risti
Revista Ibérica de Sistemas e Tecnologias de Informação
Iberian Journal of Information Systems and Technologies

Recebido/Submission: 02/05/2023
Aceitação/Acceptance: 26/07/2023

Revisão sistemática da literatura sobre viés de género e aprendizagem máquina na segurança

Luís Malheiro[1], Fernando Bessa[2], João Reis[3], Luís Saraiva[4]

malheiro.lcr@gnr.pt; a15277@hotmail.com; joao.reis@ulusofona.pt; saraiva_luis@hotmail.com

[1] Centro de Investigação Desenvolvimento e Inovação da Academia Militar, Academia Militar, Instituto Universitário Militar, Rua Gomes Freire, 1669-203, Lisboa, Portugal

[2] Centro de Investigação de Estudos de Sociologia, Avenida Forças Armadas 16 - s2, 1600-082 Lisboa, Portugal

[3] Industrial Engineering and Management, Faculty of Engineering, Lusófona University and EIGeS, Campo Grande, 1749-024 Lisbon, Portugal

[4] Centro de Investigação do Instituto Universitário Militar, Instituto Universitário Militar, Rua de Pedrouços s/n, Lisboa, Portugal

Pages: 182-194

Resumo: O presente estudo foi desenvolvido com o propósito de analisar os desafios da utilização de dados com viés de género para a aprendizagem da máquina na segurança. A procura das limitações da aprendizagem máquina e o debate sobre o contágio por via de dados desequilibrados na fase do treino, nomeadamente na dimensão do género, foram baseados no protocolo Prisma. Depois de serem definidas as palavras-chave a procurar e os critérios de exclusão, a densificação da análise das 39 investigações selecionadas foi promovida por via da utilização do *VOSviewer,* do *Connected papers* e do *Treecloud*. A conjugação das técnicas e procedimentos anteriores possibilitou concluir que: (1) os vieses cognitivos podem existir na aprendizagem máquina devido ao recurso a dados desequilibrados para treino; (2) proporcionar um treino com dados equilibrados (também sobre o género) a uma rede aumenta a sua resiliência; (3) é possível introduzir algoritmos com atributos *gender sensitive* para mitigar vieses na aprendizagem máquina; (4) existe uma lacuna sobre investigações científicas relativamente a vieses de género no campo da segurança.

Palavras-chave: Viés de género; segurança; aprendizagem máquina; dados.

Systematic literature review on gender bias and machine learning in safety

Abstract: The present study was developed to analyze the challenges of using gender-biased data for machine learning in security. The search for limitations of machine learning and the debate on contagion via unbalanced data in the training phase, namely in the gender dimension, were based on the Prisma protocol. After defining the keywords to be searched for and the exclusion criteria, the analysis of

the 39 selected investigations was expanded using VOSviewer, Connected papers, and Treecloud. The combination of previous techniques and procedures made it possible to conclude that: (1) cognitive biases may exist in machine learning due to the use of unbalanced data for training; (2) providing data-balanced training (also on gender) to a network increases its resilience; (3) it is possible to introduce algorithms with gender-sensitive attributes to mitigate biases in machine learning; (4) there is a gap on scientific investigations regarding gender biases in the security field.

Keywords: gender-bias; security; machine learning; data.

1. Introdução

Pensar a segurança no mundo líquido, onde os riscos que acreditamos identificar e que nos provocam medo são o reflexo de nós próprios e das nossas perceções culturais, obriga a que não se fique paralisado com a factualidade de que "a ciência paira às cegas sobre os limites dos perigos" (Beck, 2015, p. 78).

De facto, a sociedade do risco em que vivemos obriga a que as respostas abrangentes sejam desenhadas e implementadas e que o envolvimento de todos os recursos (humanos ou materiais) seja efetivo (Bordoni & Bauman, 2016). Assim, a procura de soluções para desafios cada vez mais globais obriga a que as respostas não possam ser locais, ou que não incorporem todo o conhecimento existente na sociedade (Zizek, 2009).

Neste contexto em que a ciência dedicada ao estudo do risco e da gestão da segurança está confrontada com o paradoxo de que o exercício de antevisão do futuro pode servir como fonte de inspiração ao cometimento de incivilidades / atrocidades, parece ser difícil superar a dialética de que os desafios atuais são consequências da modernidade (Giddens, 1990).

Nestes termos, resulta claro que é necessário recorrer, também no campo da segurança, à utilização de máquinas para conceber soluções que vão para além da capacidade de alcance da imaginação humana, como é o caso do *Alpha Zero*, do *Generative Pre-Treined Transformer* ou da descoberta da *Halicina* (Kissinger, Schmidt, & Huttenlocher, 2021).

