics: validation loss, accuracy, macro F1 score, and weighted F1 score, relative to the *description-only*baseline. Similar improvements are observed when CWE, CPE, and CVSS features are added, each contributing to increased model performance. In contrast, the inclusion of the CAPEC feature results in a decline across all four metrics, with the CAPEC Title-only extension causing a particularly notable degradation. As a result, CAPEC was excluded from the fully extended feature set. All other feature combinations outperform the description-only baseline, thereby supporting Hypothesis **H1**. Additionally, hyperparameter tuning yields a consistent performance boost over the enriched but untuned models, with gains of approximately 2% to 3% in mapping accuracy and macro F1 score, thereby supporting Hypothesis **H3**.

TABLE III. OVERALL PERFORMANCE ACROSS SUPERVISED DATASET VARIANTS.

Supervised Dataset Variant	**Validation Loss**	**Accuracy**	**Macro F1 Score**	**Weighted F1 Score**
Description Only	0.0747	0.8286	0.7948	0.9232
Description + EPSS	0.0729	0.8335	0.8138	0.9277
Description + CWE	0.0724	0.8407	0.7979	0.9248
Description + CVSS	0.0815	0.8229	0.8024	0.9163
Description + CPE	0.0746	0.8286	0.8050	0.9244
Description + CAPEC	0.0870	0.8179	0.7119	0.9011
Fully Extended (Description + EPSS + CVSS + CPE)	0.0743	0.8383	0.8144	0.9245
Fully Extended + Tuned	**0.0658**	**0.8538**	**0.8401**	**0.9347**

Table IV presents F1 scores for the *description-only* baseline, the fully enriched dataset (with and without tuning), and the Branescu et al. [14] SecRoBERTa model. The most notable improvements were observed for hard-to-predict tactics: *Initial Access* improved from 65.27% to 67.44%, *Collection* from 79.44% to 84.34%, *Impact* from 67.57% to 72.00%, and *Reconnaissance* from 37.33% to 46.15%, confirming Hypothesis **H2**. Medium-difficulty tactics, such as *Credential Access* and *Command & Control*, saw moderate gains of 1%–7%. Well-predicted tactics, such as *Defense Evasion, Discovery, Privilege Escalation, Persistence, Lateral Movement*, and *Execution* began above 90% and saw only marginal improvements (1%–2%), with fine-tuning contributing an additional 0.5%–1%. The final model achieved a 93.5% weighted F1 score.

TABLE IV. PLOTS PER-CLASS F1 SCORES FOR THE SUPERVISED DATASET VARIANT AND COMPARISON BETWEEN THE SECROBERTA PER-CLASS ON DESCRIPTION ONLY REPORTED IN [14]

Tactics	Description only (Benchmark)	Full Extended Dataset (+EPSS+CWE +CVSS+CPE)	+ Optuna Fine-Tuning	SOTA [14]
Reconnaissance	37.33%	36.73%	**46.15%**	53.84%
Resource Development	51.47%	65.81%	**65.79%**	79.13%
Initial Access	65.27%	61.52%	**67.44%**	37.18%
Execution	89.56%	89.25%	**89.95%**	74.43%
Persistence	94.42%	94.52%	**94.87%**	80.78%
Privilege Escalation	94.66%	94.90%	**95.11%**	80.46%
Defense Evasion	98.67%	98.00%	**98.41%**	91.96%
Credential Access	84.82%	89.39%	**91.81%**	67.27%
Discovery	97.24%	97.29%	**97.92%**	81.55%
Lateral Movement	92.87%	94.59%	**94.97%**	81.37%
Collection	79.44%	81.82%	**84.34%**	51.47%
Command & Control	95.21%	95.81%	**96.43%**	61.79%
Exfiltration	64.23%	74.83%	**81.01%**	88.88%
Impact	67.57%	65.73%	**72.00%**	31.11%

Compared to *Branescu et al.'s*[14] model trained on 9,986 CVEs (weighted F1: 78.88%), our approach, applied to 7,786 CVEs, achieves 93.45%. This performance gain is attributed to structured feature enrichment from NVD (CWE, CVSS, CPE) and the addition of EPSS, coupled with effective hyperparameter tuning. The results demonstrate that enriched features and tuning significantly enhance CVE-to-ATT&CK tactic mapping accuracy, especially for previously underperforming tactics.

To ensure a fair comparison with Branescu et al. [14], we grouped predicted MITRE ATT&CK tactics into three difficulty levels based on F1 scores: *hard* (<60%), *medium* (60%–80%), and *easy* (>80%). As shown in Table V, Branescu et al. [14] identified *four* hard, *four* medium, and *six* easy tactics. Using our enriched dataset and improved processing pipeline, our method reduced the number of hard tactics to *one*, with *three* medium and *ten* easy-to-predict tactics. Notably, *three* tactics previously classified as hard in [14] were reclassified as two medium and one easy, while three of the four medium tactics shifted to the easy category in our mapping methodology. The *six* easy tactics remained unchanged. These results support Hypothesis **H2**, demonstrating that dataset enrichment and enhanced modelling reduce classification difficulty for previously challenging tactics.

