Introdução

A análise de artigos científicos reveste-se de uma importância ímpar no desenvolvimento do pensamento crítico e no aprofundamento dos conhecimentos teóricos e práticos adquiridos ao longo da formação académica. Este processo analítico não só permite aos estudantes desenvolver uma capacidade crítica apurada, como também os habilita a identificar lacunas e oportunidades de investigação no campo dos sistemas de apoio à decisão e da Aprendizagem Máquina, áreas que se revelam cada vez mais cruciais na resolução de problemas complexos e dinâmicos.

No contexto da disciplina em questão, é fundamental que os estudantes se dediquem à análise de artigos que explorem, de forma direta, a intersecção entre técnicas de Data Mining e os desafios emergentes no âmbito da Cibersegurança. A relevância desta abordagem prende-se com o facto de que a utilização de algoritmos avançados de extração de dados e a aplicação de métodos de análise preditiva podem contribuir significativamente para a deteção precoce e a prevenção de ameaças cibernéticas, proporcionando um contributo robusto para a segurança dos sistemas de informação.

A ênfase deve ser colocada na análise crítica de estudos de caso que ilustrem a aplicação prática destas técnicas, permitindo uma compreensão aprofundada dos contextos reais onde as abordagens de Data Mining têm sido implementadas com sucesso para mitigar riscos e vulnerabilidades. Adicionalmente, a investigação de novas abordagens algorítmicas constitui um campo fértil para a inovação, encorajando os estudantes a explorarem e a desenvolverem métodos que se adaptam às rápidas transformações do ambiente digital.

Por fim, a comparação entre diferentes métodos de deteção de ameaças, realizada através de revisões sistemáticas e análises comparativas, oferece uma perspetiva abrangente que facilita a identificação dos pontos fortes e das limitações de cada abordagem. Este exercício de comparação não só enriquece a base de conhecimentos dos estudantes, como também os prepara para enfrentar desafios reais na implementação e na optimização de sistemas de apoio à decisão, onde a integração de técnicas de Aprendizagem Máquina se revela cada vez mais indispensável.

Em suma, a análise detalhada de literatura científica neste contexto não só fomenta uma formação académica sólida, mas também contribui para o desenvolvimento de competências essenciais que, a médio e longo prazo, terão um impacto significativo na capacidade dos profissionais em integrar e aplicar soluções inovadoras na área da Cibersegurança.

Categoria de Artigos Recomendados

Categoria	Descrição	Exemplos de Temas
Aplicações de Data Mining em Cibersegurança	Artigos que demonstrem o uso prático de técnicas de Data Mining (classificação, clustering, regras de associação, etc.) em problemas reais de cibersegurança.	- Uso de Random Forest para deteção de ataques DDoS- Implementação de Redes Neurais para deteção de malware- Comparação de técnicas de clustering na identificação de tráfego malicioso
Sistemas de Deteção de Intrusões (IDS) e IPS	Estudos sobre a aplicação de algoritmos de Data Mining e Machine Learning em IDS/IPS, com análise de desempenho e eficácia.	- Comparação entre IDS baseados em assinaturas e em anomalias- Análise de performance de IDS com KDD Cup 99 ou CICIDS2017
Estudos de Caso com Datasets Públicos	Análises empíricas utilizando datasets conhecidos em cibersegurança.	- Avaliação do dataset CICIDS2017 em deteção de ataques- Estudo comparativo entre os datasets KDD Cup 99 e UNSW-NB15
Uso de Deep Learning em Cibersegurança	Artigos que explorem a aplicação de CNN, LSTM ou Redes Neurais em deteção de malware, análise de logs e ameaças internas.	- Redes neurais convolucionais para classificação de malware- LSTM para deteção de padrões em ataques persistentes (APT)
Ameaças Internas e Análise Comportamental (Insider Threats)	Trabalhos que empreguem técnicas de Data Mining (como análise de séries temporais e NLP) para identificar comportamentos suspeitos.	- Análise de e-mails com NLP para deteção de phishing- Análise comportamental de acessos com UEBA (User and Entity Behavior Analytics)