Conforme descreve Damásio (2011), a máxima de Descartes, *cogito ergo sum*, está errada. Neste caso, não apenas porque a ausência de emoções pode prejudicar a racionalidade, mas sobretudo porque o ser humano perdeu o monopólio da superioridade da razão, da inteligência e da capacidade de pensamento.

Não sendo possível discutir as consequências do advento da inteligência artificial (IA) para a definição do próprio conceito da humanidade (o debate sobre a largura de banda que a humanidade pode ocupar em um tempo que as máquinas pensam), pretende-se contribuir para avançar com um propósito mais humilde e circunscrito. Objetivamente, esta investigação foi desenhada com o foco de *analisar os desafios da utilização de dados com viés de género para a aprendizagem da máquina na segurança*.

O contributo para o debate sobre o modo como as máquinas aprendem a pensar e os desafios que lhe estão implícitos obriga a que o estudo esteja organizado em uma análise dos conceitos e palavras-chave, análise bibliométrica e debate dos principais conteúdos, além da explicação das opções metodológicas e da conclusão.

2. Os dados como vetor de propagação do viés de género

Os trabalhos dos psicólogos quebraram os pressupostos sobre o julgamento e a tomada de decisão, revelando que o modo como pensamos não é tão racional como esperávamos. De facto, estes estudos demonstraram que só com a combinação da utilização do nosso sistema cerebral automático e reflexivo, em cada pensamento sobre determinada situação, é possível mitigar erros de julgamento decorrentes de vieses cognitivos (Allen & Gerras, 2010).

Estar desperto para estes atalhos da mente e possuir pensamento crítico para saber que o que estamos a ver resulta da conjugação de dados objetivos com dados que confirmam as nossas crenças é fundamental em todas as áreas de negócio, mas seguramente é decisivo em instituições ligadas ao *ultima ratio regum* (Jonathan, Nathan & Joe, 2015).

Neste sentido, as instituições com responsabilidade na segurança procuram fazer esforços contínuos para dotar os seus recursos com mecanismos para ultrapassar os desafios anteriores. Até à atualidade (2023), as entidades têm procurado mitigar este desafio com a valorização de capacidades de pensamento crítico, levando a que esta competências seja uma das mais valorizadas (World Economic Forum, 2023). No entanto, considerando que as máquinas começaram a aprender, surge o desafio de se perceber se a aprendizagem com recurso a máquinas pode ser um potente transmissor de vieses, em virtude de ser baseada em dados incompletos, ou por vezes pouco incorretos, que conduzem a aprendizagens deficitárias (Cervantes, García-Lamont, Rodríguez-Mazahua, & Lopez, 2020).

A questão anterior é relevante, porque em todas as aprendizagens da IA (não supervisionadas, supervisionada ou por reforço) existe uma grande utilização de dados para verificar as tendências, ainda que na aprendizagem por reforço a IA não seja um mero espectador. Assim, se os dados fornecidos tiverem limitações, tais limitações tendem a reproduzirem-se, o típico *entra lixo sai lixo* (Domingos, 2018). Considerando que, em programas da *AlphaDogFight* da *Defense Advanced Projetcts Agency*, os pilotos de IA obtiveram melhores resultados em combates simulados que os pilotos humanos, parece ser crucial estudar este domínio da segurança pelo potencial impacto que o mesmo poderá ter (Kissinger, Schmidt, & Huttenlocher, 2021).

Os autores desta área do conhecimento parecem concordar que o facto de existirem dados não balanceados pode afetar os resultados da aprendizagem e sugerem algumas formas de se ultrapassarem tais limitações, desde aumentar as amostras ou alterar a ponderação das métricas do algoritmo (Russell & Norvig, 2020). No entanto, parece ser difícil ultrapassar esta desafio recorrendo a estas soluções típicas de redimensionamento da amostra, quando uma amostra maior de dados continua a ter as mesmas limitações, como é o caso da falta de dados desagregados por género, ou para simplificar, por sexos.

A afirmação anterior está alicerçada no argumento de Beauvoir (1949) de que a humanidade é masculina e que a mulher é definida por relação com o homem e não é vista como um ser autónomo. Com este pano de fundo e sabendo-se da centralidade dos dados na atualidade, com a utilização do *big data* para diversas finalidades, o que temos no final de uma análise de dados incompletos, ou carregados de silêncios, são consequentes meias-verdades ou até mesmo perceções erradas de determinada matéria. De facto, tendo a IA a suportar diagnósticos médicos ou analisar *curricula vitae*, que

foi treinada com défices informacionais no campo do género, o que teremos será, no mínimo, informação para a tomada de decisão que objetivamente será incompleta e quiçá errada (Perez, 2019).