V. DISCUSSION

According to the CVE Key Details Phrasing Guidelines by MITRE [41], a comprehensive CVE description should articulate several key aspects, including the vulnerability type or root cause, attack vector, impact, attacker type, component identification, affected product(s) and version(s), and product vendor(s). However, despite the importance of these details, many CVE descriptions and their associated references suffer from inconsistencies, a lack of structure, or insufficient information [17][22][42], which poses a significant challenge for downstream tasks such as mapping vulnerabilities to offensive tactics and techniques, particularly within the MITRE ATT&CK framework.

TABLE V. COMPARISON BETWEEN THE PREDICTED ATT&CK TACTICS BY BRANESCU ET AL. [14] AND OUR PROPOSED APPROACH

Difficulty Level	ATT& CK Tactics in Branescu et al. [14]	ATT&CK Tactics in our proposed approach
Hard	Reconnaissance, Collection, Initial Access, Impact	Reconnaissance
Medium	Resource Development, Credential Access, Execution, Command & Control	Resource Development, Initial Access, Impact
Easy	Privilege Escalation, Discovery, Persistence, Exfiltration, Defense Evasion, Lateral Movement	Privilege Escalation, Discovery, Persistence, Exfiltration, Defense Evasion, Lateral Movement, Execution, Credential Access, Collection, Command & Control

This research demonstrates that supplementing CVE descriptions with structured data, such as software weaknesses (CWE), platform identifiers (CPE), exploit prediction scores (EPSS), attack patterns (CAPEC), and vulnerability scoring metrics (CVSS vector strings), can significantly enhance the completeness and utility of CVE records. By enriching the original textual descriptions with these standardized attributes, the proposed approach improves the effectiveness of both supervised and unsupervised models for mapping CVEs to ATT&CK techniques.

Vulnerability descriptions serve as a critical foundation in the identification and communication of security weaknesses in software, systems, and hardware. High-quality descriptions not only support threat assessment and mitigation but are also essential for enabling automated systems to aid in vulnerability prioritization and response. This study contributes to the growing body of work aimed at automating the mapping of CVEs to adversary behavior models, thereby advancing vulnerability analysis and threat-informed defense.

VI. LIMITATION

A key limitation of this research is the dynamic nature of the MITRE ATT&CK framework and the fast-paced evolution of the cyber threat landscape, which may lead to misalignment between the framework and the most current adversary tactics, techniques, and procedures (TTPs). Despite the enhanced approach, the *Reconnaissance* tactic remains difficult to predict. While this could be attributed to its underrepresentation in the dataset, this explanation is insufficient, as *Exfiltration*, similarly sized (see Table I), achieves significantly better performance and falls into the more predictable category. Another limitation stems from the exclusion of CVEs that lack extended fields, which introduces bias toward well-documented vulnerabilities and excludes zero-day threats. These cannot be included until evaluated by the NVD and assigned relevant attributes, such as CPE. To address this, future work may explore partial feature selection to enable broader coverage until a fully extended dataset becomes available. Furthermore, LLMs can enhance textual vulnerability descriptions by utilising historical data, enabling the system to comprehend new

vulnerabilities without requiring retraining of the LLM. Additionally, the CAPEC feature was poorly represented due to incomplete web scraping, which extracted only CAPEC Titles. This limited the utility of the CAPEC data and negatively impacted performance.

VII. Conclusion and Future Work

This research demonstrates that augmenting CVE descriptions with extended features, including EPSS, CWE, CVSS, CPE, and CAPEC, significantly improves mapping accuracy to MITRE ATT&CK tactics. In unsupervised experiments, enrichment increased the number of correct mappings and reduced misclassifications. In supervised experiments, the proposed SecRoBERTa-based model outperformed the current state-of-the-art models. Accurate CVE-to-ATT&CK mapping enables Security Operations Centers (SOCs) to prioritize and mitigate unpatched vulnerabilities more effectively. As CVE descriptions often lack consistency and technical detail, enriching them with well-processed structured features leads to more reliable mapping outcomes.

Future work will focus on developing a CAPEC API to streamline the integration of CAPEC features. Building on the high mapping accuracy to ATT&CK tactics, the next phase will extend this approach to ATT&CK techniques, using a supervised methodology while constraining predictions to the relevant parent tactic. Beyond enriching textual vulnerability descriptions with structured information such as EPSS, CWE, CAPEC, CVSS vector strings, and CPE configurations, future work will explore methods for detecting and augmenting missing key aspects of CVE entries. This can be approached through machine learning techniques that predict the labels of absent attributes based on known vulnerability characteristics or through software feature inference. Such enhancements have the potential to improve downstream applications that rely on CVE data, including vulnerability severity prediction, automated alignment with adversary tactics and techniques, the development of exploitation prediction models, and automated vulnerability classification.

