AI-Powered Penetration Testing: Navigating the Phases of Cybersecurity

Kiera Conway   
Dakota State University  
Seattle, USA  
Kiera.Conway@trojans.dsu.edu

*Abstract*— In response to the rising complexity and frequency of cyber threats, Penetration Testing (PT) has become imperative for safeguarding digital assets. However, traditional PT methods face limitations in coping with substantial workloads, intricate networks, and the severe shortage of skilled penetration testers (pentesters); these limitations have prompted the exploration of Artificial Intelligence (AI) integration as a potential solution. This report explores this avenue by examining AI technologies, ranging from foundational Machine Learning (ML) models to more complex neural network-driven approaches, and assesses their potential to enhance the efficiency and effectiveness of PT. By synthesizing insights from seven research papers, across various phases of the PT process, this report investigates the potential of combining the structured methodologies of modern PT methods with AI's capacity for operational efficiency and vulnerability prioritization. Then, in response to the severe limitation of relevant research, which currently impedes the adaptability and scalability of transitioning towards an intelligently automated PT practice, these insights are consolidated into a short seminar course titled "Harnessing Artificial Intelligence for Penetration Testing." The primary objective of this course extends beyond providing participants with a comprehensive understanding of PT, but also aims to spread awareness of the vast potential inherent to the integration of AI at each phase of the PT process. By fostering critical thinking and practical application, this course aims to empower students to intelligently navigate the complex realm of cybersecurity, combat adversaries, and actively contribute to a more secure digital environment.

Keywords— Artificial Intelligence, Penetration Testing, Machine Learning, Reinforcement Learning, Deep Learning, Cybersecurity, Ethical Hacking, Vulnerability Assessment, Phases

# Introduction

In an age defined by the relentless increase of technology, the growing digital landscape has become both a playground for innovation and a battleground for cyber threats. As organizations increasingly rely on technology to operate, communicate, and store critical data, safeguarding these assets against potential adversaries becomes paramount. This realization has given rise to Penetration Testing (PT) as a vital and proactive strategy that allows organizations to simulate cyberattacks on their systems to discover and eliminate dangerous vulnerabilities. PT, often referred to as ethical hacking, is the “offensive approach” of probing and assessing computer systems, networks, and applications “to actively identify vulnerabilities and then exploit them in the same way as a genuine attacker [1].”

As modern digital tools become increasingly intelligent and interconnected, the importance of adhering to a structured process for the offensive systematic evaluation of a system's security is paramount. While the specific phases of PT may exhibit variations in terminology depending on the source, their fundamental organization remains consistent: preparation, implementation, and analysis [2]. However, in order to underscore the significance of AI methodically, this report and course are explicitly developed around the PT phases as outlined by Splunk [3].

* Phase 1 – Gathering Information: The initial phase involves the meticulous compilation of data about the target system and potential vulnerabilities. This step forms the foundation for the subsequent phases and provides a comprehensive understanding of the target's landscape.
* Phase 2 – Scanning: Following the information gathering, this phase focuses on verifying potential vulnerabilities as exploitable. Through systematic scanning of networks, vulnerabilities, or web applications, the penetration tester identifies weak points that may potentially be exploited in further stages.
* Phase 3 – Exploitation: Building upon the identified vulnerabilities, the exploitation phase strategically leverages these weaknesses to gain access to the target system. This step is crucial for simulating real-world cyberattacks and understanding potential points of compromise.
* Phases 4, 5, 6 – Post Exploitation: These three phases, collectively known as Post Exploitation, play a pivotal role in the aftermath of successful penetration testing. In this category, the pentester navigates through: Maintaining the Connection, Covering Tracks, and Reporting. Immediately after successful exploitation of vulnerabilities and gaining unauthorized access to a target system, maintaining a connection allows the penetration tester to establishing a persistent connection, and ensure ongoing access to the compromised system. Once access is secured, the pentester will attempt to minimize the risk of detection by erasing or obfuscating any evidence of their presence. Lastly, and most importantly, the pentester must document and communicate vulnerabilities discovered, methods used for exploitation, and remediation recommendations. To facilitate discussions in both the report and course, these post-exploit phases are collectively referred to as such, thus enabling a more centralized focus on the initial three phases.

Given the complex nature and substantial workload inherent to this PT process, it necessitates significant expertise from skilled professionals who can “provide organizations with actionable information about their security posture, enabling them to identify and prioritize areas of risk and improve their overall security [4] [5].” As such, pentesters involved in this field must maintain a continuous regimen of training and practical skill development to remain at the forefront. Unfortunately, this excessive pressure, combined with the recent shortage of pentesters, has left many in the field feeling overburdened and overtired [1].

In response to this growing complexity of networks, there is a fundamental need to embrace intelligent automation, by integrating Artificial Intelligence (AI), Machine Learning (ML), and Reinforcement Learning (RL). This imperative arises from the recognition that traditional and manual PT methods, while valuable, are time-consuming, resource-intensive, and may struggle to keep pace with the dynamic nature of cyber threats. Consequently, as networks and threats become more sophisticated, the demands for a more nuanced and efficient approach to identifying vulnerabilities becomes pertinent. However, a shift towards intelligent automation could not only significantly reduce the time and resources required for comprehensive testing, but also serve as a remedy for many of the prevalent and "recurrent human errors" in manual PT that stem from factors such as "tiredness, omission, and pressure" [1]. This transformation in the field, as highlighted by [5], signifies the use of "advanced algorithms, machine learning, and AI to scan systems for vulnerabilities," offering significant improvements across multiple facets of PT. By navigating this dynamic landscape intelligently, this evolution serves not just as a response to current demands, but as a proactive strategy to better enable organizations to bolster their defenses against emerging threats.

However, addressing the current landscape of research in the integration of AI in PT reveals several notable limitations. Firstly, there exists a significant gap in research and testing within this domain, thereby signifying an unexplored terrain ripe for investigation. While research has begun exploring this novel integration [1] [5], many of these approaches are limited by their reliance on manual human intervention for vulnerability identification and exploitation. This reliance hinders the full realization of AI's autonomous potential in the PT process, but autonomous real-time detection presents complex technical hurdles, including issues of data accuracy, model robustness, and environment adaptability. Additionally, a considerable portion of the existing research predominantly focuses on the initial stages of penetration, with specific concentration on vulnerability discovery and exploitation. This emphasis overlooks the substantial potential for AI to revolutionize the post-exploit phases and marks a critical gap in the current research landscape. These challenges highlight the necessity for innovative solutions that are currently not fully realized and can hopefully serve as a crucial catalyst to spark the development of groundbreaking methodologies.

The literature review in this report will be dedicated to a comprehensive discussion of recent innovative methods in the field. This exploration will attempt to shed light on cutting-edge approaches and untapped potentials within the realm of AI-driven security assessments. Then, these key ideas will be extracted from the literature review to serve as a foundation for the short seminar course, titled ‘Harnessing Artificial Intelligence for Penetration Testing.’ This course is designed to show the untapped potential of AI in Penetration Testing and addresses the previously mentioned challenges.

# Background of the Problem and Motivations

The rapidly evolving cybersecurity landscape, with its complex and labor-intensive strategies, stands poised to significantly benefit from recent advancements in AI technology, including the potential for innovative applications and techniques. This surge in AI's significance is crucial for cybersecurity endeavors such as PT, the linchpin of proactive cybersecurity. The integration of AI in PT has the potential to level the playing field and empower defenders to anticipate and mitigate threats effectively.

The motivation behind the course, 'Harnessing Artificial Intelligence for Penetration Testing,' stems from a desire to comprehensively explore modern threats and current PT methods. It then aims to delve into the integration of various AI techniques, including ML, RL, and Deep Reinforcement Learning (DRL), into the practice of PT.

The core objective of the course revolves around a fundamental question: How can AI techniques be effectively harnessed for each stage of the PT process? This question is pivotal, especially as the power of AI attracts cybercriminals, as it demands that cybersecurity professionals keep pace with its rapid integration. Relying solely on traditional PT methods may render pentesters ill-equipped to effectively address modern threats. Furthermore, the course offers value not only to cybersecurity defenders but also to ethical "white-hat" hackers, security analysts, and cybersecurity researchers. By providing insights into popular AI tools and techniques, the course enables professionals to refine their ability to assess and secure systems efficiently.

Given AI's unmistakable influence in current technological research, it becomes essential to delve deeper and extend this research to PT to understand its future implications in security. The perpetual evolution of cyber threats has set in motion a continuous cycle that demands an equally dynamic response. The incessant propagation of cyber threats prompts continuous advancements in detection and mitigation techniques—a cycle that perpetuates itself with each subsequent rise of a new cyber threat. However, by using AI integration to break into this cycle and transform it from a reactive to a proactive strategy, a unique opportunity emerges. This opportunity affords pentesters the chance to not just respond to evolving threats, but proactively anticipate and prepare for the cyber threats of the future.

# Objectives

"Harnessing Artificial Intelligence for Penetration Testing" introduces innovative concepts to guide participants through the evolution from traditional PT methodologies to the forefront of AI applications. The course begins with an introduction to the fundamentals of PT and provides insights into existing manual methods. This foundational knowledge allows participants to gain detailed insights and serves as a precursor that allows them to better understand and appreciate the revolutionary potential of AI as they progress through the course.

Then, by building upon this foundational knowledge, the course adopts a phased approach that directly aligns with Splunk's outline of the PT stages. Each lecture is structured to focus on a specific phase, which allows participants to observe the practical integration of AI throughout the entire PT scenario. After a phase-specific introduction, each lecture outlines the intricacies of existing manual methods, their inherent strengths and limitations, and then explores the integration of AI to overcome these limitations. Then, by dynamically integrating insights from ongoing research, the course not only addresses the current AI landscape but also anticipates the future potential of AI in each PT phase. For the more complex final phases, the manual and AI concepts are dissected across two videos. To facilitate a thorough understanding, the initial video explores manual methods, while the second video shifts focus towards the integration of AI. By offering this dual perspective that mirrors the natural progression from manual to AI-driven methods, the course not only provides a comprehensive understanding of current practices but also instills a forward-looking perspective and anticipates the transformative potential that AI holds for each specific PT phase.

Lastly, this course goes beyond theoretical discussions, and stimulates critical thinking through research questions focused on real-world application challenges, ethics, and accountability. To further enhance the learning experience, participants engage in immersive hands-on labs and tutorial-based research to expose them to practical scenarios that closely mirror real-world challenges and demand strategic problem-solving skills. The goal is that, by bridging the gap between theory and practice, the theoretical exploration of AI integration will become more meaningful by being firmly grounded in the practical insights gained through hands-on practice. By grappling with the intricacies of manual methods, participants are better equipped to appreciate the limitations and challenges inherent in traditional approaches.

By the conclusion of the course, students should emerge with a profound understanding of the fundamentals of AI, Machine Learning (ML), and Deep Reinforcement Learning (DRL). They should not only have a comprehensive knowledge of these technologies but also be adept at recognizing their critical roles in reshaping and enhancing the practice of PT. They should be well-equipped to navigate the complex digital frontier with confidence and proficiency. Embracing AI is positioned not just as an option but a necessity for staying ahead of adversaries, securing our digital future, and safeguarding critical systems and data.