Esta reprodução das desigualdades continua a ser uma realidade, não irá desaparecer com o tempo e não é apenas uma questão geracional. Por exemplo, apenas em 2016 é que a *Unicode Consortium* passou a adotar o género para a linguagem do *emoji* que até então eram exclusivamente masculinos, por exemplo, não existia *policewoman* e *policeman* (Bosque, 2018).

Áreas tão cruciais como a saúde não são exceção. Até há bem pouco tempo não se sabia que os níveis de tolerância de referência à radiação e aos químicos para o Homem (estavam incluídas as mulheres) afinal não eram seguros para as mulheres (Ministério da Saúde, 2013).

Na segurança, só recentemente foram disponibilizados coletes balísticos adaptados para a ergonomia feminina e o tipo de arma continua a não ser ajustado ao tamanho da mão da mulher, ainda que em alguns países seja possível ajustar o punho da arma de serviço (*e.g*. Espanha).

Neste contexto de aparente falta de consideração na recolha de dados desagregados por sexos e, por conseguinte, na falta de customização de soluções para um outro ator social de suma importância no seio de qualquer população – as mulheres –, parece tornar-se evidente que é necessário que as mulheres tenham mais inclusão e visibilidade no campo da ciência dos dados, sobretudo para uma área basilar dos Estados – a segurança. De facto, sobressai a relevância de se avançar com a consolidação do objetivo número cinco do desenvolvimento sustentável da Organização das Nações Unidas (2023), sobre a igualdade de género em todas as dimensões da sociedade (Agenda 2030).

Com o propósito de se identificarem as principais tendências neste campo e de se contribuir para debelar estas realidade, procurou-se identificar e sistematizar as principais tendências nesta área, dedicando-se a secção seguinte à explicação do protocolo seguido.

3. Opções metodológicas

A procura de contributos para o objetivo de investigação foi guiada pela realização de um estudo ético, ontologicamente posicionado no relativismo e epistemologicamente no interpretativismo (Saunders, Lewis, & Thornhill, 2019).

Seguindo-se um percurso do geral para o particular, estamos perante um raciocínio dedutivo, com um desenho de estudo de caso, circunscrevendo-se o mesmo ao estudo da literatura sobre viés de género e aprendizagem máquina no campo da segurança para os últimos seis anos, desde 2016 (Santos & Lima, 2019).

A recolha de dados foi conduzida com recurso ao *software Publish or Perish* (para a *Scopus* e *Google Scholar*) e ao *site* da *ScienceDirect*, tendo por base as palavras-chave que se revelaram como cruciais da análise efetuada na divisão anterior do texto: *gender-bias; security; machine learning*. Após esta definição de palavras-chave a usar na procura dos dados, foram definidos critérios de filtragem para se apurarem os artigos científicos

mais relevantes onde se centrou a revisão sistemática da literatura, nomeadamente pela verificação do fator de impacto de cada artigo científico na plataforma *Scimago Journal & Country Rank* (Easterby-Smith, Thorpe, & Jackson, 2015).

Posteriormente, tendo por base uma análise bibliométrica efetuada com recurso ao *VOSviewer*, foram desenvolvidos mapas de ligações entre os principais autores.

Em suma, procurou-se seguir o protocolo Prisma (Charbonneau, *et al.,* 2020), onde inicialmente foi identificado o que procurar e onde procurar (palavras-chave e locais de pesquisa), foram identificados e eliminados os registos duplicados nas diversas plataformas e definidos critérios de exclusão. Os critérios de exclusão passaram por não analisar os documentos que não tratavam das áreas estabelecidas nas palavras-chave, que eram anteriores a 2016 e que não eram estudos empíricos. Esta análise possibilitou um redimensionamento do número de artigos a estudar, de 127 para 39 (Apêndice A), e que são a base do debate que consta na próxima divisão do texto.

Adicionalmente, também se procurou utilizar a ferramenta designada por *Connected papers* para se obter mais informação e/ou acesso a ligações sugeridas pelos principais estudos analisados.

Finalmente, para se reforçar as conclusões de tudo o quanto se analisou, foram desenhadas árvores de palavras com os textos selecionados, refletindo a sua proximidade semântica tendo por base o *Treecloud,* resultados que também foram utilizados no debate efetuado na próxima divisão do texto.

4. Análise e discussão dos resultados

O recurso ao primeiro elemento de análise, o *VOSviewer*, permitiu identificar que o principal autor de ligação entre os clusters que surgem da procura efetuada é o Qadir, sobretudo no seu trabalho em coautoria (Butt, Qayyum, Ali, Al-Fuqaha, & Qadir, 2023), conforme demonstra a figura 1.