Fontes Recomendadas

- **arXiv:** Repositório de acesso livre que reúne preprints de artigos científicos em diversas áreas (tais como Física, Matemática, Informática, entre outras), permitindo o acesso antecipado a inovações na investigação.
- **B-On:** Plataforma que agrega conteúdos científicos provenientes de bibliotecas portuguesas, oferecendo acesso a publicações nacionais e internacionais, o que facilita a pesquisa e a consulta a estudos relevantes.
- **Google Scholar:** Ferramenta de pesquisa académica que consolida a literatura científica disponível em múltiplas bases de dados, permitindo a identificação rápida de estudos e a verificação de citações e referências.
- **Scopus:** Base de dados multidisciplinar de elevada credibilidade, que disponibiliza uma vasta coleção de artigos científicos e oferece ferramentas para análises bibliométricas e a identificação de tendências de investigação.
- **Web of Science:** Plataforma reconhecida internacionalmente para a pesquisa científica, permitindo o acesso a publicações de alta qualidade e a realização de análises detalhadas do impacto dos estudos através de métricas específicas.
- **IEEE Xplore:** Biblioteca digital especializada em engenharia, informática e cibersegurança, que disponibiliza artigos, atas de conferências e normas técnicas, sendo essencial para a pesquisa em sistemas de apoio à decisão e Aprendizagem Máquina.
- **ACM Digital Library:** Repositório de publicações científicas e técnicas na área de computação e tecnologia da informação, que contribui para a investigação em Aprendizagem Máquina, cibersegurança e outras áreas correlatas.
- **SciELO**: Biblioteca eletrónica focada na divulgação de revistas científicas de países lusófonos e latino-americanos, promovendo o acesso a estudos em língua portuguesa e espanhola e contribuindo para a disseminação do conhecimento regional.

Critérios para Escolha do Artigo

A escolha de um artigo científico de qualidade deve basear-se num conjunto de critérios rigorosos que permitam uma compreensão aprofundada do tema em estudo, bem como a validação da contribuição científica apresentada. Para tal, é fundamental que o artigo disponha de uma metodologia detalhada, que inclua uma descrição pormenorizada dos datasets utilizados, esclarecendo a sua origem, dimensão e características relevantes. Esta transparência não só possibilita a replicabilidade do estudo, como também permite uma avaliação crítica da adequação dos dados face ao problema investigado.

Além disso, é imperativo que o artigo descreva de forma clara os algoritmos e as técnicas implementadas, justificando a escolha dos mesmos e detalhando os parâmetros e configurações utilizados durante a sua aplicação. Esta explicitação metodológica contribui para evidenciar as vantagens e limitações das abordagens propostas e, sempre que possível, compara-as com métodos alternativos, enriquecendo assim a discussão científica.

Outro aspeto determinante reside na apresentação de resultados experimentais robustos. O artigo deve demonstrar a eficácia dos métodos através de dados quantitativos e, se aplicável, qualitativos, que permitam verificar o desempenho dos algoritmos. A inclusão de comparações entre diferentes abordagens, apoiada na utilização de elementos visuais como gráficos, tabelas e diagramas, facilita a interpretação dos resultados e realça as implicações práticas do estudo, promovendo uma análise crítica aprofundada.

Por fim, o enquadramento referencial constitui um critério essencial. O artigo deve situar a investigação no contexto do estado da arte, recorrendo a referências relevantes e atualizadas que evidenciem os trabalhos pioneiros e os avanços recentes na área. A utilização de fontes de elevada credibilidade garante que o estudo se insira num panorama científico sólido e coerente com as evoluções no domínio da cibersegurança, do Data Mining e da Aprendizagem Máquina.