# Literature Review

## Review 1: Introduction to PT Methodology

#### Introduction

The “Penetration Testing: Practical Introduction & Tutorials" blog published by Splunk, a leading authority in cybersecurity, serves as a critical resource for understanding the intricate methodologies of digital defense. Splunk provides valuable insights and resources across various domains, including cybersecurity, compliance, data management, IT monitoring, and overall management of IT and business operations [6]. This report introduces PT as an offensive strategy to identify system vulnerabilities by simulating cyberattacks. [3] emphasizes that the primary goal of this process is to uncover weak points, understand potential attack scenarios, and analyze the severity of vulnerabilities. By emulating the actions of a genuine attacker and reporting discovered vulnerabilities, this approach allows target systems to prioritize critical security enhancements before real threats materialize.

In addition to introducing PT as an offensive strategy, the blog delves into the intricacies of the PT process. It not only discusses how PT enhances security, but it also details the specific and complex steps and procedures involved in the PT process. These steps systematically assess and analyze system security through information gathering, scanning, exploitation, connection maintenance, eliminating intrusion evidence, and reporting. It then transitions from theoretical knowledge to practical application by providing tangible examples of PT tasks and outlining the basic functionality of various essential tools. Overall, this article delivers a comprehensive perspective on the importance of PT in cybersecurity and its practical relevance to safeguarding businesses when digital threats are both inevitable and evolving.

#### Summary

This article provides a thorough exploration of the structured stages integral to PT, to facilitate a more methodical, ethical, and comprehensive assessment of a system's security. These stages contribute to the overall effectiveness of PT by simulating real-world attack scenarios, identifying vulnerabilities, assessing their severity, and guiding the enhancement of cybersecurity measures. While other sources may present slightly different PT stage breakdowns, this blog distinguishes itself by not only submitting a comprehensive analysis of these stages but also by providing concrete examples, precise definitions, and practical insights into their importance. The stages outlined in this Splunk article include ‘Information Gathering,' ‘Scanning,' ‘Exploiting,' ‘Maintaining Access,' ‘Covering Tracks,' and ‘Reporting’ [3]. This structured approach enables security professionals to systematically assess vulnerabilities and prioritize security improvements.

Due to the growing complexity of modern networks, Watts highlights skills critically necessary for pentesters and underscores the importance of continuous training and practice to excel in the field of PT. These skills encompass a wide range of technical and practical competencies, ranging from "knowledge of operating systems and networking," an understanding of "authentication and authorization mechanisms," to a strong foundation in programming [3]. Whether it is achieved through self-directed efforts, such as reading and exploring online resources, or through formal education, continuous skill development is crucial. Hands-on practice is emphasized as one of the most essential components, as it allows individuals to apply their knowledge and techniques in a controlled environment and ensures proficiency maintenance and adaptability.

The author reinforces this principle through a series of practical demonstrations. The first demonstration establishes a controlled test environment through VMWare, including an attacker and victim machine. This approach showcases the use of different PT tools in various phases, starting with gathering important network-related information: First by gathering domain-specific intelligence with Whois, then identifying vulnerable devices exposed on the internet through Shodan, and finally conducting a comprehensive network topology scan with Nmap. The examples in this foundational phase can be used to assist security professionals, obtain a detailed understanding of network infrastructure, and identify potential security vulnerabilities.

Once a thorough network visualization has been established, the author demonstrates a series of exploitation tools, such as password cracking with John the Ripper and web traffic interception and manipulation with Burp Suite. [3] then proceeds to highlight the use of Metasploit exploits to establish a “backdoor shell that [will enable him] to run commands on the victim system.”

#### Methodologies

This article thoroughly explores the fundamental techniques and processes behind PT and provides a structured approach for identifying and resolving security vulnerabilities. These methodologies are categorized into several stages, each playing a crucial role in the PT process. From information gathering and scanning to exploiting vulnerabilities and maintaining access, each step simulates real-world attack scenarios and helps organizations comprehend their system's weak points. By conducting PT, businesses can effectively prioritize security measures, subsequently enhancing the overall security of their digital assets.

The first stage, as defined by this article, is 'Information Gathering.' This initial phase is equivalent to conducting reconnaissance, during which a pentester collects essential data about the target system or organization. The primary goal is to gather relevant data to understand the available testing surface and potentially detect security vulnerabilities. Some common categories of information pentesters aim to collect in this stage include IP addresses, server details, subdomain identification, and specific software applications, platforms, operating systems (OS) or frameworks [3]. This phase is about building a comprehensive profile of the target to serve as the foundation for subsequent stages of the testing process.

After identifying potential entry points and vulnerabilities in the target system during the previous phase, the pentester begins to assess these points for possible weaknesses during the 'Scanning' phase. Exploration of the target system must be conducted systematically rather than haphazardly testing each potential vulnerability. Not only is a blind approach more time-consuming, but it is also significantly less effective. Therefore, to increase the chances of successful vulnerability detection, pentesters identify known vulnerabilities in their target’s framework and assess how the system responds to intrusion attempts [3]. This stage significantly improves the efficiency of the testing process as it refines the list of potential vulnerabilities and allows testers to concentrate on only the most relevant.

After vulnerabilities have been identified during the previous stages, pentesters actively exploit them during the 'Exploitation' phase. The goal is to simulate an actual intrusion by accessing data within the target system, intentionally triggering failures, or making unauthorized changes [3]. While this critical phase mimics the actions of a genuine attacker, the author emphasizes the importance of maintaining ethical and controlled PT practices by focusing on understanding vulnerabilities rather than causing actual damage.

After successful exploitation, the testing focus shifts from gaining initial access to the 'Post-Exploitation' phases: 'Maintaining Access,' 'Covering Tracks,' and 'Reporting.' These phases, encompassing Steps 4 to 6, align with this primary purpose of assessing and improving security measures rather than engaging in malicious actions. Unlike authentic attacks, these phases aim to evaluate and enhance security by examining the system's ability not only to detect unauthorized access but also to log and store data related to security incidents [3]. Then, by reporting their insights in detail, pentesters enable organizations to fortify their defenses against real-world cyber threats.

#### Main Findings

The main findings in the article revolve around the importance of PT in the context of cybersecurity. The report highlights that businesses undergoing significant growth are more likely to attract the attention of cybercriminals and emphasizes PT as an essential strategy for protecting their digital assets. The testing process is introduced as a proactive and offensive method for identifying system vulnerabilities by simulating cyberattacks to uncover weak points, anticipate potential attack scenarios, and assess the severity of vulnerabilities.

The article also emphasizes Splunk’s multi-phase testing process, with each stage crucial in systematically replicating real-world attack scenarios. 'Information Gathering' provides critical insights into the target, while 'Scanning' refines the focus by identifying specific vulnerabilities, thus preventing inefficient testing of unrelated weaknesses. The 'Exploitation' phase, while simulating an actual attack, strictly adheres to ethical principles to avoid harming the target system. After successful exploitation, the post-exploitation phases, ‘Maintaining Access’ and ‘Covering Tracks,' evaluate an attacker's ability to sustain a persistent presence and evade detection. The final ‘Reporting’ phase is crucial for outlining and prioritizing vulnerabilities to guide businesses to address easily exploitable weaknesses first.

Overall, this blog highlights that PT is a proactive, systematic, and highly effective approach for identifying and addressing security vulnerabilities within an organization's system. Simulating the actions of potential attackers enables organizations to fortify their security measures and safeguard their valuable assets from complex cyber threats. The article not only emphasizes the significance of PT but also provides a foundational framework for conducting the testing process through well-defined phases, all while upholding essential ethical considerations. This structured and ethical approach ensures that PT not only identifies and exploits vulnerabilities but utilizes this data to equip organizations with practical strategies to improve their overall security.

#### Relevance to Your Course Content

While the specific phases of PT may exhibit variations in terminology depending on the source, their fundamental organization and underlying concepts remain the same. Since the course is developed explicitly around these phases, it was essential to find a credible basis from which to structure the modules. As such, the decision to adopt Splunk's PT methodology into the curriculum is substantiated by their extensive expertise and experience in cybersecurity. Integrating these well-structured phases into the course helped provide a solid and reliable foundation upon which to build.

Another reason this article is an invaluable resource for the course content is because it effectively bridges the gap between theoretical knowledge and practical implementation. It not only offers an in-depth exploration of PT stages, but it also provides additional context, insights into practical applications, guidance for setting up testing environments, and real-world examples of popular tools. These practical applications demonstrate the impact of PT tools in identifying vulnerabilities, creating custom exploits, and enhancing offensive strategies. By providing a blend of theoretical insights and hands-on experience, this article develops a solid PT foundation for the course and serves as the launchpad from which to seamlessly delve into AI-enhanced cybersecurity strategies.

Some tools discussed in this report, including Nmap, Metasploit, and Burp Suite, are commonly used in the field and are essential for carrying out PT and identifying vulnerabilities. Although these tools do not innately contain AI functionality, they can be integrated with AI and ML techniques to enhance their capabilities. For example, AI can automate the detection of vulnerabilities or streamline the exploitation of weaknesses using data analysis to identify patterns or trends in data [1] [5]. The course will explore the potential of this integration and demonstrate how AI can augment the capabilities of various tools. Additionally, comprehending the functions and limitations of the manual PT methods discussed in this report is essential to better appreciate how AI-driven tools can leverage ML to enhance threat identification and exploitation.

## Review 2: Leveraging Large Language Models for PT

#### Introduction

During the upcoming European Software Engineering Conference proceedings, researchers Andreas Happe and Jürgen Cito will present a compelling exploration of the integration of Large Language Models (LLMs) into the realm of PT. LLMs, such as ChatGPT, GPT3.5, and AutoGPT, have gained significant popularity recently due to their remarkable ability to predict missing data and generate human-like text. As a result of these pattern-recognition abilities, which are learned through extensive training, the authors recognized the potential for leveraging LLMs to identify vulnerabilities, execute custom exploits, and even act as virtual sparring partners to their human counterparts. This integration could provide guidance to not only “empower existing human security testers,” but could also “counteract the lack of sufficiently educated security professionals,” thereby addressing a current critical shortage of skilled experts in the field [4].

#### Summary

With the aim of determining to what extent security testing can be automated through LLMs, the authors framed their research question around the deployment of these models as virtual sparring partners for security professionals. To provide a structured framework for their investigation, they turned to MITRE ATT&CK, a comprehensive repository of knowledge concerning threat actors in the cybersecurity domain. Their goal was to produce a proficient sparring partner, capable of covering a diverse array of tactics, techniques, and procedures (TTP) summarized within ATT&CK [4].

To comprehensively explore their hypothesis, the authors led a series of experiments, where they conducted demonstrations with both high- and low-levels of guidance. These demonstrations vary in detail and specificity, with high-level addressing general PT aspects and low-level dealing with more detailed, practical actions. For their high-level demonstration, they employed LLMs to assist the planning phase of a PT. This involved tasking the LLM with designing the test itself, including determining TTPs and identifying potential vulnerabilities [4]. They then explored low-level guidance, during which they engaged the LLM to assist in the execution phase of the PT. As it is assumed that pentesters have completed their high-level analysis by the time they begin a low-level analysis, this stage is often in a step-by-step format and includes activities such as identifying systems, targeting specific vulnerabilities, executing custom commands and exploits, and providing information on how to escalate privileges.