Com este estudo específico, os seus autores sublinham a relevância que a IA está a ter no quotidiano das pessoas, desde o agendamento de atividades diárias até recomendações de compras personalizadas, sendo que o sucesso desta mudança deve ser medido na justa proporção do benefício para os seres humanos. Atendendo a que a IA está a alargar o perímetro de segurança que estávamos habituados e que os algoritmos estão intimamente ligados aos humanos, é importante olhar para essas tecnologias de IA por meio de uma lente centrada no ser humano, sempre conscientes de que é necessário corrigir vieses que lhes estão subjacentes.

Complementarmente, os estudos mais centrais também sugerem que a robustez das redes está ligada ao modo como as mesmas são treinadas, sublinhando a relevância de se efetuarem treinos em redes pequenas e com dados equilibrados em diversas dimensões (entre as quais a dimensão género), e só depois ampliar o modelo, adequadamente, entenda-se mantendo o equilibro relativo do mesmo (Wu, Chen, Cai, He, & Gu, 2020).

Utilizando o *Connected Papers* (figura 2), constatamos que o elemento que mais sobressai é a publicação de Mehrabi, Morstatter, Saxena, Lerman e Galstyan (2019) sobre a importância na justiça/equidade da conceção dos sistemas de IA, sempre na

ótica de que os mesmos podem ser usados para tomar decisões disruptivas nas nossas vidas. No fundo, os autores sugerem que é crucial garantir que essas decisões não reflitam um comportamento discriminatório, sobretudo porque estamos no início da consciencialização de que esta realidade/dimensão representa um desafio, sendo necessário implementar soluções para mitigar vieses em sistemas de IA, nomeadamente os de género.

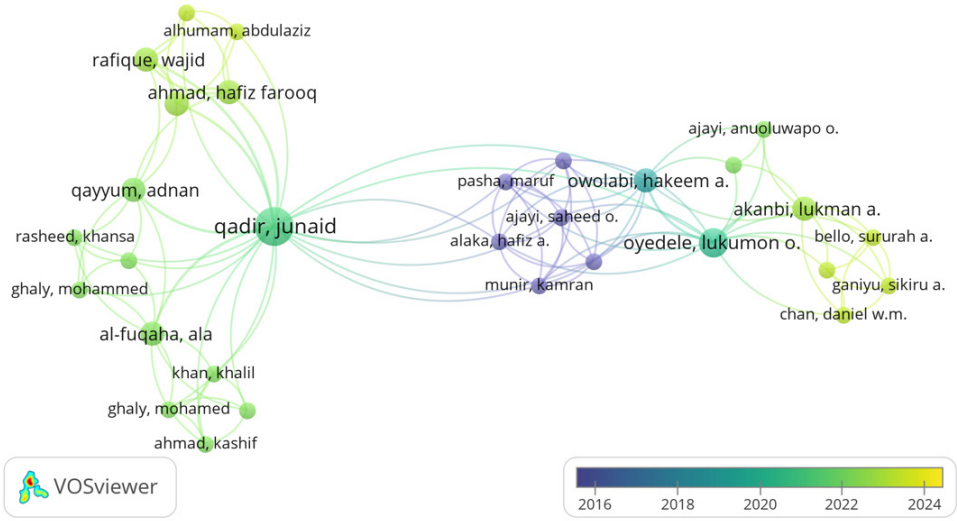

Figura 1 – Análise bibliometria VOSviewer

Outros estudos, que se interligam com o anterior, sugerem que os vieses dos modelos aumentam à medida que os conjuntos de dados se tornam mais desequilibrados e as informações permanecem na representação latente, mesmo quando os algoritmos de mitigação de vieses são aplicados (Reddy, *et al.*, 2021). Elemento que faz sobressair a criticidade da factualidade de apenas recentemente existir a recolha de dados desagregados por géneros.

Tendo por base as investigações selecionadas (Apêndice A), foi introduzido o conteúdo das mesmas no *Treecloud* e foi estabelecido um limite máximo de apresentação de 50 palavras e a sua relação de proximidade, conforme demonstra a figura 3.

A primeira constatação que surge da análise da figura 3 é que a palavra *segurança* não consta na mesma, enquanto as restantes que foram eleitas como fundamentais para serem pesquisadas surgem, ainda que de modo incompleto, como no caso do viés de género.