Em resumo, a seleção de um artigo científico de excelência passa pela verificação da existência de uma metodologia detalhada, a apresentação de resultados experimentais consistentes e a utilização de referências pertinentes, o que, em conjunto, promove o desenvolvimento do pensamento crítico e contribui para a construção de uma base de conhecimento robusta e atualizada.

Estrutura para Análise do Artigo

A análise deverá ser apresentada numa folha A4, impressa em frente e verso, e deve contemplar os seguintes elementos essenciais, organizados de forma clara e estruturada.

O primeiro elemento a incluir é a **Identificação do Artigo**, que deve conter o título, os autores, o nome da conferência ou revista onde o estudo foi publicado e o respetivo ano de publicação. Esta informação é fundamental para situar o artigo no seu contexto académico e permitir a sua fácil localização e referência.

Na sequência, deve ser elaborado um **Resumo do Artigo**, que consiste numa breve descrição do objetivo do estudo e da principal contribuição apresentada. Este resumo tem o propósito de fornecer uma visão global do conteúdo, permitindo ao leitor compreender rapidamente a essência da investigação.

O terceiro elemento é a **Metodologia**, onde se deverá detalhar, de forma pormenorizada, as técnicas de Data Mining utilizadas. É imperativo que se descrevam os datasets empregados (incluindo a sua origem, dimensão e características), os algoritmos implementados e as métricas de avaliação utilizadas para aferir o desempenho dos métodos aplicados. Esta secção deve ser suficientemente detalhada para possibilitar a replicabilidade do estudo.

A seguir, deve ser apresentada a secção de **Resultados e Discussão**, na qual se exponham as principais conclusões do estudo. Nesta parte, é importante evidenciar o desempenho dos modelos através de resultados experimentais e, quando possível, proceder à comparação entre diferentes abordagens. A utilização de elementos visuais, como gráficos e tabelas, pode facilitar a interpretação dos dados e reforçar a análise crítica dos resultados obtidos.

Posteriormente, a análise deverá contemplar as **Limitações** do estudo, onde se descrevam os desafios enfrentados durante a implementação dos métodos e se apontem as potenciais melhorias que poderiam ser adotadas em futuras investigações. Esta secção demonstra o rigor crítico e a consciência das restrições inerentes à investigação científica.

Por fim, a **Conclusão** deve sintetizar uma opinião crítica sobre a relevância do estudo e a sua aplicação prática, resumindo os contributos e sugerindo possíveis linhas de investigação futura. Complementarmente, a análise deve incluir uma secção de **Referências**, na qual se enumere todas as fontes citadas, assegurando a transparência e a credibilidade do trabalho apresentado.

Avaliação do artigo

A avaliação do artigo, no âmbito deste trabalho, incumbe de verificar não só a qualidade do conteúdo crítico e analítico, mas também a capacidade do aluno em comunicar eficazmente os resultados da sua análise. Assim, a avaliação será composta por duas componentes fundamentais: a análise escrita do artigo e a respetiva apresentação oral, ambas desempenhando um papel crucial na apreciação global do desempenho do aluno.

1. Análise Escrita do Artigo

A componente escrita deverá refletir um tratamento aprofundado do artigo selecionado, demonstrando um pensamento crítico e uma capacidade analítica robusta. Os seguintes aspetos serão considerados na avaliação:

- **Clareza e Objetividade**: O texto deve ser bem estruturado, coeso e redigido de forma clara e precisa, sem ambiguidades ou redundâncias.
- **Profundidade da Análise**: A capacidade de examinar criticamente as premissas, a metodologia, os resultados e as conclusões do artigo, evidenciando um domínio do tema.
- **Riqueza Argumentativa**: O aluno deve apresentar argumentos bem fundamentados, apoiando-se em referências académicas e exemplos pertinentes.
- **Rigor Metodológico**: A discussão deve demonstrar um entendimento sólido da abordagem metodológica do artigo, destacando eventuais limitações ou potenciais melhorias.