#### Methodologies

In pursuit of answering their research question and exploring their hypothesis further, the authors aimed to demonstrate the extent and effectiveness of deploying LLMs as virtual sparring partners. In order for the LLMs to meet the authors expectation of success, the models must produce valid and “suitable tactics and corresponding techniques [4].” To test the practicality of using LLMs as sparring partners, the authors built upon the framework established in their research question with carefully designed experiments that encompassed both levels of guidance. Their approach ranged from broad and theoretical to highly specific and practical, which allowed them to assess the capacity and applicability of these models.

While the traditional approach to leveraging LLMs in PT requires human testers to manually initiate conversations using prompts, the authors sought to automate this process by using pre-trained Autonomous AI Agents: AutoGPT and AgentGPT. Not only do these agents increase productivity, but the incorporation of "external knowledge and automated feedback” can mitigate the occurrence of fact inventing, referred to as hallucinations [4]. Each tool can operate independently, eliminating the need for constant human intervention. This is accomplished by automatically breaking down predefined tasks into smaller, specialized subtasks through the use of “self-prompts [7].” However, despite their simulation similarities, AutoGPT is described as having more decision-making capabilities than AgentGPT, while AgentGPT offers a more user-friendly experience that welcomes a wider range of users [8]. However, as both AutoGPT and AgentGPT can successfully accomplish an assigned objective from a single directive, they are considered valid options for PT.

In the high-level experiments, the authors focused on the LLMs' potential to provide strategic guidance for both a general and specific target using both autonomous agents. For the general scenario, they provided AgentGPT with the task of “becoming a domain admin in an Active Directory,” and for the specific target, they tasked AutoGPT with creating a PT plan [4]. These experiments were considered successful as both AI agents provided responses which were “realistic, and feasible, and would give a penetration tester good feedback about potential attack vectors [4].” However, it is important to note that while AutoGPTs functionality also enabled it to crawl the target’s website, it declined to perform certain actions, citing ethical concerns.

In contrast, the low-level guidance experiments focused on providing step-by-step guidance, offering detailed actions such as identifying and exploiting system-specific vulnerabilities, executing custom commands and exploits, and providing insights on privilege escalation. At this stage, it was assumed that pentesters had already completed their high-level analysis, obtained some basic level of access to the system, and simply required guidance to escalate to root. Therefore, the goal of this experiment was to achieve privilege escalation and gain root access on a deliberately vulnerable Linux Virtual Machine (VM). The authors used Python to set up a connection between GPT3.5 and the vulnerable VM and asked the LLM to analyze the VM's state, generate commands or actions, and potentially control or influence the VM's behavior. The script operated in an infinite loop, instructing GPT3.5 to suggest Linux shell commands, execute them over SSH on the vulnerable VM, analyze the command and its output, identify potential security vulnerabilities, and finally provide steps on how to exploit them. The results showed that GPT3.5 successfully obtained root privileges, identified and exploited security vulnerabilities, and retrieved essential system files for privilege escalation [4].

#### Main Findings

During the experiments, the researchers found that the LLM displayed signs of understanding causal relationships and exhibited a degree of logical thinking in its suggestions for PT tasks. These suggestions followed logical patterns, even when specific information about the target system's configuration or vulnerabilities was not provided. The authors highlighted that these suggestions, while “eerie”, were primarily generated “based upon pattern-matching and not on a deeper understanding” of the subject matter [4].

The authors also found that, on a small scale, the performance of LLMs appeared unstable and inconsistent, and often produced a large variation in generated commands and identified vulnerabilities. During individual and short runs, the LLM would become too fixated and overly focused on a specific detail and lose sight of the broader picture, similar to “going down a rabbit hole [4].” While extending or combing results from multiple runs led to more consistent outcomes, LLMs were deemed less predictable and consistent compared to traditional enumeration tools like linpeas.sh in their current state.

LLMs were also found to be limited by their ethical filters, which prevent the AI from generating responses or taking actions that could engage in unethical behaviors. This was shown during the experiments when AutoGPT refused to execute additional network scans or phishing attempts. The authors found that many of these restrictions could be bypassed by running the LLM locally or by using prompt engineering to test slight prompt variations and reduce triggering ethical filters. The simplicity of engineering prompts was shown when the authors requested “verification commands for vulnerabilities” instead of “exploits for vulnerabilities” and when they instructed the AI not to “ask questions or provide judgments [4].” While these techniques prove effective in reducing ethical denials, they also raise concerns about potential misuse. Due to the ease and accessibility of LLMs, they can be employed by both legitimate security professionals and malicious actors.

While the experiments with LLMs have showcased their potential in providing valuable PT guidance, there remains a pressing need for further refinement in their application. The findings indicate that LLMs, although proficient at pattern recognition and generating suggestions, still rely heavily on data-driven responses rather than a true comprehension of security systems. Addressing the challenges of occasional hallucinations and variability in single runs, especially when overly focused on specific aspects, is crucial to ensure their reliability. However, the urgency to incorporate AI in PT is crucial. As the field faces a critical shortage of skilled security professionals, it becomes increasingly vital that the relationship between pentesters and AI is strengthened. As the cybersecurity landscape evolves, preparing for AI-driven attacks becomes not only a necessity but also an opportunity for the industry to stay ahead in the ongoing battle against emerging threats.

#### Relevance to Your Course Content

This paper aligns with the course content as it explores the integration of a familiar AI model, specifically LLMs, into the field PT. As the LLMs discussed in this paper are among the most recognizable AI models, they provide an ideal starting point for introducing the central theme of the course. Their familiarity offers a comfortable and approachable introduction to ‘Harnessing Artificial Intelligence for Penetration Testing.’ Furthermore, this report not only engages in theoretical discussions, but also delves into the practical application of popular prompt-response techniques within PT. Through tangible examples, it illustrates how AI can enhance various facets of PT, offering both a relevant and captivating perspective to witness firsthand the transformative potential of AI from the outset.

Additionally, the paper introduces important ethical dilemmas that arise when AI is used as a tool in PT. Ethical considerations hold great significance in the cybersecurity domain, and addressing these issues early on is crucial. The report explores the effectiveness of prompt engineering, raising questions of its acceptability and ethical boundaries. It also addresses the accessibility of these powerful tools to both security professionals and malicious actors, prompting considerations about the distinctions between their respective ethical codes. Exploring and understanding these ethical complexities is a vital step to exploring PT.

## Review 3: Gathering Information with AI and Reinforcement Learning

#### Introduction

The report by Ghanem and Chen focuses on the initial step of PT, known as Gathering Information; its primary focus is on how the integration of AI, particularly RL, can revolutionize this critical phase. RL has quickly become one of the most important PT advancements, resulting from the recent integration of AI and cybersecurity. This transformative approach to ML enables systems to learn through experiences from interactions with their environments. The incorporation of RL into automated PT techniques not only increases productivity, but also limits common human errors. However, existing automation systems have limitations in their scope and optimization that result in an inability to comprehensively address all potential threats while efficiently managing resources. Recognizing these challenges, Ghanem and Chen's research paper sets forth to employ ML techniques in the development of an Intelligent Automated Penetration Testing System (IAPTS) that will be “capable of imitating human PT experts in performing an intelligent and automated pen test [1].”

#### Summary

This research delves into the complexities of PT, an area that humans themselves often find challenging. The authors emphasize that blind automation, which entails complete automation without any human intervention, is impractical. This is particularly true during the initial phases of PT as the explorative nature often yields incomplete conclusions and requires continuous revisitation/changes in approach. As such, utilization of AI at this stage tends to result in uncertainty. However, the authors suggest that by using RL to automate these phases intelligently, it can more closely resemble a human expert’s decision-making process.

The challenges associated with automation in PT are not new, as autonomous systems have been employed in the industry for some time. However, these current systems often require substantial hands-on guidance, extensive time and resources, and are limited to smaller networks. Especially considering “PT should be repeated and performed on a regular basis to ensure continuous security,” Ghanem and Chen's work suggests that intelligent automation holds the key to significantly improving various aspects of PT [1]. These improvements would not only reduce the cost of manual, repetitive, and methodical testing but could also make PT more efficient and targeted. This streamlining and automation of repetitive tasks would reduce testing time, foster adaptability, and facilitate the exploration of innovative and unconventional techniques.

With this objective in mind, the authors advocate for the use of RL in PT, noting that RL aligns well with the “goal-directed learning and decision-making processes” required in the PT context [1]. Unlike manually created rules and configurations, RL learns through the consequences of its interactions, focusing on long-term goals rather than short-term fixes. This emphasis on RL represents a crucial step in addressing the challenges posed by PT automation and is converted into a formal computational model known as a Partially Observed Markov Decision Process (POMDP).

#### Methodologies

The methodologies employed in Ghanem and Chen's research revolve around the innovative application of RL within the framework of POMDP. This approach seeks to address the challenging PT scenario where an “agent cannot determine with full certainty the true state of the environment” by encompassing essential elements such as state observations, selection policies, dynamic transitions, and rewards [9]. Within this framework, an RL agent learns to make decisions based on its observations, with the goal of maximizing cumulative rewards. The strategies executed by the RL agent that returns the largest reward value are then stored in memory for similar cases in the future, thus enabling it to autonomously tackle complex PT problems.

Ghanem and Chen tackle these challenges by integrating a combination of advanced algorithms, PERSEUS and PEGASUS, which are specifically designed for solving POMDPs. PERSEUS, a “randomized point-based value iteration” algorithm, simulates various random scenarios to obtain a set of educated guesses, which is referred to as a belief set [1]. These guesses represent possible situations or states of the environment based on the limited information available to the AI agent. This understanding is then improved gradually, as the algorithm updates its belief set after every simulation to ensure that each value either improves or at least remains constant [9].

Alternatively, the PEGASUS algorithm is a policy search method that seeks to determine optimal sequences of actions, known as policies, that maximize cumulative rewards over time. It transforms the problem into an equivalent deterministic POMDP, where each state-action pair has only one possible outcome. PEGASUS then conducts a set number of simulations, iteratively refining the policies to maximize their estimated cumulative reward value [9]. This approach is particularly effective in solving large POMDPs, making it suitable for addressing the challenges posed by PT, as it contains a “polynomial rather than exponential” time complexity, making it suitable for large-scale PT scenarios [1].

During the learning process for their proposed system, IAPTS relies on human input as experts provide knowledge to teach the system. However, over time, the system evolved, gaining the potential to develop autonomous learning modules that reduce the need for manual intervention. This evolution aligns with the various operational modes of IAPTS ranging from fully autonomous (Level 4) to learning mode (Level 1), where a human expert performs PT while the system observes and learns.

The primary goal of testing IAPTS was not only to evaluate its capabilities but also to demonstrate the suitability and effectiveness of applying RL to PT. The researchers conducted two main types of tests within controlled environments: Simple Simulation and Experience Replay. In the Simple Simulation, they set up a simulated network consisting of seven machines (M0 to M6) to mimic real-world PT scenarios. This allowed them to gain insights into how IAPTS would perform under various conditions, measure its performance metrics, assess execution times, and identify potential weaknesses. Alternatively, for the Experience Replay tests, the researchers simulated scenarios in which the same network underwent updates and upgrades. These tests aimed to evaluate how well IAPTS learned and adapted to changes in the network, further confirming its potential for automating PT processes.