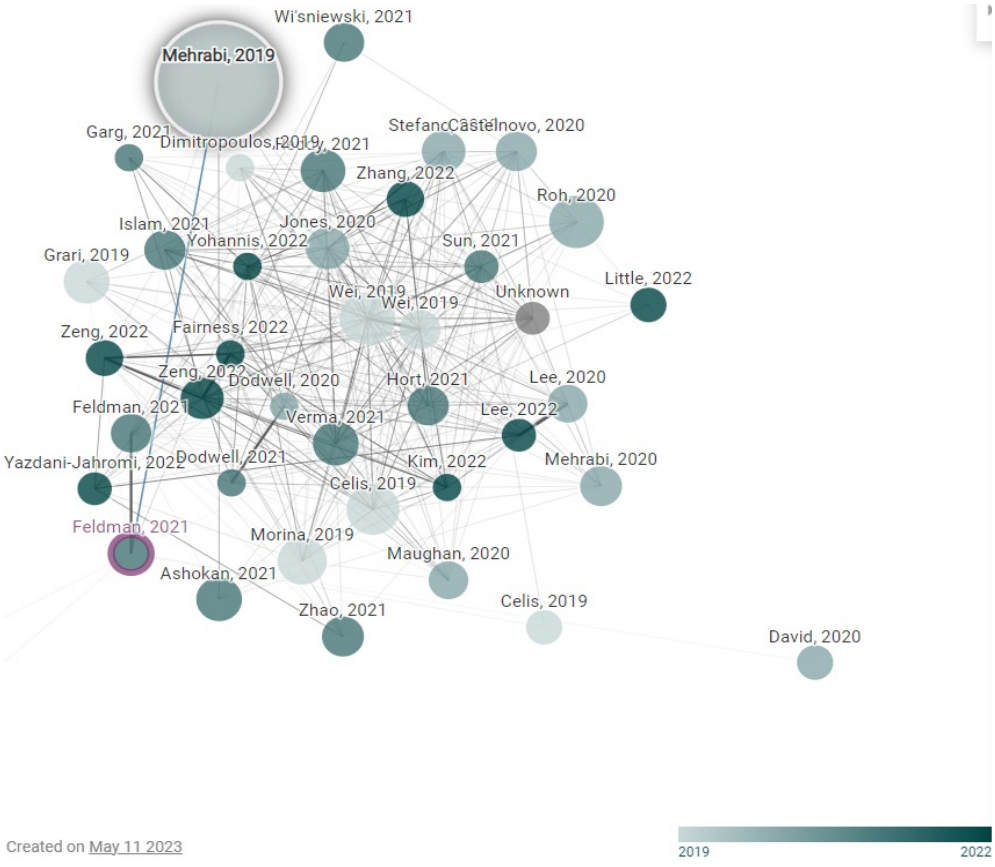

Figura 2 – Análise *Connected papers*

A segunda constatação que se pode retirar é que os modelos de aprendizagem máquina (ML – *machine learning*), podem realmente ser influenciados por vieses (também os de género), caso os algoritmos não tenham sido pensados para mitigar tais atalhos, entenda-se introdução de algoritmos com atributos *gender sensitive* cuja abordagem fomente a justiça e a precisão na aprendizagem que é fornecida/proporcionada à máquina.

5. Conclusões

A presente investigação teve como principal propósito contribuir para *analisar os desafios da utilização de dados com viés de género para a aprendizagem da máquina na segurança*.