2. Apresentação Oral em Aula

Para além da análise escrita, a apresentação oral constitui um elemento imprescindível da avaliação, com uma duração máxima de **20 minutos**. Esta apresentação deverá ser estruturada de forma a garantir a máxima clareza e impacto na transmissão da informação. A avaliação incidirá nos seguintes aspetos:

2.1 Estrutura e Clareza da Exposição

- A apresentação deve ter um **início claro**, onde se contextualiza o artigo e se definem os objetivos da análise.

- O **desenvolvimento** deve enfatizar os pontos-chave do artigo, garantindo uma explicação acessível e bem fundamentada.

2.2 Qualidade da Análise e Interpretação

- A capacidade de identificar e explicar os aspetos mais relevantes do artigo.
- A profundidade na análise metodológica, destacando eventuais limitações ou implicações dos resultados.

2.3 Comunicação e Expressão Oral

- **Clareza e objetividade** na exposição das ideias.
- **Uso adequado de linguagem técnica e científica**, sem excessos de jargão desnecessário.
- **Postura e dicção**, garantindo que a comunicação seja fluída e envolvente.
- **Apoio visual** (slides ou outros materiais) bem organizado e complementar à apresentação.

3. Critérios de Avaliação

A avaliação final será ponderada com base nos seguintes critérios:

Critério	Subcritério	Peso (%)
Análise Escrita	Clareza, profundidade e estrutura do texto	30%
Apresentação Oral	Estrutura e clareza da exposição	20%
	Qualidade da análise e interpretação	20%
	Comunicação e expressão oral	15%
Interação e Debate	Resposta às questões e estímulo à participação	15%

Exemplos de Artigos Indicados

1. **Piskozub, A.** (2024). *Data mining for threat detection in Active Directory*. ResearchGate. https://www.researchgate.net/publication/388886846
2. **Kumar, P., Kushwaha, C., Sethi, D., Ghosh, D., & Gupta, P.** (2023). Investigating the performance of multivariate LSTM models to predict the occurrence of Distributed Denial of Service (DDoS) attack. *PloS ONE*, 18(7). https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0313930
3. **D'Antonio, S., & Uccello, F.** (2022). Artificial intelligence applications in healthcare security. In *Advances in Cybersecurity Technologies* (pp. 120-134). Springer. https://link.springer.com/chapter/10.1007/978-3-031-70775-9_9
4. **Guru, A., Gopal, A. V., & Bandarupalli, S. S. B.** (2023). Uncovering threats: Data mining techniques for cybersecurity. *Risk Assessment and Management Journal*, 15(2), 34-45. https://ramd.reapress.com/journal/article/view/52
5. **Wu, E. H. K., & Lin, Y. D.** (2021). TRACE: Relationship analysis and causal factor extraction in cyber threat intelligence reports. *IEEE Transactions on Dependable and Secure Computing*, 18(9), 874-889. https://ieeexplore.ieee.org/abstract/document/10851819
6. **Velasquez, J. D., Pant, M., Pan, J. S., & Snasel, V.** (2022). On the fuzzy entropy and the rankability of data. *SSRN Electronic Journal*. https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=5134638
7. **Song, D.** (2023). Penetration testing automation with inverse soft-Q learning: An imitation learning method. *DiVA Portal*. https://www.diva-portal.org/smash/record.jsf?pid=diva2:1935147
8. **Zhao, X., Leng, X., Wang, L., & Liu, Y.** (2023). Efficient anomaly detection in tabular cybersecurity data using large language models. *Scientific Reports*, 13(4), 567-578. https://www.nature.com/articles/s41598-025-88050-z
9. **Papoutsis, A., & Panagiotou, P.** (2020). AI-based holistic framework for cyber threat intelligence management. *IEEE Transactions on Information Forensics and Security*, 15(7), 1024-1039. https://ieeexplore.ieee.org/abstract/document/10851288