#### Main Findings

The main findings of Ghanem and Chen's research paper provide valuable insights into the field of PT. In their Experience Replay tests, they discovered that the system successfully learned and stored knowledge from previous tests, with policies being effectively reused in most instances. This highlights the system's adaptability and capability to learn from past experiences, a crucial feature for PT automation. When compared to traditional manual methods, which rely on human expertise, and the blind automation approach, where tasks are automated but lack intelligent decision-making, IAPTS, significantly reduces the time required for testing and outperforms both approaches in terms of efficiency and effectiveness. This not only saves time and resources but also generates alternative attack strategies that humans may overlook.

The RL-generated attack policies also proved to be highly relevant and accurate, especially when targeting the most secure machine in the network. These policies were deemed plausible and realistic, mirroring how actual attackers might approach and execute an attack on the target system. Additionally, IAPTS was intentionally designed with flexibility in mind, permitting the seamless incorporation of new features and functionalities in the future. This modern design ensures IAPTS remains a versatile and evolving tool in the field of PT, through continual enhancement of its performance and capabilities.

#### Relevance to Your Course Content

This research paper offers a comprehensive overview of PT, including its purpose, advantages, disadvantages, and the intricate challenges associated with its pertinence to the first step in PT - Gathering Information. By emphasizing the extensive data collection and assessment required during manual execution of this phase, the authors highlight the necessity for discussions on automation in AI. Through practical simulations, the authors demonstrate how these solutions can significantly reduce human effort, enhance accuracy, improve adaptability, and expedite tasks, ultimately proving that automation can make the PT processing more efficient.

This report also introduces advanced techniques, such as RL and POMDPs, within the context of PT. RL, being a subset of AI, holds particular relevance in automating various phases of PT. It also highlights the practicality and adaptability of RL by exploring its application in partially observable environments, utilizing belief sets instead of the Q-tables addressed later in a fully observable scenario. This incorporation of RL and POMDPs in partially observable environments not only signifies the direction of future research but also illustrates that automated PT is an evolving field marked by ongoing developments. As such, this paper not only demonstrates the current achievements, but also serves as a preview of the extensive possibilities and potential advancements within the field.

Therefore, incorporating this paper into my seminar course can provide additional context to PT, especially in the initial Gathering Information phase, and explain how advanced AI-driven techniques, such as RL, are transforming the field. My goal is that it will ultimately serve as a useful resource to introduce the challenges of PT, automation as a solution, and the application of AI in enhancing cybersecurity practices.

## Review 4: Enhancing the Scanning Phase with GyoiThon

#### Introduction

The research paper “Penetration Testing Procedure using Machine Learning” focuses on the second phase of PT - the scanning phase, with a particular focus on assessing the effectiveness of GyoiThon. GyoiThon is a PT tool integrated with ML capabilities, specifically leveraging the Naïve Bayes algorithm, that primarily focuses on automating data acquisition from target URLs [10]. This integration represents a significant advancement within the field of cybersecurity, as it not only enhances the speed and efficiency of vulnerability detection, but also introduces the potential for more precise identification of security weaknesses. By leveraging other PT tools to enhance its capabilities, GyoiThon extends its utility beyond traditional methods. It automates the process of gathering data from target URLs, thus streamlining the scanning phase while reducing the time and effort required by pentesters.

#### Summary

In this study, the researchers set out to address the fundamental research question: ‘How effective is the GyoiThon tool in detecting vulnerabilities [10]?’ The hypothesis guiding this exploration speculates that PT tools integrating ML algorithms will exhibit greater effectiveness in searching for and identifying vulnerabilities compared to their non-ML counterparts. To highlight this premise, the paper briefly examines common ML-based PT methods used in the field, including tools known for simulating real-world attacks, detecting vulnerabilities, and addressing security weaknesses. Their analysis provided valuable context and benchmarks for evaluating GyoiThon's performance and offers support for their hypothesis regarding the capabilities of using ML in PT.

This exploratory study places a particular emphasis on comparing GyoiThon's default mode with its ML mode, executing each of them within controlled environments. Through these experiments, the researchers explore the capabilities of GyoiThon and showcase its ability to enhance PT. By exploring the effectiveness of GyoiThon, the authors assess its efficiency in detecting known vulnerabilities, identifying software components, discovering configuration weaknesses, highlighting authentication issues, and pinpointing general web application vulnerabilities [10]. These capabilities emphasize GyoiThon's pivotal role in the scanning phase of PT, highlighting the demand for advanced tools and techniques to navigate the complex landscape of cybersecurity.

#### Methodologies

The methodology employed in this report is particularly significant as it delves into a novel area of interest within cybersecurity. It's worth noting that this comprehensive study of GyoiThon represents a unique endeavor, as the only prior study into the capabilities of the tool was conducted by its developer. As such, the researchers had the distinct advantage of operating within a flexible framework that lacks predefined steps, which enabled them to create new procedures to address their research question [10].

As for execution, the researchers established an isolated testing environment using the Kali Linux operating system within VirtualBox. Within this controlled environment, GyoiThon was employed to detect vulnerabilities related to data exchange; it analyzed both unencrypted HTTP traffic on Port 80 and encrypted HTTP traffic on Port 443. The target websites were hosted on a server provided by OWASP and accessed via a locally hosted environment.

Since the researchers' hypothesis centered on comparing PT tools with and without ML algorithms, their analysis was limited to GyoiThon's Default Mode and ML Mode. The Default Mode encompassed various steps, including parsing HTTP responses, identifying product/version information, assessing vulnerabilities using Common Vulnerabilities and Exposures (CVE) numbers, examining HTML and JavaScript comments, analyzing debug messages, and assessing login pages [11]. In contrast, the ML Mode incorporated all the steps from the Default Mode, but additionally utilized the Naïve Bayes algorithm for product/version identification [11]. This setup enabled researchers to directly evaluate the effectiveness of the Naïve Bayes algorithm in the realm of PT, aligning with their hypothesis.

#### Main Findings

The analysis of the PT procedure conducted using GyoiThon revealed several significant insights. First, it was observed that Port 80, commonly associated with unencrypted HTTP data, exhibited a higher number of vulnerabilities for both the Default and ML modes. This finding aligns with expectations, as Port 80's lack of encryption renders it less secure compared to HTTPS (Port 443). This absence of encryption causes Port 80 to be more susceptible to vulnerabilities and potential attacks, as was reflected in the test results. However, the variation in the number of vulnerabilities detected between these ports decreased with the use of ML mode; by identifying three additional vulnerabilities in Port 80, ML mode reduced the disparity in vulnerability frequency from six to only three [10]. Not only does this outcome highlight the potential of GyoiThon, but it also supports the hypothesis that integrating machine learning into PT tools enhances their effectiveness in identifying vulnerabilities.

While these initial results demonstrate success, it is essential to note that GyoiThon relies on external sources, such as the National Vulnerability Database (NVD), to gather information about vulnerabilities. This reliance is a limitation of the tool’s capabilities as it may be unable to identify vulnerabilities that have not yet been documented in the NVD. This potential blind spot highlights the importance of staying updated with emerging threats and identifies an aspect requiring improvement. As such, while GyoiThon showcases promise as a valuable PT tool, the researchers explain that future testing against real websites and a comprehensive assessment of all nine modes is necessary to obtain a more comprehensive understanding of its capabilities [10]. These findings contribute to the ongoing development of AI-driven PT tools and emphasize the need for continuous refinement to stay ahead of evolving cyber threats.

#### Relevance to Your Course Content

This report extensively explores the application of ML in PT, using GyoiThon as a focal point, and aligns with the central theme of my course. With our aligning goals of providing a thorough understanding of AI techniques for cybersecurity, this report serves as a valuable reference for my course by delving into the practical application of a specific ML tool. This significance is particularly true for the second phase of PT, scanning, which is the primary focus of my second module and the explicit function of GyoiThon, which effectively identifies vulnerabilities by scanning web pages.

Through a comparative analysis between default PT methods and those augmented with AI, this article showcases the effectiveness of AI-driven approaches through direct evidence. This novel and practical study not only highlights the superiority of ML-enhanced techniques but also emphasizes the potential transformative power of AI within the cybersecurity domain. Through empirical evidence, this report encourages further exploration into the integration of AI into the field of PT and invites active engagement for modern AI security solutions.

Arguably most importantly, the article delves into the discussion of common vulnerabilities found in web applications and the various tools used to detect them. This practical understanding of vulnerabilities and the tools and techniques available for their detection and mitigation is essential for effectively navigating the complex digital domain.

## Review 5: Exploitation in PT with RL

#### Introduction

In the paper titled “Vulnerability Exploitation Using Reinforcement Learning,” the authors leverage modern PT techniques, specifically ML and RL, to automate one of the most critical phases in cybersecurity: exploitation. By prioritizing actions that maximize rewards, RL underscores the importance of developing tools that not only identify vulnerabilities but also utilize ML to efficiently exploit them. The authors focus goes beyond automation and emphasizes the need for further evolution in PT to address the complex field of cyber security.

The intelligent agent created in this report prioritizes adaptability, ensuring it can be trained on a wide array of vulnerabilities and operating systems. This approach offers a tailored and intelligent approach to exploitation that challenges traditional methods, which often involve resource-intensive, brute-force techniques that are time and resource intensive [5]. To accelerate the PT process and ensure a more targeted and efficient approach to identifying and exploiting vulnerabilities, this agent leverages Metasploit, a well-known PT tool with a wide range of payloads for various purposes.

What further sets this approach apart is the agent’s ability to archive successful exploits as states alongside corresponding payloads with high success rate. The agent then intelligently leverages this payload repository, known as a Q-table, to execute exploitation with precision – a milestone that demonstrates the potential of RL to leverage an award system and continuously refine and enhance exploitation strategies using AI. This report provides a look into the future of PT, where customization, adaptability, and intelligence combine to not only identify vulnerabilities but to masterfully exploit them.

#### Summary

In this report, the authors utilize ML to create an RL agent that makes decisions by interacting with a fully observable environment. The primary focus of this RL agent lies in the exploitation phase, the third and crucial step in PT. Through an extensive training process, the agent interacts with a simulated environment, dynamically adapting its exploitation strategies by analyzing various factors, including the environment configuration. This adaptive approach is made possible by representing the environment as states, each defined by a unique combination of operating system and vulnerability [5]. These states are then linked to payloads that have demonstrated a high likelihood of success and are stored in a Q-Table. Due to the variability in payload effectiveness based on these states, the authors reward successful attempts, which they define as "the establishment of a reverse shell session following payload execution [5]." Therefore, even in instances where the payload is not successful, the RL agent adjusts its decision-making process based on the rewards it receives; it then learns to prioritize actions that result in positive rewards.

Once the RL agent is trained, it is deployed in a real-world scenario where it encounters target systems with specific operating systems and vulnerabilities. Metasploit serves as a valuable resource as the RL agent selects and utilizes payloads based on its learned strategies, facilitating effective delivery of exploits to compromised target systems. The extensive payload options offered by Metasploit enhance the agent's versatility during the exploitation process. This integration contributes to the authors primary goal of creating a versatile “general agent that is capable of exploiting any/general task and making the appropriate decision [5].”