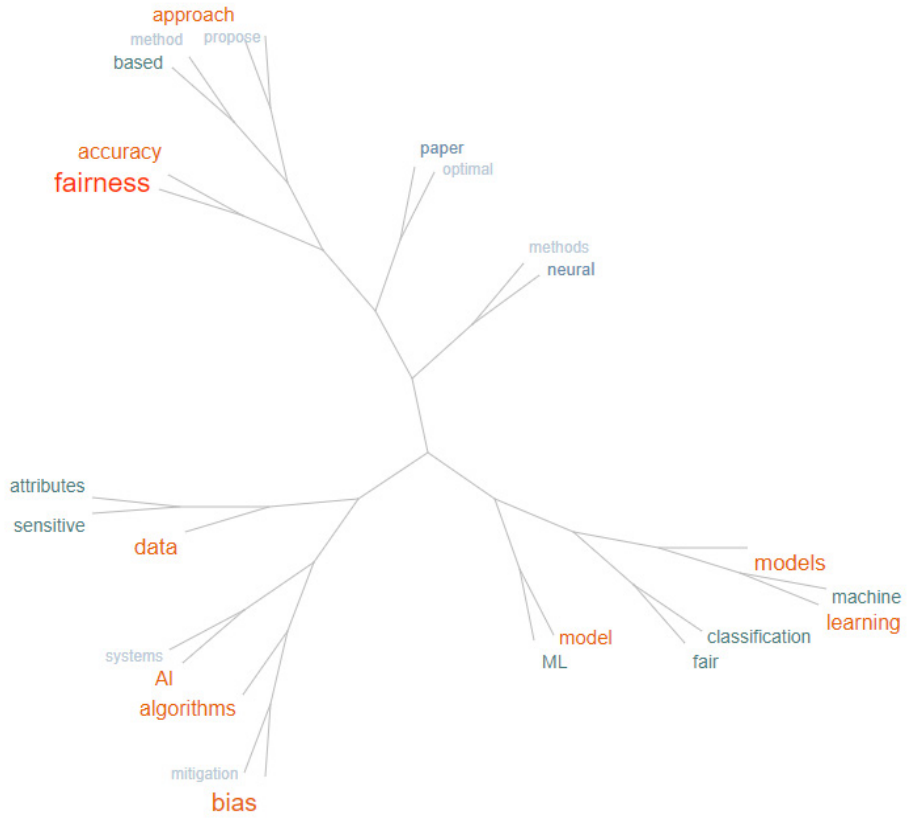

Figura 3 – Análise *Treecloud*

O objetivo de investigação surge com a dialética de que o pensamento deixou de ser um monopólio da humanidade para ser mais um domínio de competição. Neste ambiente de livre mercado, em que as máquinas começam a demonstrar competências que o ser humano nem imaginava, é obrigatório efetuar exercícios de verificação das limitações deste pensamento desprovido de alma, nomeadamente se este pensamento pode ser influenciado pelos mesmos vieses que normalmente afetam os falíveis seres mortais.

Considerando a diversidade de vieses cognitivos conhecidos, foi selecionado o viés de género para ser estudado. A escolha deste viés teve por base o potencial efeito de *spill over* que esta área pode proporcionar; foi ponderado o atraso relativo que o tema possuiu em áreas cruciais como a segurança, além de se constatar que se nada for feito, as desigualdades tenderão a ser reproduzidas para o mundo digital, mas desta feita com resultados potenciais mais nefastos devido ao efeito multiplicador inerente à inteligência artificial.

Tendo por base o protocolo Prisma, foram identificados 127 artigos em diversas plataformas, tendo em consideração as palavras-chave definidas: *gender-bias; security; machine learning*. Após a definição de critérios de exclusão e a eliminação de resultados repetidos, foi efetuada uma análise bibliométrica no *VOSviewer* às 39 investigações selecionadas. Após se identificarem os mapas de ligações entre os principais autores, também se recorreu ao *Connected papers* para obter mais informação sobre as ligações.

A conjugação dos procedimentos anteriores e ainda a utilização de árvores de palavras com base na proximidade semântica no *Treecloud* permitem sublinhar que: (1) os vieses cognitivos podem existir na aprendizagem máquina devido ao recurso a dados desequilibrados/incompletos para treino; (2) proporcionar um treino com dados equilibrados (também sobre género) a uma rede aumenta a sua resiliência; (3) é possível introduzir algoritmos com atributos *gender sensitive* para mitigar vieses na aprendizagem máquina; (4) existe um *gap* sobre investigações científicas relativamente a vieses de género no campo da segurança.

Formando a livre convicção de que o tema não se esgota com esta investigação, sendo desejável alargar o número de plataformas utilizadas, calibrar as palavras-chave e os critérios de exclusão, afigura-se que este estudo de cariz exploratório pode servir de base para se aumentar o conhecimento sobre estas *drivers* de mudança da gestão da segurança.

Acknowledgements

This research was developed during the PhD program in Military Science at Instituto Universitário Militar, as part of the PhD thesis: *Integrated approach to gender in gendarmeries*.

Funding

This research was funded by CINAMIL and the European Commission within the EU programme for education, training, youth, and sports (Erasmus+), grant number 2020-1-PT01-KA203-078544. The APC was funded by the Military Gender Studies project–https://erasmus-plus.ec.europa.eu/projects/search/details/2020-1-PT01-KA203- 078544.

Referências

Allen. D. Charles & Gerras. J. Stephen. (2010). Como Desenvolver Pensadores Criativos e Críticos. *Military Review*. Retirado de https://www.armyupress.army.mil/Portals/7/military-review/Archives/Portuguese/MilitaryReview_20101031_art006POR.pdf

Beauvoir, S. (1949). *The Second Sex*. Londres: Parshley.

Beck, U. (2015). *Sociedade de risco mundial*. Lisboa: Almedina.

Bordoni, C., & Bauman, Z. (2016). *Estado de Crise*. Lisboa: Relógio D'Água.

Bosque, I. (2018). *Sexismo linguístico y visibilidade de la mujer*. Retirado de https://elpais.com/cultura/2012/03/02/actualidad/1330717685_771121.html

Butt, M., Qayyum, A., Ali, H., Al-Fuqaha, A., & Qadir, J. (2023). Towards secure private and trustworthy human-centric embedded machine learning: An emotion-aware facial recognition case study. *Computers & Security*, https://doi.org/10.1016/j.cose.2022.103058.

Cervantes, J., García-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges, and trends. *Neurocomputing* https://doi.org/10.1016/j.neucom.2019.10.118), 189-215.

