

Attività 3.1.2 - Principali problematiche e rischi per la società nell'uso dell'IA

Durata stimata	40 minuti
Requisiti	Moduli 1 e 2 Attività 3.1.1. Se effettuato in presenza, serve uno spazio abbastanza grande in cui gli studenti possano discutere e svolgere presentazioni in gruppi
Strumenti software e hardware utilizzati	In remoto: Software - programma per le video conferenze. Software per condividere online in tempo reale documenti di testo e presentazioni. Hardware - Computer fisso/portatile
Tipo di valutazione	La valutazione si basa sulla comprensione delle questioni discusse in classe, evidenziata attraverso la presentazione di argomenti che evidenziano l'impatto negativo dell'IA e obiezioni che mettono in discussione tali argomenti.

Scopo di questa attività è facilitare la comprensione del modo in cui l'uso delle tecnologie di IA possa avere un impatto negativo sul mondo reale, servendosi di argomenti ed esempi di supporto. Saranno incoraggiati ed esaminati anche i commenti e le domande che mettono in discussione tali argomenti.

Procedimento (fasi)

Introduzione

- Gli insegnanti possono ricordare agli studenti uno dei sistemi di intelligenza artificiale studiati in precedenza (ad esempio chatbot, assistenti digitali, motori di ricerca, riconoscimento di immagini come il rilevamento dei volti). Gli insegnanti fanno nuovamente cenno agli SDG presentati nell'attività precedente, per spiegare che l'uso dell'IA può anche sortire effetti opposti e allontanarci dal raggiungimento di tali obiettivi e illustrano alcuni esempi per spiegare questo aspetto.

- Il caso dei bias algoritmici. L'insegnante spiega che gli esseri umani a volte prendono decisioni assistite da algoritmi (sequenze di azioni o passaggi necessari per risolvere un problema); queste decisioni sono, quindi, in parte automatiche. Le indicazioni date dal computer possono trasmettere un'impressione di razionalità o infallibilità e le persone possono quindi essere portate a seguirle ciecamente. In ogni caso, a seconda del modo in cui è stato "addestrato" a funzionare e in base alle informazioni che gli esseri umani forniscono al PC (apprendimento automatico), l'algoritmo può suggerire decisioni "discriminatorie", cioè favorire certi tipi di persone rispetto ad altre. Si tratta dei cosiddetti "bias dell'IA" (bias = distorsione del processo decisionale sulla base di un pregiudizio).
- L'insegnante propone un esempio: un impiegato di banca può decidere se un cliente può chiedere un prestito alla banca o meno basandosi su una valutazione del cliente effettuata automaticamente da un algoritmo. Se l'algoritmo si basa su dati relativi a un periodo in cui nessuna donna poteva chiedere un prestito in banca, rischia di dare l'indicazione che i clienti migliori a cui concedere un prestito sono gli uomini. Questo comporta una discriminazione, cioè una disparità di trattamento, nei confronti delle donne. Il che contrasta con l'SDG 5 (Parità di genere).
- L'insegnante illustra altri possibili problemi: questioni di privacy nell'uso degli algoritmi di riconoscimento facciale in luoghi pubblici, oppure problemi di sicurezza e privacy collegati all'uso dei sistemi che elaborano la voce e i suoni come gli assistenti digitali. Altri esempi riguardano le difficoltà degli algoritmi di riconoscimento vocale nell'identificare con precisione i comandi provenienti da voci di persone anziane o di bambini, perché sono stati addestrati su una popolazione adulta di mezza età. I software di riconoscimento delle immagini o i motori di ricerca basati sulle immagini possono presentare risultati falsati da bias (si veda <https://en.unesco.org/artificial-intelligence/ethics/cases#biasedai>)

Istruzioni dirette

- Gli insegnanti dividono gli studenti in gruppi.
- Ogni gruppo esamina un esempio di implicazioni negative o di rischi causati dalle applicazioni concrete dell'IA.
- Ogni gruppo dovrà riflettere sulle possibili implicazioni di tali esempi in altri settori/rispetto ad altre questioni chiave.

Pratica guidata e comprensione

- Ogni gruppo, utilizzando il materiale fornito dall'insegnante, elabora un modo per rappresentare l'esempio di cui sta discutendo per evidenziare i modi in cui le tecnologie di IA creano rischi e hanno un impatto negativo rispetto alla questione specifica, nonché gli argomenti contrari a tale applicazione. Ogni studente potrebbe rappresentare un aspetto, un agente o una funzionalità dell'intero sistema.

Attività pratica indipendente

- Gli studenti di ogni gruppo presentano all'intera classe l'esempio di applicazione dell'IA e gli argomenti contrari.

Chiusura

- Il resto della classe mette in discussione gli argomenti contrari all'utilizzo dell'IA in quel particolare settore servendosi di commenti critici e domande. Gli insegnanti moderano la discussione e chiariscono qualsiasi eventuale fraintendimento o dubbio degli altri studenti.