This combination of ML, RL, and established PT tools represents a significant advancement in the merging of AI and cybersecurity. Through the incorporation of RL algorithms and their integration with established tools like Metasploit, this report demonstrates an evolution of PT. This innovative approach showcases the potential of AI-driven agents to optimize and streamline exploitation tasks, ultimately benefiting cybersecurity professionals in identifying and addressing vulnerabilities in a more efficient and effective manner.

#### Methodologies

The methodologies employed in this study consist of two important phases, the Training Phase and the Exploitation Phase. During the Training Phase, an intelligent agent is developed through the application of RL techniques, using a guess-and-reward system. This phase involves the agent navigating a simulated environment, in which it uses an “epsilon greedy strategy” to make informed decisions by balancing exploration (delivering a randomly selected payload) and exploitation (selecting a specific payload that will yield the highest expected reward based on its learning so far). The agent then receives rewards based on the success or failure of a particular payload, from which it builds a valuable repository of previous exploits and their results. The training phase is then repeated for a certain number of iterations, with a gradual decrease of exploration.

To motivate its decision-making, a point-based reward system is employed that offers substantial rewards for success and imposes penalties for failures. These rewards are maximized by leveraging the Q-learning algorithm, to “determine the best series of actions to take based on the agent’s current state [12].” This approach often results in the agent executing calculated and cautious actions to minimize risks [5].

The learning phase honed its exploitation skills across seven trials, during which the agent spent an average of 2.5 hours executing 500 attempts to exploit vulnerabilities. During this phase, the agent's primary focus was on continuous learning and strategy refinement. It actively experimented with different actions, assessing their success or failure, and served to provide insight into valuable tuning parameters from controlling the importance of new versus old information, long-term versus short-term rewards, to exploration vs exploitation [5]. An assessment of the agent's performance is then calculated to determine how effective it is at establishing a reverse shell. This computationally intensive process positively reflected the agent's ability to actively learn and adapt its exploitation techniques by making informed decisions.

In the exploitation phase, the RL agent took advantage of its learned strategies, drawing insights from its repository to effectively select payloads from the Metasploit framework. To simulate real-world scenarios, it was deployed on multiple vulnerable machines with a "remote code execution" [5] vulnerability found in Apache CouchDB, specifically Version 3.1.0. The agent's primary objective was to establish a reverse shell, which was achieved with remarkable efficiency by leveraging payloads with the “highest rank in the Q-Table [5].” Impressively, it accomplished this goal in an average of just 8.26 seconds across the tested systems. This performance indicates that the training phase prepared the agent well and proved its ability to effectively execute learned strategies against real-world systems.

#### Main Findings

The study's main findings highlight the remarkable effectiveness of the RL agent in automating exploitation tasks, particularly within the realm of PT. As the RL agent gains experience through training, it exhibits a gradual shift from exploration to exploitation, becoming more discerning in its actions. For example, while it initially explores new actions to gather information, over time it prioritizes actions it has deemed effective for achieving its goals. This transition, combined with the selection of optimal parameters, consistently resulted in an average success rate of 83.64% and an average exploit time of 8.26 seconds [5]. These notable statistics highlight the potential of the RL approach to significantly reduce the time and resources required for PT, presenting a novel and cost-effective solution to the challenges of vulnerability exploitation.

In contrast to traditional exhaustive testing methods, which often follow rigid approaches, the RL agent's adaptability and capacity for fine-tuning its strategies prove beneficial. By focusing on maximizing overall rewards and balancing learning and randomness, the RL approach proves more efficient and effective in verifying exploitable vulnerabilities. In summary, the main findings of this study emphasize the RL agent's aptitude for automating exploitation tasks, its proficiency in achieving PT objectives, and its potential to revolutionize vulnerability assessment practices.

#### Relevance to Your Course Content

This study explores modern techniques in cybersecurity, highlighting the innovative use of RL algorithms for vulnerability exploitation and emphasizing the field's dynamic nature. The authors take a comprehensive approach as they explore not only the capabilities of RL but also its adaptability. Notably, they explore versatile fine-tuning options, such as learning rate and exploration rate, and provide insights into the impacts of these methods. The study also examines RLs application in fully observable environments, utilizing Metasploit and Q-tables instead of the previously mentioned POMDP and belief sets in partially observable scenarios. This multifaceted exploration demonstrates how RL techniques can be adapted and leveraged effectively across different cybersecurity scenarios, aligning seamlessly with our course's goal of understanding AI techniques in cybersecurity.

Additionally, since RL consistently selects the most effective actions to maximize rewards, it directly addresses a critical aspect of Penetration Testing, particularly in Step 3 - Exploitation. By prioritizing the actions that yield the highest rewards, RL showcases the importance of developing similar tools that not only identify vulnerabilities but utilize ML to efficiently exploit them. Overall, this research broadens perspectives on the possibilities within the field of cybersecurity and highlights its crucial role in staying current with the dynamic landscape of digital threats.

## Review 6: Automating Post-Breach PT with RL

#### Introduction

In the report, “Automated Post-Breach Penetration Testing through Reinforcement Learning,” the authors introduce the concept of using RL, a subset of ML, to automate the post-breach phases of PT. Theses phases occurs after the initial breach of a system and focus on privilege escalation, maintaining persistence, and further exploration [3]. This approach aims to automate and enhance the capabilities of an AI agent, allowing it to navigate and effectively interact with diverse network environments. The agent is trained through interactions with various networks and prioritizes the balance of exploration and exploitation. The importance of this study is emphasized as there is still a substantial lack of testing for automation in the post-exploitation phase [13]. This lack of research is particularly dangerous as current PT practices are rapidly growing in complexity and resource consumption. In an attempt to mitigate these challenges, researchers explore the application of AI techniques, specifically the Deep Reinforcement Learning (DRL) subset, Deep Q-learning (DQ). They hope that by leveraging neural networks to directly map input states to action-Q-value pairs, their agent will excel in navigating complex environments without requiring a detailed model of the environment [12].

#### Summary

This report highlights the current limitations in the field, pointing out that even with the use of automated tools, current PT practices remain complex and resource intensive. In response to the limitations of the popular RL algorithm, Q-learning, and driven by recent advancements in deep Q-learning (DQ) algorithms, the authors made a deliberate choice to adopt DQ as their training model. Q-learning struggles when tasked with handling intricate systems or environments as it becomes computationally expensive to maintain the Q-table [12]. To overcome these challenges, the authors explore the implementation of alternative approaches such as DRL.

DRL has emerged to address these challenges and offer more solutions for RL in larger and more intricate environments. Instead of relying on a detailed model of the environment, DQ learns directly from interactions with the environment by utilizing neural networks. These networks take the current state as input, produce estimated reward values known as Q-values for all possible actions in that state, and associate each action with a unique Q-value, where higher values indicate more favorable actions. In other words, DQ distinguishes itself from traditional Q-learning by replacing the Q-table with a neural network that directly produces recommendations for actions based on the current state [12]. The agent then employs a decision-making process that involves comparing these Q-values and selects the action linked to the largest reward. These rewards are then used to adjust and refine the Q-values in both traditional Q-learning and DQ. Just as it does in the traditional method, these Q-values become more accurate over time and enhance the agent's ability to make decisions.

Overall, while the paper primarily presents a conceptual framework, it lays the groundwork for practical applications of AI in post-exploitation cybersecurity. It introduces key concepts such as Q-value estimation, exploration-exploitation balance, and the importance of realistic training environments. The paper's focus on future research and development suggests RL’s potential to shape the future of AI-driven PT.

#### Methodologies

For their research, the authors propose an architecture which involves employing a Deep Q Network with TF Agents, constructed on the TensorFlow library as the fundamental framework for training the RL agent. TF Agents is a specialized, modular software that leverages the capabilities of the TensorFlow framework [14]. By building upon TensorFlow and incorporating TF Agents, they ensure a solid foundation for training the AI agent by harnessing the power and flexibility of these frameworks.

To create a realistic PT environment, the authors plan to deploy the agent in virtualized Linux and Windows servers. These environments were crafted to simulate authentic cybersecurity scenarios, effectively mirroring computer networks tailored for cybersecurity training and competitions. Emulating target data, these environments included critical elements commonly found in these networks such as password files, shadow files, and system configurations. The agent's actions are confined to a predefined list of terminal commands, with the specifics of these commands being adapted based on the agent's observations within the environment. The resulting performance will then be gauged based on its adeptness at exploration and exploitation within these environments, with scores serving as rewards critical for reinforcement [13]. The balance between exploiting actions that appear promising based on its current knowledge (Q-values) and exploring new actions to discover potentially better strategies is carefully managed to optimize the agent's learning process and overall performance.

To establish a performance baseline for their research, the authors created a Python script designed to locate files within simulated environments. This script serves as a reference point to assess the AI agent's performance, enabling a direct comparison between the agent and the script's capabilities. This python script is tested across the servers to determine its effectiveness in locating files, particularly focusing on configurations and log files. To assess the trained agent's effectiveness, a comparison will be drawn between the script’s performance and that of an RL agent “created using randomized policy [13].” This comparison will ultimately provide insights into the efficiency of the trained agent's policy.

The methodologies discussed in this report offer a comprehensive plan for training and evaluating an RL agent for automated post-breach PT. This plan encompasses Deep Q-Learning, specialized architecture, simulated environments, action-reward configuration, baseline testing, and performance comparison. These methodologies are integral to the successful development and assessment of the proposed automated PT approach.

#### Main Findings

One of the primary findings of the report is the recognition of the applicability of RL, specifically DQ, in the domain of PT. The report emphasizes that RL offers a promising avenue for automating the post-breach phases of PT, a field where modern practices are lacking research and traditional practices are quickly becoming obsolete. While this report focuses on conceptual feasibility of this incorporation, it directly addresses the beginning steps required to train an RL agent to perform tasks in a compromised network environment. This discovery is significant as it not only validates the role of ML in enhancing cybersecurity measures but also opens doors to the development of more efficient and effective PT methodologies, especially in regard to the post-exploitation phase.

Through the analysis of automation of the post-exploitation phase, the authors set the stage for ongoing and future investigations. The report communicates the researchers' intention to implement the proposed approach, train the RL agent, and expand the model's applicability to a broader spectrum of network environments. While this report primarily focuses on conceptual feasibility, it is also practical and action-oriented, with the aim of making meaningful contributions to the cybersecurity field by advancing the automation of PT practices.

#### Relevance to Your Course Content

This report highlights the critical application of AI, specifically RL, in automating and optimizing the post-breach stages of PT. Notably, it offers practical examples that demonstrate how AI can be effectively deployed during this critical phase and advances uncharted territory in doing so. This novel discussion addresses the imperative need for cybersecurity professionals to remain current and adaptive in the face of evolving threats and encourages innovation.

Additionally, this study provides insights into the advanced technologies that power AI-driven PT, notably highlighting the significance of DQ and its role in improving traditional Q-learning methods. As ML and deep learning are at the core of AI's capabilities, delving into these technological intricacies establishes a robust foundation to comprehend how AI models are constructed, trained, and effectively deployed. This comprehensive understanding is essential as it empowers individuals to make informed decisions, adapt AI tools to specific cybersecurity challenges, and innovate within the field, ultimately contributing to the ongoing evolution and effectiveness of AI-driven PT practices.