Charbonneau, É., Lévesque, J., Tchokouagueu, F., Tadjiogue, Y., Tchinda, K., & Mongong, M. (2020). Camouflaged or deserted? A systematic review of empirical military research in public administration. *Asia Pacific Journal of Public Administration*, 10.1080/23276665.2020.1839351.

Damásio, A. (2011). *O Erro de Descartes Emoção, razão e cérebro humano*. Lisboa: Círculo de Leitores.

Domingos, P. (2018). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books, Inc.

Easterby-Smith, M., Thorpe, R., & Jackson, P. (2015). *Management and Business Research 5th Edition*. London: Sage.

Giddens, A. (1990). *The consequences of Modernity*. Combridge: Polity.

Jonathan, Nathan & Joe. (2015). Como Preparar os Militares para a Incerteza. *Military Review*. Retirado de https://www.armyupress.army.mil/Portals/7/military-review/Archives/Portuguese/MilitaryReview_20150630_art011POR.pdf

Kissinger, H., Schmidt, E., & Huttenlocher, D. (2021). *A era da inteligência artificial*. NY: Dom Quixote.

Ministério da Saúde. (2013). *Guidelines for Surveillance in Job Related Cancer*. Obtido de https://bvsms.saude.gov.br/bvs/publicacoes/diretrizes_vigilancia_cancer_relacionado_2ed.pdf

Ninareh, M., Morstatter, F., Saxena, M., Lerman, K., & Lerman, A. (2019). A Survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, https://doi.org/10.1145/3457607.

Organização das Nações Unidas (2023). *Goal 5: Achieve gender equality and empower all women and girls*. Retirado de https://www.un.org/sustainabledevelopment/gender-equality/

Perez, C. (2019). *Mulheres Invisíveis*. Lisboa: Relógio de Água.

Reddy, C., Sharma, D., Mehri, S., Romero-Soriano, A., Shabanian, A., & Honari, S. (2021). Benchmarking bias mitigation algorithms in representation learning through fairness metrics. *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track* , 1-25.

Russell, S., & Norvig, P. (2020). *Artificial intelligence a modern approach*. Pearson Education.

Santos, L., & Lima, J. (2019). *Orientaçõesesmetodológicas para a elaboração de trabalhos de investigação (2.ª ed., revista e atualizada). Cadernos do IUM*. Lisboa: Instituto Universitário Militar.

Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research Methods for Business Students 8 th ed*. Harlow: Pearson Education.

World Economic Forum. (2023). *The Future of Jobs Report*. Retirado de https://www.weforum.org/agenda/2020/10/top-10-work-skills-of-tomorrow-how-long-it-takes-to-learn-them/?DAG=3&gclid=EAIaIQobChMI_JzT14Hf_gIVWv1RCh20Qw-kEAAYASAAEgLYsvD_BwE

Wu, B., Chen, J., Cai, D., He, X., & Gu, Q. (2020). Do Wider Neural Networks Really Help Adversarial. *Neural Information Processing Systems*, 1-20.

Zizek, S. (2009). *Violência*. Lisboa: Relógio D'Água.