Acknowledgments

This work was supported by the NICYBER2025 programme funded by Innovate UK.

References

[1] The MITRE Corporation, "Metrics," Retrieved: July 2025, 2025, [Online]. Available: https://www.cve.org/about/Metrics.

[2] W. Baker, "The pithy p2p: 5 years of vulnerability remediation & exploitation research," Retrieved: July 2025, 2025, [Online]. Available: https://www.cyentia.com/pithy-p2p/.

[3] P. Garrity, "State of exploitation - a peek into the last decade of vulnerability exploitation," Retrieved: July 2025, 2024, [Online]. Available: https://vulncheck.com/blog/state-of-exploitation-a-decade.

[4] Cyentia Institute, "A visual exploration of exploitation in the wild: The inaugural study of epss data and performance," Retrieved: July 2025, 2024, [Online]. Available: https://www.cyentia.com/wp-content/uploads/2024/07/EPSS-Exploration-Of-Exploits.pdf.

[5] Google Cloud Security, "Mandiant m-trends 2025 report," Retrieved: September 2025, 2025, [Online]. Available: https://services.google.com/fh/files/misc/m-trends-2025-en.pdf.

[6] J. Baker, "2023 r&d roadmap to advance threat-informed defense," Retrieved: July 2025, 2023, [Online]. Available: https://medium.com/mitre-engenuity/2023-r-d-roadmap-to-advance-threat-informed-defense-cf726d30e583.

[7] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, "Mitre att&ck: Design and philosophy," in *Technical report*, The MITRE Corporation, 2018.

[8] The MITRE Corporation., "Att&ck®," Retrieved: August 2025, 2025, [Online]. Available: https://attack.mitre.org/.

[9] J. Baker, "Cve + mitre att&ck® to understand vulnerability impact," Retrieved: July 2025, 2021, [Online]. Available: https://medium.com/mitre-engenuity/cve-mitre-att-ck-to-understand-vulnerability-impact-c40165111bf7.

[10] D.-Y. Kim, S.-S. Yoon, and I.-C. Euom, "V2tsa: Analysis of vulnerability to attack techniques using a semantic approach," *IEEE Access*, 2024.

[11] O. Grigorescu, A. Nica, M. Dascalu, and R. Rughinis, "Cve2att&ck: Bert-based mapping of cves to mitre att&ck techniques," *Algorithms*, vol. 15, no. 9, p. 314, 2022.

[12] A. Kuppa, L. Aouad, and N.-A. Le-Khac, "Linking cve's to mitre att&ck techniques," in *Proceedings of the 16th International Conference on Availability, Reliability and Security*, 2021, pp. 1–12.

[13] B. Abdeen, E. Al-Shaer, A. Singhal, L. Khan, and K. Hamlen, "Smet: Semantic mapping of cve to att&ck and its application to cybersecurity," in *IFIP annual conference on data and applications security and privacy*, Springer, 2023, pp. 243–260.

[14] I. Branescu, O. Grigorescu, and M. Dascalu, "Automated mapping of common vulnerabilities and exposures to mitre att&ck tactics," *Information*, vol. 15, no. 4, p. 214, 2024.

[15] B. Ampel, S. Samtani, S. Ullman, and H. Chen, "Linking common vulnerabilities and exposures to the mitre att&ck framework: A self-distillation approach," *arXiv preprint arXiv:2108.01696*, 2021.

[16] O. Mendsaikhan, H. Hasegawa, Y. Yamaguchi, and H. Shimada, "Automatic mapping of vulnerability information to adversary techniques," in *The Fourteenth International Conference on Emerging Security Information, Systems and Technologies SECUREWARE2020*, 2020, pp. 53–59.

[17] Vulcan Cyber, "Cve to t&ts: Using cve attributes for mitre att&ck mapping," Retrieved: July 2025, 2023, [Online]. Available: https://web.archive.org/web/20240429155732/https://l.vulcan.io/hubfs/Ebooks-and-White-Papers/Vulcan-Cyber-Mapping-CVEs-to-MITRE.pdf.

[18] C. Adam, M. F. Bulut, D. Sow, S. Ocepek, C. Bedell, and L. Ngweta, "Attack techniques and threat identification for vulnerabilities," *arXiv preprint arXiv:2206.11171*, 2022.

[19] Q. Li, W. Tang, X. Chen, and H. Ren, "Vuldifffinder: Discovering inconsistencies in unstructured vulnerability information," *Computers & Security*, p. 104 447, 2025.