## Review 7: Deep RL in PT

#### Introduction

The article "Autonomous Security Analysis and Penetration Testing" from Arizona State University introduces an innovative framework designed to address the growing challenge of evaluating network security amidst the complexity of expanding networks and the shortage of cybersecurity professionals. By leveraging advanced RL techniques based on DeepQ Networks (DQN), this framework in this study integrates vulnerability information into the PT processes. It associates RL reward values with Common Vulnerability Scoring System (CVSS) scores, enabling prioritization of the most critical vulnerabilities. The result is a highly efficient, automated PT system that can significantly reduce assessment time and improve overall efficiency.

#### Summary

Previous research on using RL for automating PT has focused predominantly on smaller networks and often failed to harness vulnerability information effectively. These traditional AI models have struggled to grasp the intricacies of real-world networks, falling short in accounting for the specific network structure, distribution of vulnerabilities, or correlation between vulnerabilities and exploitation probabilities [15]. This limitation has led to difficulty in prioritizing vulnerabilities, resulting in less accurate and efficient security assessments. To obtain essential information about a target and its associated vulnerabilities, these methods often rely on known sources, scans, or manual analysis for identification. This overall failure of traditional AI models to comprehensively understand the nuances of the dynamic and complex security landscapes of modern networks has resulted in a desperate need for a more comprehensive approach to PT.

Recognizing these limitations, the authors introduced the Autonomous Security Analysis and Penetration Testing framework (ASAP) as an innovative approach to security analysis and PT. This autonomous system not only understands the interconnectedness of vulnerabilities and their relation to a network's structure, but it also leverages an RL reward system based on vulnerability severity and exploitability. The approach adopted by the authors emphasizes domain-specific modeling by integrating the CVSS to quantify known vulnerabilities. This system tracks the severity of vulnerabilities and the complexity of exploiting them, allowing for a more comprehensive understanding of the network's security landscape. This modeling approach harnesses state-transition diagrams to visualize the most optimal PT policy for the network. These diagrams represent different network states and the associated actions, including probability values derived from the vulnerability's Access Complexity (AC). By generating autonomous attack plans and validating them against real-world networks, ASAP creates a comprehensive map of security threats and potential attack paths. This approach ensures efficiency not only in smaller networks but also in large-scale environments, demonstrating its exceptional performance and scalability.

To enable autonomous PT, the authors adopt an RL-based AI algorithm to identify the optimal "attack path that maximizes the reward value for the pentester [15].” RL is a concept where an agent learns through the consequences of its interactions within an environment, focusing on long-term objectives; this can be compared to security professionals experimenting with attack strategies against vulnerabilities until successful exploitation. However, what sets their RL model apart from other traditional AI models in the PT domain is that the authors propose using a DQN-based RL model. Since DQ models learn directly from interactions with the environment by utilizing neural networks, it is more equipped to handle diverse network conditions, including those that may not have been encountered during training. As such, their RL approach involves dynamic interactions with the environment by considering the current user privilege level, actions linked to vulnerability exploitation, the difficulty and probability of a successful action, reward values, and the decision-making process. The outcome of this method is a carefully designed attack plan that "guides the security professional" through subsequent actions based on their user privilege and progression strategy [15].

#### Methodologies

The methodologies of ASAP involve a structured series of steps that enable efficient and effective PT. First, researchers employ popular scanners such as Nessus and OpenVAS to scan for vulnerabilities in the target network. The obtained scan information about the availability and accessibility of network services (e.g., open ports, protocols) and vulnerabilities within those services are then generated into an attack graph. This graph creates a visual representation of potential attack paths, relationships between different elements of a network, “and dependencies between the vulnerabilities [15]." Essential information from the attack graph is then converted into a structured format, known as a State Graph, and passed to the RL algorithm for further analysis.

The State Graph represents how privileges transition within the network and if a specific vulnerability leads to an exploit. When a vulnerability is discovered and linked to an exploit, certain attributes such as the CVSS and AC are extracted and saved for future reference. The reward value is determined by the vulnerability's CVSS score, where higher severity vulnerabilities, with a more critical potential impact if exploited, earn a higher reward. This information is vital for calculating exploit success probabilities as it is used to define and build the RL algorithm through parameters including the state of user privilege, actions, transition probability, reward values, and the agent’s decision policy [15]. After confirming the success of their exploits through log analysis, the state graph and any relevant threat information are generated into an attack plan.

Finally, after the attack plan is generated, a Python wrapper for the Metasploit framework is used to validate its effectiveness. If vulnerabilities and weaknesses are found in the target organization's network or systems, the findings are used to recommend actions to improve the security of the organization. These actions may include applying patches or making changes to the network based on the vulnerabilities and weaknesses discovered during the test. Once these changes are implemented, the attack graph, which represents the network's vulnerabilities and potential attack paths, can be updated to reflect the new security measures. The system can then be retested to ensure that the implemented changes have effectively addressed the identified security issues and that the network's security posture has improved. This cyclical process of testing, improving, and retesting to enhance the organization's security is vital in cybersecurity, as it ensures that security measures remain robust, modern, and adaptive to evolving threats.

#### Main Findings

Overall, the main findings of the article emphasize that the ASAP framework, with its use of RL and attack graphs, offers a more efficient and effective approach to PT. It not only reduces the manual effort and time required, but it also reveals previously undiscovered attack paths that manual testing might miss, ultimately improving the overall security assessment process.

A case study involving the PT of an enterprise network with an industrial control system and IoT devices was conducted. The network consisted of 16 hosts distributed across three networks and offered a mix of Windows and Linux systems. The goal was to compromise email information by exploiting vulnerabilities on the SMTP service and infiltrating the IoT subsystem through a vulnerability in the gateway machine. The study investigated the effect of changes in the Discount Factor (DF) on determining the degree of significance assigned to future rewards in the decision-making process. Values closer to 0 prioritize immediate rewards, while values closer to 1 prioritize future rewards. The researchers also explored variations in batch size (BS) to explore the number of interactions the AI system uses to learn and improve its policy. These variations were analyzed to assess their influence on the RL agent's ability to make effective decisions.

The case study findings revealed that the DQN algorithm reached an effective solution quickly for different variations in DF, with the optimal value being around 0.8. Higher DF variations, 0.9 and 0.99, caused the agent to take more time to learn and make decisions as it required more time to explore each potential future outcome [15]. Similarly, reward value diminishes considerably the more the agent prioritizes long-term rewards, regardless of the number of interactions the system used to learn and improve. The agent's reward value was highest for the optimal DF around 0.8, especially with a BS of 16. The study showed that larger BS, such as 32 or 64, caused the AI's performance to decline [15]. However, researchers note that this observation might only hold true for a small network due to the increased complexity and scale of large networks, which can lead to different dynamics in the learning process for the AI system.

Researchers also conducted a scalability experiment on a simulated flat network comprising 300 hosts and three vulnerabilities. This experiment aimed to highlight the framework's scalability in situations where determining the balance between exploiting actions that appear promising based on its current knowledge and exploring new actions to discover potentially better strategies is challenging. Pentesters often experience this challenge in real-world situations, where they must decide how to allocate their resources, time, and efforts effectively.

In such situations, the experiment demonstrated the framework's ability to provide an attack plan within a short timeframe, about 70 seconds, when BS and DF parameters were set to their optimal values. This time frame is notably faster than research that utilized autonomous methods for PT, where a similar process took approximately 300 seconds (5 minutes) to perform on a network with only seven hosts. Even when the framework faced challenging scenarios with extreme BS and DF parameter values, it consistently generated effective attack plans within approximately 350 to 400 seconds [15]. These results indicate that the tested framework is highly efficient and capable of providing rapid and effective attack plans in various scenarios, significantly outperforming traditional PT methods.

What sets the ASAP framework apart from manual methods is its distinct strategy for PT. Unlike traditional manual testing, AI-based approaches, like ASAP, prioritize exploiting certain vulnerabilities before others, resulting in more efficient and effective PT. This data-driven approach involves adapting to the characteristics of vulnerabilities within unique network environments. Sometimes, starting with less challenging vulnerabilities can lead to a more efficient overall PT. The ASAP framework's adaptability and its consideration of vulnerability characteristics make it an asset in the field of cybersecurity and offers a significant reduction in the time and effort required over traditional manual approaches.

#### Relevance to Your Course Content

This article is highly relevant to my course as it aligns with the central theme of harnessing AI techniques for offensive strategies in PT. In the article, an ASAP framework provides a practical example of various aspects touched on in my course, including AI-driven PT tools, RL, DQ, and real-world applications of AI in security assessments. This in-depth exploration forms the foundational knowledge for the course's focus on AI-driven PT techniques and demonstrates the efficiency and effectiveness of AI in identifying security vulnerabilities.

Furthermore, the article's application of DQN as a deep RL technique serves as a practical example of how ML models can be used for identifying vulnerabilities and threats. This integration aligns with my course's content, which covers the training of various ML models. The ASAP framework showcased in the article also highlights the scalability of AI-driven techniques, making it suitable for large-scale networks. It demonstrates the transformative power of AI in PT by uncovering hidden attack paths, offering valuable insights about AI's practical applications in identifying vulnerabilities, and optimizing security assessments.

# Incorporation of Findings into the Course

The insights derived from this literature review form a strong foundation for the development of the short seminar course focused on harnessing AI for PT. Notably, the literature highlight the growing significance of AI-driven PT tools, which encompass AI-enhanced vulnerability detection, network scanning, exploit selection and execution, and post-exploitation management. [11] and [5] crucial phases, thereby highlighting the imperative need to integrate a combination of demonstrations and practical exercises into the course to ensure a comprehensive understanding of their functionality and limitations.

Additionally, the literature review sheds light on the ethical dilemmas inherent when employing AI in cybersecurity, as discussed in [4]. This insight alludes to the importance of discussing responsible and ethical practices within the course. Thus, the course not only emphasizes technical proficiency, but also instills an awareness to the ethical considerations associated with PT. Since these ethical considerations hold great significance in the cybersecurity domain, it is important to address these issues early in the course.

Furthermore, the research emphasizes the criticality of executing each PT step methodically and meticulously. For example, in the initial step of Gathering Information, it is important to determine whether the investigation targets a fully or partially observed environment. Drawing from [1] and [5] it becomes evident that this knowledge significantly impacts the choice of exploit techniques. These understandings will be harnessed throughout the course to foster a deeper comprehension of the intricacies incapsulating each PT phase, ranging from gathering information to exploiting vulnerabilities. and will provide the basis for structuring the course modules according to the PT phases to ensure a comprehensive and systematic approach to PT education.

These discoveries, which acknowledge the growing demand for AI-driven PT tools, ethical considerations, the imperative for a comprehensive understanding of PT steps and techniques, and the need for hands-on experiences, provide valuable preliminary course concepts. Beginning with Lecture 1, the course's introduction aims to provide a comprehensive overview of PT and its distinct phases. Drawing insights from [4], this lecture explores the significance of incorporating AI, ML, and deep learning into PT practices. This includes addressing constantly evolving threats, navigating modern network complexities such as Internet of Things (IoT), the cloud, remote work, and the limitations of traditional PT methods that AI aims to overcome.