Apêndice A – Artigos selecionados

Título	Autores	Ano
A Survey on Bias and Fairness in Machine Learning	Ninareh Mehrabi, Fred Morstatter, N. Saxena, Kristina Lerman, A. Galstyan	2019
A Generative Approach to Mitigate Bias in Face Matching using Learned Latent Structure	Jamal Alasadi; Ahmed AlHilli; Pradeep K. Atrey; Vivek K. Singh	2022
Gender Bias in Artificial Intelligence	Enrique Latorre Ruiz & Eulalia Pérez Sedeño	2023
Gender Differences and Bias in Artificial Intelligence	Valentina Franzoni	2023
Towards secure private and trustworthy human-centric embedded machine learning: An emotion-aware facial recognition case study	Muhammad Atif Butt a, Adnan Qayyum a, Hassan Ali a, Ala Al-Fuqaha b, Junaid Qadir	2023
A systematic review of socio-technical gender bias in AI algorithms	Paula Hall, Debbie Ellis	
Understanding and Mitigating Gender Bias in Information Retrieval Systems	Amin Bigdeli, Negar Arabzadeh, Shirin Seyedsalehi, Morteza Zihayat & Ebrahim Bagheri	2023
Analysis Effect of Bias lying in Artificial Intelligence: A Survey	Bangdao Liu, Keyao Wang, Mutian Yang	2023
Better security through diversity – a personal perspective	Leila Iranmanesh	2022
Mitigate Gender Bias using Negative Multi-Task Learning	Liyuan Gao, Huixin Zhan, Austin Chen, Victor Sheng	2022
Dealing with Gender Bias Issues in Data-Algorithmic Processes: A Social-Statistical Perspective	Juliana Castaneda, Assumpta Jover, Laura Calvet, Sergi Yanes, Angel A. Juan, Milagros Sainz	2022
Peer review and gender bias: A study on 145 scholarly journals	F. Squazzoni, Giangiacomo Bravo, M. Farjam, A. Maruši ć, B. Mehmani, Michael Willis, Aliaksandr Birukou, Pierpaolo Dondio, F. Grimaldo	2021
No evidence of any systematic bias against manuscripts by women in the peer review process of 145 scholarly journals	F. Squazzoni, Giangiacomo Bravo, Pierpaolo Dondio, M. Farjam, A. Maruši ć, B. Menhami, Michael Willis, Aliaksandr Birukou, F. Grimaldo	2020
Data and Techniques Used for Analysis of Women Authorship in STEMM: A Review	V. Bhagat	2019
Championing the Success of Women in Science, Technology, Engineering, Maths, and Medicine	Suw Charman-Anderson, Lauren Kane, Alice Meadows, B. G. Tzovaras, Stacy Konkiel, L. Wheeler	2017
Measuring the developmental function of peer review: a multi-dimensional, cross-disciplinary analysis of peer review reports from 740 academic journals	D. García-Costa, F. Squazzoni, B. Mehmani, F. Grimaldo	2021
Gender differences among active reviewers: an investigation based on publons	Lin Zhang, Yuanyuan Shang, Ying Huang, G. Sivertsen	2022
Mind the (submission) gap: EPSR gender data and female authors publishing perceptions	C. Closa, Catherine Moury, Z. Novakova, M. Qvortrup, Beatriz Ribeiro	2020
Women Authorship of Scholarly Publications in STEMM: Authorship Puzzle	V. Bhagat	2019
Gender and international diversity improves equity in peer review	Dakota S. Murray, Kyle Siler, V. Larivière, Wei Mun Chan, Andrew M. Collings, Jennifer Raymond, C. Sugimoto	2018

Título	Autores	Ano
Representation of Women in HPC Conferences	Eitan Frachtenberg, Rhody D. Kaner	2021
How Gendered Is the Peer-Review Process? A Mixed-Design Analysis of Reviewer Feedback	T. König, G. Ropers	2021
A gender bias in the European Journal of Political Research?	E. Grossman	2020
Large-scale language analysis of peer review reports	I. Buljan, D. García-Costa, F. Grimaldo, F. Squazzoni, A. Marušić	2020
Discovering Inclusivity in Remote Sensing: Leaving No One Behind	K. Joyce, C. Nakalembe, Cristina Gómez, G. Suresh, Kate C. Fickas, Meghan Halabisky, M. Kalamandeen, M. Crowley	2022
Unprofessional peer reviews disproportionately harm underrepresented groups in STEM	N. Silbiger, Amber D. Stubler	2019
Implementation of the article "Representation of Women in High-Performance Computing Conferences"	Eitan Frachtenberg, Rhody D. Kaner	2020
Females Are First Authors, Sole Authors, and Reviewers of Entomology Publications Significantly Less Often Than Males	K. A. Walker	2019
Investigation of potential gender bias in the peer review system at Reproduction	Marie Biolková, T. Moore, K. Schindler, Karl Swann, A. Vail, Lindsay Flook, Helen Dick, G. Fitzharris, C. Price, Norah Spears	2023
Underrepresentation of women in computer systems research	Eitan Frachtenberg, Rhody D. Kaner	2022
Gender differences in scientific careers: A large-scale bibliometric analysis	Hanjo D. Boekhout, Inge van der Weijden, L. Waltman	2021
Gender gap among highly cited researchers, 2014–2021	Lokman I. Meho	2022
Gender Differences in Collaboration Patterns in Computer Science	Josh Yamamoto, Eitan Frachtenberg	2022
Disparities in a Flagship Political Science Journal? Analyzing Publication Patterns in the Journal of Politics, 1939–2019	Joseph Saraceno	2020
Gender trends in computer science authorship	Lucy Lu Wang, Gabriel Stanovsky, Luca Weihs, Oren Etzioni	2019
Gender differences in patterns of authorship do not affect peer review outcomes at an ecology journal	C. Fox, C. S. Burns, Anna D. Muncy, Jennifer A. Meyer	2016
Analysis of scientific society honors reveals disparities.	Trang T. Le, Daniel S. Himmelstein, Ariel A. Hippen, Matthew R. Gazzara, C. Greene	2021
Women's representation and experiences in the high performance computing community	A. Frantzana	2019
A bibliometric approach for detecting the gender gap in computer science	Sandra Mattauch, Katja Lohmann, Frank Hannig, Daniel Lohmann, J. Teich	2020