10. Bollmann, C. A., Tummala, M., & McEachen, J. C. (2021). Resilient real-time network anomaly detection using novel non-parametric statistical tests. *Computers & Security, 102, 102146*. *doi:https://doi.org/10.1016/j.cose.2020.10214*

11. Gibert, D., Mateu, C., Planes, J., & Marques-Silva, J. (2021). Auditing static machine learning anti-Malware tools against metamorphic attacks. *Computers & Security, 102, 102159*. *doi:https://doi.org/10.1016/j.cose.2020.102159*

12. Krumay, B., Bernroider, E. W. N., & Walser, R. (2018). *Evaluation of Cybersecurity Management Controls and Metrics of Critical Infrastructures: A Literature Review Considering the NIST Cybersecurity Framework, Cham*.

13. Lin, W.-C., Ke, S.-W., & Tsai, C.-F. (2015). CANN: An intrusion detection system based on combining cluster centers and nearest neighbors. *Knowledge-Based Systems, 78, 13-21*. *doi:https://doi.org/10.1016/j.knosys.2015.01.009*

14. Mitchell, R., & Chen, I.-R. (2014). A survey of intrusion detection techniques for cyber-physical systems. *ACM Comput. Surv., 46(4), Article 55*. *doi:10.1145/2542049*

15. Casas, P., Mazel, J., & Owezarski, P. (2012). Unsupervised Network Intrusion Detection Systems: Detecting the Unknown without Knowledge. *Computer Communications, 35(7), 772-783*. *doi:https://doi.org/10.1016/j.comcom.2012.01.016*

16. García-Teodoro, P., Díaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. Computers & Securit

17. Data Mining for Cyber Security, V.Chandois *et al.*, in Data Warehousing and Data Mining Techniques for Computer Security, Springer, 2006.

18. Data mining methods for anomaly detection KDD-2005 workshop report, Margineantu *et al.,* ACM SIGKDD Explorations Newsletter, Volume 7 Issue 2, December 2005.

19. On the efficacy of data mining for security applications, Ted E. Senator, ACM SIGKDD Workshop on CyberSecurity and Intelligence Informatics -CSI-KDD '09, 2009.

20. Metrics for mitigating cybersecurity threats to networks, IEEE Internet Computing, 14, 1, Jan-Fev 2010.
21. A Combined Fusion and Data Mining Framework for the Detection of Botnets, Kiayias *et al.*, Conference For Homeland Security, 2009. CATCH '09. Cybersecurity Applications & Technology, March 2009
22. A study of Spam Detection Algorithms on Social Media Networks, Jacob Soman Saini, International Conference on Computational Intelligence, Cyber Security, and Computational Models, Coimbatore, India, December 2013.
23. Comparative Study of Two- and Multi-Class-Classification-Based Detection of Malicious Executables Using Soft Computing Techniques on Exhaustive Feature Set. Shina Sheen, R. Karthik and R. Anitha; International Conference on Computational Intelligence, Cyber Security, and Computational Models, Coimbatore, India, December 2013
24. Botnets: A Study and Analysis, G. Kirubavathi and R. Anitha, International Conference on Computational Intelligence, Cyber Security, and Computational Models, Coimbatore, India, December 2013
25. The VoIP intrusion detection through a LVQ-based neural network, Zheng Lu ; Taoxin Peng, International Conference for Internet Technology and Secured Transactions, 2009. ICITST 2009.
26. Detection of applications within encrypted tunnels using packet size distributions, Mujtaba,G.,Parish, D.J., International Conference for Internet Technology and Secured Transactions, 2009. ICITST 2009.
27. Email classification: Solution with back propagation technique, Ayodele et al. International Conference for Internet Technology and Secured Transactions, 2009. ICITST 2009.
28. Malware detection using statistical analysis of byte-level file content, Tabish et al., CSI-KDD '09 Proceedings of the ACM SIGKDD Workshop on CyberSecurity and Intelligence Informatics, 2009