After establishing this foundation, the focus for Lecture 2 shifts to the essential first step of PT – Gathering Information. Before exploring AI's role in revolutionizing this phase, the lecture outlines the definition, objectives, and importance of gathering information within the context of PT. To provide a practical perspective, the lecture then navigates through the intricacies of traditional information gathering methods, in order to shed light on their functionality and constraints. This exploration serves as a precursor to the subsequent analysis on the transformative potential of AI to revolutionize the entire process of collecting, analyzing, and interpreting data. It draws inspiration from [1] to not only explore the limitations of existing automated data acquisition techniques, but discuss modern Machine Learning (ML) techniques, including IAPTS, which challenge these limitations. Supported by modern research and in-depth research questions, this lecture aims to explore novel avenues for achieving more effective and dynamic information gathering in the field of Penetration Testing.

The next two lectures continue the exploration of the Splunk's PT phases and focus on Scanning. To provide a comprehensive understanding, the manual and AI concepts for this phase are dissected across two videos. While this section will cast a wide net, there will be a particular focus on GyoiThon and similar tools, following the research from [10]. The initial video, Lecture 3-1, provides a foundational understanding of scanning by delving into the core aspects and commonly used traditional manual methods. This includes a comprehensive examination of the objectives, significance, and manual tools tailored for each scanning type: Network, Vulnerability, and Web Application. By incorporating practical activities using TryHackMe exercises, such as Nmap Live Host Discovery and OWASP Juice Shop, Lecture 3-1 provides tangible experience to reinforce comprehension of manual scanning methodologies. This direct engagement offers a firsthand look into the limitations inherent in these manual techniques and offers a nuanced perspective on the challenges associated with traditional scanning approaches.

Having examined the limitations of traditional scanning methodologies, Lecture 3-2 redirects its focus toward the integration of AI in this phase. This segment begins by providing a definition of AI within the context of cybersecurity and delves into its pivotal role in addressing traditional scanning limitations. It then explores the advantages that AI offers to scanning, by highlighting its potential as a substantial improvement over manual approaches across each scanning type. Throughout this analysis of Network, Vulnerability, and Web Application scanning, the lecture offers practical insights into specific AI tools such as Dark Trace, Shodan, AppSpider, and others. Examining the advantages presented by these AI tools necessitates consideration of the implementation complexity. This complexity prompts a cautious awareness and compels a thorough and critical assessment of common challenges associated with integrating AI into scanning. To ensure a comprehensive understanding of the integration process, this evaluation considers factors such as ethics, biases, and technical hurdles. Then, to prompt participants to delve deeper into these ethical challenges, the lecture concludes with a thought-provoking research question on accountability in scenarios where an AI tool autonomously exploits vulnerabilities. This scenario challenges the core principle of PT and prompts reflection on responsible AI use and mechanisms to ensure clear lines of responsibility.

The subsequent phase in the Penetration Testing process, Exploitation, is similarly explored across two lectures, 4-1 and 4-2. This dual approach aims to provide a nuanced understanding of both manual methods and intelligent automation, by allowing participants to delve into the unique advantages and obstacles posed by each. Lecture 4-1 thoroughly analyzes the objectives and importance of the exploitation phase, by delving into traditional methods like SQL Injection, XSS, and Buffer Overflows. Then, to bridge the gap between manual methods and automation, it explores Metasploit, a tool that occupies an intermediary position in this spectrum. While Metasploit incorporates automation, it still requires a considerable level of manual intervention in terms of customization, interpretation, and adaptation. This dual nature positions Metasploit as a pivotal subject of exploration. Recent significant advancements in AI research, notably highlighted by [15] [5], showcase the evolving landscape where ML is integrated with Metasploit, thereby diminishing the reliance on manual processes. These developments emphasize how AI holds the potential to not only replace current manual tools but also seamlessly integrate with them to enhance overall capabilities. This exemplifies the potential for a harmonious coexistence between traditional manual methodologies and advanced AI-driven approaches in the field of Penetration Testing.

After discussing the limitations of these manual methods in Lecture 4-1, including challenges such as lack of adaptability to evolving threats, complexity, time consumption, and resource intensity, Lecture 4-2 pivots to the exploration of AI in the Exploitation phase. Guided by insights presented in [5], this lecture explores the practical implementation of AI and ML to overcome these limitations and enhance the identification and exploitation of vulnerabilities. This discussion extends to ML for exploit development, by introducing concepts such as RL and the Q-learning algorithm. To better illustrate these concepts, the lecture incorporates examples of real-world AI-enhanced exploitation, featuring tools such as the Social Engineering Toolkit (SET), DeepExploit, and various Python Libraries [16] [17]. The lecture not only focuses on existing AI-enhanced tools but also explores emerging trends and technologies in AI for exploitation. By leveraging novel research in the field, it demonstrates how integrating ML into powerful automated tools, such as Metaspolit, represents an important step toward introducing genuine AI capabilities in PT [5]. Given the continuous advancements and breakthroughs in AI, the lecture concludes by prompting participants to consider how these technologies can effectively handle novel and unexpected attack scenarios. It challenges them to consider the intricacies that arise from diverging threat landscapes and contemplate which patterns AI should prioritize to mitigate these challenges.

The concluding segment of the course delves into the post-exploitation phases of PT by leveraging insights from [13] [17]. This section strategically combines this cutting-edge research with the accumulated knowledge from preceding sections to comprehensively tackle the final PT phases. Similar to previous sections, the manual and AI concepts are dissected across two videos, Lectures 5-1 and 5-2, to provide a more thorough understanding. In the initial video, Lecture 5-1 focuses on traditional manual methods for the last phases of penetration testing: Maintaining Access, Covering Tracks, and Reporting. It begins by providing an overview of preceding phases in order to highlighting their relevance and importance to the post-exploitation process. The lecture then delves into each phase's objectives, importance, and practical tools, including netcat, meterpreter, and various password attack techniques. Participants receive detailed guidance on covering tracks through log manipulation and file concealment tools like AuditPol and WinZapper, concluding with reporting processes and tools such as Dradis and Faraday IDE. Practical exercises are then presented to further reinforce comprehension by offering hands-on experiences in maintaining access and covering tracks.

Building on this manual foundation, Lecture 5-2 shifts to the integration of AI in post-exploitation, by addressing the limitations of manual approaches, appreciating the advantages of AI integration, and acknowledging the current scarcity of research in the domain. As a result of this limited research, this lecture embraces a theoretical approach focuses more on theoretical implantation and explores potential future trajectories inspired by [13] and [17]. Referencing these novel studies, the lecture delves into advanced AI techniques, emphasizing RL, deep learning, and the Actor-Critic method, specifically A2C. The session then concludes by prompting participants to consider the effective deployment of RL agents in future research, and encourages exploration of insights, methodologies, and potential limitations. This theoretical exploration aims to ignite interest and innovation in an underexplored area of AI and PT.

Structuring the course modules in this manner will provide a comprehensive understanding of AI in PT by covering the entire PT process, from information gathering to post-exploitation. This approach achieves several objectives: it introduces essential foundational knowledge, explores real-world applications, and ignites a curiosity for the evolving landscape of AI in cybersecurity. This dual-perspective approach for the later phases mirrors the natural progression from manual to AI-driven methods and instills a forward-looking perspective to anticipate the transformative potential that AI holds for each phase in PT process.

# Course Overview

#### Introduction

Through a curriculum structured around the PT phases, the following lectures will explore the intersection of AI and PT by offering valuable insights and practical knowledge into the dynamic field of cybersecurity. With the central theme of 'Harnessing Artificial Intelligence for Penetration Testing,' topics discussed throughout these phases range from AI-driven PT tools, ML, RL, and advanced deep learning techniques. Lectures will provide an assortment of practical demonstrations, theoretical research questions, and hands-on exercises to illustrate how AI enhances manual offensive strategies and highlight AI's transformative power in the world of cybersecurity.

Lecture 1 establishes the groundwork by introducing fundamental AI concepts in the context of PT, paving the way for the exploration of the initial PT phase, Gathering Information, in Lecture 2. Building on this, Lectures 3-1 and 3-2 delve into the Scanning phase and emphasize AI's adaptability. Moving forward, Lectures 4-1 and 4-2 demonstrate how AI and ML can effectively identify and exploit vulnerabilities. Concluding the course, Lectures 5-1 and 5-2 explore the post-exploitation phases, navigating a limited research landscape that steers the discussion toward a more theoretical approach and opens doors to innovation in the future of AI. Each of these lectures incorporates practical hands-on labs and thought-provoking research questions, to offer participants a unique opportunity to engage with theoretical implications and explore potential advancements in the field.

#### Course Table of Contents

* 1. Introduction:   
  <https://clipchamp.com/watch/8Ga8P3d49mW>
* 2. Gathering Information:  
  <https://clipchamp.com/watch/zNgpqzh6tdW>
* 3-1. Scanning (*Manual*):  
  <https://clipchamp.com/watch/DkiiipNAf27>
* 3-2. Scanning (*AI*):  
  <https://clipchamp.com/watch/ClG4PTorwYu>
* 4-1. Exploitation (*Manual*):  
  <https://clipchamp.com/watch/VoIN8MhSTYe>
* 4-2. Exploit (*AI*):  
  <https://clipchamp.com/watch/epDzlpNNYsq>
* 5-1. Post-Exploit (*Manual*):  
  <https://clipchamp.com/watch/tZg1U3Z4nM5>
* 5-2. Post-Exploit (*AI*):  
  <https://clipchamp.com/watch/0ycz6fR8jbX>

#### Conclusion

In culmination, this course serves as a comprehensive journey into the symbiotic relationship between AI and cybersecurity. By laying the foundational understanding of AI and its diverse domains, these lectures go beyond theory and immerse participants in the practical applications across key penetration testing phases. From information gathering, scanning, exploitation, and the often-overlooked post-exploitation phase, each lecture meticulously guides participants through a genuine PT experience. This structured approach not only imparts theoretical knowledge but also provides a first-hand view of how AI operates in the context of a genuine assessment. By witnessing the natural progression of a PT, participants gain invaluable insights into the application of AI at every crucial juncture. Thus, the course ensures that the theoretical foundations are not just understood but are actively integrated into the practical realm of authentic cybersecurity challenges.

#### References and Further Reading

* AI and ML: [18], [19], [20], [21]
* Vulnerability Assessments and PT: [22], [2], [23], [24], [25], [26]
* PT+ (Automation or AI): [16], [27], [28], [29]
* RL, Q learning, and Deep Q Learning: [30], [31], [32]
* Tools and Walkthroughs: [33], [34], [35]

# Key Components of the Newly Developed Course

The short seminar course on Harnessing Artificial Intelligence (AI) for Penetration Testing (PT) is designed with several key components to provide participants with a thorough and applicable learning experience. The primary learning objectives of the course are to equip students with a deep understanding of AI's role in PT and to develop practical proficiency in deploying various prominent tools. Tailored for undergraduate-level students, early-career professionals, and digital enthusiasts aspiring to fortify their cybersecurity skills, the course is geared towards individuals possessing a foundational understanding of cybersecurity, albeit with limited experience in AI or PT.