[20] Y. Dong, W. Guo, Y. Chen, X. Xing, Y. Zhang, and G. Wang, "Towards the detection of inconsistencies in public security vulnerability reports," in *28th USENIX security symposium (USENIX Security 19)*, 2019, pp. 869–885.

[21] Y. Chen, A. E. Santosa, A. Sharma, and D. Lo, "Automated identification of libraries from vulnerability data," in *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Software Engineering in Practice*, 2020, pp. 90–99.

[22] H. Guo, S. Chen, Z. Xing, X. Li, Y. Bai, and J. Sun, "Detecting and augmenting missing key aspects in vulnerability descriptions," *ACM Transactions on Software Engineering and*

Methodology (TOSEM), vol. 31, no. 3, pp. 1–27, 2022. DOI: 10.1145/3498537.

[23] T. Chen *et al.*, "Vullibgen: Identifying vulnerable third-party libraries via generative pre-trained model," *CoRR*, 2023.

[24] Center for Threat-Informed Defense, "Mapping mitre att&ck® to cves for impact," Retrieved: July 2025, 2021, [Online]. Available: https://github.com/center-for-threat-informed-defense/attack_to_cve.

[25] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2019, pp. 2623–2631.

[26] P. Bhosale, W. Kastner, and T. Sauter, "Mapping ics vulnerabilities: Prioritization and risk propagation analysis with mitre att&ck framework and bayesian belief networks," in *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*, IEEE, 2024, pp. 1–8.

[27] The MITRE Corporation (MITRE), "Common weakness enumeration," Retrieved: September 2025, 2025, [Online]. Available: https://cwe.mitre.org/.

[28] Forum of Incident Response and Security Teams, "Common vulnerability scoring system," Retrieved: July 2025, 2025, [Online]. Available: https://www.first.org/cvss/.

[29] NIST, "Official common platform enumeration (cpe) dictionary," Retrieved: July 2025, 2025, [Online]. Available: https://nvd.nist.gov/products/cpe.

[30] J. Jacobs, S. Romanosky, B. Edwards, I. Adjerid, and M. Roytman, "Exploit prediction scoring system (epss)," *Digital Threats: Research and Practice*, vol. 2, no. 3, pp. 1–17, 2021.

[31] The MITRE Corporation, "Common attack pattern enumeration and classification," Retrieved: July 2025, 2023, [Online]. Available: https://capec.mitre.org/.

[32] NIST, "Vulnerabilities: Cve api," Retrieved: July 2025, 2025, [Online]. Available: https://nvd.nist.gov/developers/vulnerabilities.

[33] Forum of Incident Response and Security Teams, "Epss api," Retrieved: July 2025, 2025, [Online]. Available: https://www.first.org/epss/api.

[34] Y. Lakhdhar and S. Rekhis, "Machine learning based approach for the automated mapping of discovered vulnerabilities to adversial tactics," in *2021 IEEE Security and Privacy Workshops (SPW)*, IEEE, 2021, pp. 309–317.

[35] NIST, "Common vulnerability scoring system calculator - cvss version 2.0," Retrieved: July 2025, 2025, [Online]. Available: https://nvd.nist.gov/vuln-metrics/cvss/v2-calculator.

[36] NIST, "Common vulnerability scoring system calculator - cvss version 3.0, cvss version 3.1," Retrieved: July 2025, 2025, [Online]. Available: https://nvd.nist.gov/vuln-metrics/cvss/v3-calculator.

[37] Kun jackaduma, "Secroberta," Retrieved: July 2025, 2023, [Online]. Available: https://huggingface.co/jackaduma/SecRoBERTa.

[38] Y. Liu *et al.*, "Roberta: A robustly optimized bert pretraining approach," *arXiv preprint arXiv:1907.11692*, 2019.

[39] F. Castaño, A. Gil-Lerchundi, R. Orduna-Urrutia, E. F. Fernandez, and R. Alaiz-Rodríguez, "Wave-27k: Bringing together cti sources to enhance threat intelligence models," in *Proceedings of the First International Conference on Natural Language Processing and Artificial Intelligence for Cyber Security*, 2024, pp. 119–126.

[40] T. Wolf *et al.*, "Transformers: State-of-the-art natural language processing," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Online: Association for Computational Linguistics, 2020, pp. 38–45. DOI: 10.18653/v1/2020.emnlp-demos.6.

[41] J. Evans, "Mitre key details phrasing," Retrieved: September 2025, 2020, [Online]. Available: https://cveproject.github.io/docs/content/key-details-phrasing.pdf.

[42] C. Madden, "Vulnerability description quality checks and data analysis," Retrieved: September 2025, 2025, [Online]. Available: https://github.com/CyberSecAI/VulnerabilityDescriptionQualityChecker.