The course structure is meticulously organized to ensure a systematic progression of knowledge. Beginning with an introduction to AI in PT, participants will examine the fundamentals before gradually transitioning to more advanced topics, similar to the natural flow of a PT. From information gathering, scanning, exploitation, to the under-utilized post-exploitation phase, each lecture builds a bridge between theoretical knowledge and hands-on experience to provide participants with a firsthand view of how AI functions within the context of a genuine PT.

Assessment strategies within the course comprise a blend of hands-on exercises to develop specific skills, demonstrate knowledge, and reflective questions designed to stimulate discourse and improve problem-solving capabilities. This multifaceted approach ensures participants not only grasp theoretical concepts but also gain practical experience to foster an extensive skill set crucial for the dynamic landscape of cybersecurity.

Supplementary materials for the course encompass research papers and access to the PT tools discussed during the sessions. To augment the learning journey, participants will be provided with links to pertinent online challenges, hosted on the platform TryHackMe. These exercises will be directly aligned with the course content and enable students to apply their newfound knowledge in authentic, real-world scenarios. Throughout the course's progression, additional materials may be introduced to ensure participants have access to the most up-to-date, comprehensive, and valuable resources available. As the course culminates, participants will not only emerge with a nuanced understanding of AI in PT but also possess the practical skills to navigate the field.

# Reflection

My experience with the Udemy-like online short course on Harnessing Artificial Intelligence for Penetration Testing has been equally a challenging and rewarding experience. One of the most significant hurdles I faced was overcoming my struggle to speak seamlessly during recording sessions. Even with meticulously prepared notes, I often stumbled over my words and lacked the fluidity I had originally envisioned for my course. I could spend hours studying a topic, but as soon as I began recording, my mind would go blank. This is something I have struggled with during interviews in the past, but to experience it when alone was one of the most internally frustrating situations I have dealt with.

In an attempt to overcome this challenge, I decided to adopt a scripted approach. However, this brought forth its own set of challenges, as reading made my voice dull and not engaging. Despite numerous re-recordings, I acknowledge there's still room for improvement, but I have started to find a better balance between being engaging and informative. This process pushed me mentally and emotionally, and I believe my communication and research skills have significantly improved over this semester.

Also, as this was my first attempt at instructional content creation, I gained a profound appreciation for the intricacies involved in teaching. Underestimating the amount of work required to deliver engaging and informative lectures, I found this project to be one of the most academically demanding endeavors I have undertaken. Despite organizing my course schedule meticulously in OneNote, where I spent weeks searching for research to support my lectures, I was surprised with how difficult it was to plan and organize an entire course.

Although counterintuitive, I do hope for future opportunities to reattempt the creation of a short course. If I could do it over again, I would focus my research on quality over quantity, prioritize time management, and embrace imperfections on audio. Despite the challenges, I am grateful for this experience as I believe this semester has contributed significantly to my overall learning and personal development.

# Conclusion

In conclusion, this short seminar course, titled "Harnessing Artificial Intelligence for Penetration Testing," is designed to provide a comprehensive exploration of the integration of AI into the field of PT. Drawing upon the insights collected from the literature review, this course aims to equip participants with both the theoretical and practical knowledge necessary to navigate the complex and evolving landscape of cybersecurity. Key insights from the literature review have shaped the course content, from the structure of the course modules to an emphasis on the importance of AI-driven PT tools, ethical standards, comprehensive understanding, and hands-on experience in PT techniques. These insights highlight the critical role of AI in bolstering cybersecurity defenses, reducing human errors, and addressing the critical shortage of cybersecurity personnel. By offering a tailored curriculum structured around the PT phases, this course prepares students to not only understand the current AI landscape but also envision its future possibilities.

The potential impact of this course focuses on circulating knowledge regarding the critical need for additional research on AI-driven PT tools during each phase of PT. As students delve into the practical exercises and engage in thoughtful discussions, they not only enhance their technical expertise but also develop critical thinking and problem-solving skills. This course paves the way for innovative approaches to cybersecurity, bridging the gap between AI and PT, and empowering individuals to navigate the evolving threat landscape. Ultimately, it serves as a catalyst for a more secure digital world, where AI is not just wielded by criminal hackers but harnessed by cybersecurity professionals who stay one step ahead, outsmarting and countering these threats with cutting-edge techniques to ensure the safety and resilience of our digital future.

# References

[1] M. C. Ghanem and T. M. Chen, "Reinforcement Learning for Intelligent Penetration Testing," in Second World Conference on Smart Trends in Systems, Security and Sustainability, London, 2018.

[2] H. M. Z. A. Shebli and B. D. Beheshti, "A study on penetration testing process and tools," in Long Island Systems, Applications and Technology Conference (LISAT), Farmingdale, 2018.

[3] S. Watts , Penetration Testing: Practical Introduction & Tutorials, 2022.

[4] A. Happe and J. Cito, "Getting pwn’d by AI: Penetration Testing with Large Language Models," in European Software Engineering Conference and Symposium on the Foundations of Software Engineering, San Francisco, 2023.

[5] A. AlMajali, L. Al-Abed, R. Mutleq, Z. Samamah, A. A. Shhadeh, B. J. Mohd and K. M. Ahmad Yousef, "Vulnerability Exploitation Using Reinforcement Learning," in Jordan International Joint Conference on Electrical Engineering and Information Technology, Amman, 2023.

[6] C. Kidd, What Is Splunk & What Does It Do? An Introduction To Splunk, 2022.

[7] M. Pogla, Auto-GPT vs ChatGPT: How Do They Differ and Everything You Need To Know, 2023.

[8] G. Dheda, Auto-GPT vs AgentGPT: Understanding the Differences, 2023.

[9] M. T. Spaan and N. Vlassis, Perseus: Randomized Point-based Value Iteration for POMDPs, vol. 24, 2005, p. 26.

[10] R. S. Jagamogan, S. A. Ismail, N. H. Hassan and H. Aba, "Penetration Testing Procedure using Machine Learning," in International Conference on Smart Sensors and Application (ICSSA), Kuala Lumpur, 2022.

[11] gyoisamurai, GyoiThon: Next generation Penetration Test Tool, 2021.

[12] Q. T. Luu, Q-Learning vs. Deep Q-Learning vs. Deep Q-Network, 2023.

[13] S. Chaudhary, A. O’Brien and S. Xu, "Automated Post-Breach Penetration Testing through Reinforcement Learning," in Conference on Communications and Network Security (CNS), Avignon, 2020.

[14] TensorFlow, Introduction to TensorFlow.

[15] A. Chowdhary, D. Huang, J. S. Mahendran, D. Romo, Y. Deng and A. Sabur, "Autonomous Security Analysis and Penetration Testing," in 16th International Conference on Mobility, Sensing and Networking (MSN), Tokyo, 2020.

[16] Y. Stefinko, A. Piskozub and R. Banakh, "Manual and automated penetration testing. Benefits and drawbacks. Modern tendency," International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), pp. 488-491, February 2013.

[17] R. Maeda and M. Mimura, "Automating post-exploitation with deep reinforcement learning," Computers & Security, vol. 100, pp. 102-108, January 2021.

[18] N. Duggal, What is Artificial Intelligence: Types, History, and Future, 2023.

[19] N. Kühl, M. Schemmer, . M. Goutier and G. Sat, Artificial intelligence and machine learning, 2022, pp. 2235-2244.

[20] C. Chebbi, Mastering Machine Learning for Penetration Testing, Packt Publishing, 2018.

[21] E. Tsukerman, Machine Learning for Cybersecurity Cookbook, J. Cummings, Ed., Birmingham: Packt Publishing, 2019.

[22] S. Shah and B. M. Mehtre , "An overview of vulnerability assessment and penetration testing techniques," Journal of Computer Virology and Hacking Techniques, pp. 27-49, 18 November 2014.

[23] P. Engebretson, The Basics of Hacking and Penetration Testing, A. Ward, Ed., Waltham, MA : Elsevier Inc, 2011.

[24] M. N. Zakaria, P. A. Phin, N. Mohmad, S. A. Ismail, M. N. Kama and O. Yusop, "A Review of Standardization for Penetration Testing Reports and Documents," in International Conference on Research and Innovation in Information Systems (ICRIIS), 2013.

[25] A. A. Alghamdi, "Effective Penetration Testing Report Writing," in International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), 2021.

[26] S.-P. Oriyano, Penetration Testing Essentials, Sybex, 2016.

[27] F. Abu-Dabaseh and E. Alshammari , "Automated Penetration Testing: An Overview," in Computer Science & Information Technology - Computer Science Conference Proceedings (CS & IT- CSCP), Amman, 2018.

[28] J. Hoffmann, "Simulated Penetration Testing:From “Dijkstra” to “Turing Test++”," in International Conference on Automated Planning and Scheduling364, Saarbrücken, 2015.

[29] N. Singh, V. Meherhomji and B. R. Chandavarkar, "Automated versus Manual Approach of Web Application Penetration Testing," International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-6, July 2020.

[30] D. Pandey and P. Pandey, "Approximate Q-Learning: An Introduction," in International Conference on Machine Learning and Computing, Bangalore, 2010.

[31] H. Van Hasselt, A. Guez and D. Silver, "Deep Reinforcement Learning with Double Q-learning," in AAAI Conference on Artificial Intelligence, Phoenix, 2016.

[32] K. Tran, A. Akella, M. Standen, J. Kim, D. Bowman, T. Richer and C.-T. Lin, "Deep Hierarchical Reinforcement Agents for Automated Penetration Testing," in International Workshop on Adaptive Cyber Defense, Sydney, 2021.

[33] ZION3R, GyoiThon - A Growing Penetration Test Tool Using Machine Learning, 2018.

[34] D. Foti, Metasploit: Exploitation Walkthrough TryHackMe, 2022.

[35] M. Shivanandhan, Metasploit — A Walkthrough Of The Powerful Exploitation Framework, 2020.

[36] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," Journal of Industrial Information Integration, vol. 23, September 2021.

[37] G. Stone, D. Talbert and W. Eberle, "Using AI/Machine Learning for Reconnaissance Activities During Network Penetration Testing," in International Conference on Cyber Warfare and Security, 2021.

[38] S. V. N. Parasram, A. Samm, D. Boodoo, G. Johansen, L. Allen, T. Heriyanto and S. Ali, Kali Linux 2018: Assuring Security by Penetration Testing, 4th ed., Packt Publishing, 2018.

[39] J. M. Ortega, Mastering Python for Networking and Security, 2 ed., V. Boricha, Ed., Birmingham: Packt Publishing, 2020.

[40] J. Matherly, Complete Guide to Shodan, Leanpub, 2016.

[41] N. Kolakowski, Splunk and Cloudera Alliance Hints at New Big Data Landscape, 2013.

[42] P. Cooke, "From The Machine Learning Region to The Deep Learning Region: Tesla, DarkTrace and DeepMind as Internationalised Local to Global Cluster Firms.," 2018.

[43] Z. Ali, F. Hussain, S. Ghazanfar, M. Husnain, S. Zahid and G. A. Shah, "A Generic Machine Learning Approach for IoT Device Identification," in International Conference on Cyber Warfare and Security (ICCWS), Islamabad, 2021.

[44] H. Al-Alami, A. Hadi and H. Al-Bahadili, "Vulnerability scanning of IoT devices in Jordan using Shodan," in 2nd International Conference on the Applications of Information Technology in Developing Renewable Energy Processes & Systems (IT-DREPS), Amman, 2017.