Risoluzione dei problemi

Possibile problema	Possibile soluzione
Gli esempi di applicazioni concrete dell'IA non sono chiari per gli studenti, oppure gli studenti non riescono a collegare i concetti appresi in precedenza con le applicazioni dell'IA che stanno cercando di spiegare.	Rassicurazione da parte dell'insegnante e, se necessario, nuova spiegazione dei concetti più complessi.
Gli studenti non sono disposti a rappresentare fisicamente un'applicazione concreta delle tecnologie di IA.	Collaborare con un insegnante di teatro e rendere la rappresentazione parte della lezione in un ambiente più appropriato, come un teatro scolastico o uno spazio ampio.
Gli studenti non contestano gli esempi e le argomentazioni presentati da un gruppo	L'insegnante pone delle domande che incoraggino il dibattito critico su quella specifica applicazione dell'IA

Risorse per gli insegnanti

Risorse che possono essere utilizzate in classe per introdurre esempi che riguardano le applicazioni dell'IA e la dimensione etica

Titolo	Descrizione	Link	Lingue
Etica e IA: Privacy & the Future of Work (Etica e IA: Privacy e futuro del lavoro)	Video che illustra come l'etica e la privacy siano collegate all'IA e come quest'ultima possa influenzare i lavori futuri.	https://www.youtube.com/watch?v=zNxx5gJtHLc https://youtu.be/PVDguCvRwEs?list=PLIz0pQKPcu8	INGLESE con sottotitoli autogenerati INGLESE con sottotitoli in

		9NB4AdYYuixHPnN_2QocuZ	GRECO
Etica e IA: Pari opportunità e bias algoritmici	Video che descrive il potenziale dell'IA a favore della società, focalizzandosi su etica, opportunità ed equità	https://www.youtube.com/watch?v=tJQSyzBUAew https://www.youtube.com/watch?v=yUSj9VsM82U&list=PLIzOpQKPcu89NB4AdYYuixHPnN_2QocuZ&index=6	INGLESE con sottotitoli autogenerati INGLESE con sottotitoli in GRECO
ως λειτουργεί η Τεχνητή Νοημοσύνη - Τεχνητή Νοημοσύνη: Δεδομένα και Μεροληψία	Un breve video che spiega il legame tra i dati di addestramento nell'IA e i bias.	https://www.youtube.com/watch?v=7a2aNF-dRo8&list=PLIzOpQKPcu89NB4AdYYuixHPnN_2QocuZ&index=3	INGLESE con sottotitoli in GRECO
Come funziona il riconoscimento facciale?	Un breve video che spiega come il riconoscimento facciale possa essere usato a favore della società o come possa essere usato a danno della privacy delle persone	https://youtu.be/BqQT4slOYA0	INGLESE con sottotitoli autogenerati
ИЗКУСТВЕНИЯТ интелект и ЗАЩО Илон МЪСК се обърна срещу себе си?	Video sull'IA e Elon Musk, con altri esempi di applicazione dell'IA	https://www.youtube.com/watch?v=wFPp2IEQPog	BULGARO
Algoritmo e pregiudizio	Video sul modo in cui i pregiudizi e gli stereotipi umani possono influenzare gli algoritmi	https://www.raiscuola.rai.it/tecnologia/articoli/2021/05/Paola-Velardi-a-Maestri-d557376c-fe22-41e8-880f-a09413aec33c.html	ITALIANO
Dove nascono le Idee: la Sicurezza Intelligente	Video sulle applicazioni dell'IA in materia di sicurezza	https://www.leonardocompany.com/it/news-and-stories-detail/-/detail/the-intelligent-security	ITALIANO

Impatto negativo dell'IA/rischio rappresentato dalle risorse di IA:

Distorsioni degli algoritmi

- [Amazon scraps secret AI recruiting tool that showed bias against women](#)
- [Google will pay \\$2.6 million to workers over claims its hiring and pay practices were biased against women and Asians](#)
- [Algorithm-driven Hiring Tools & Disability Discrimination](#)

- Köchling, A., Wehner, M.C. Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Bus Res* **13**, 795–848 (2020).
<https://doi.org/10.1007/s40685-020-00134-w>
- Wired. “Machines Taught by Photos Learn a Sexist View of Women: Algorithms showed a tendency to associate women with shopping and men with shooting”.
- [in italiano] Festival dei Diritti Umani 2021 - Algoritmocrazia
“Nuove tecnologie, vecchi bias” - Educational session
Includes resources on use of images & algorithms, digital citizenship, hate speech (anche Modulo 4)
<https://festivaldirittiumani.it/event/nuove-tecnologie-vecchi-bias/>
Vedere anche: <https://festivaldirittiumani.it/irony-di-radheya-jegatheva/>
- [Centre for Democracy and technology, Responsible Use of Data and Technology in Education: Community Engagement to Ensure Students and Families Are Helped, Not Hurt](#)

Questioni relative alla privacy

- [Centre for Democracy and Technology, Student privacy protection Training Module](#)
- Stoilova, Mariya; Livingstone, Sonia; Khazbak, Rana (2021). Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children’s internet use and outcomes, Innocenti Discussion Papers no. 2021-01, UNICEF Office of Research - Innocenti, Florence.
<https://www.unicef-irc.org/publications/1183-investigating-risks-and-opportunities-for-children-in-a-digital-world.html>
Per ulteriori risorse dell’UNICEF in materia di diritti dei minori nell’era digitale:
<https://www.unicef-irc.org/research/child-rights-in-the-digital-age/>
- Save the Children (2020). Digital Safeguarding For Migrating And Displaced Children: An overview of the current context and trends, potential risks and practical next steps.
https://resourcecentre.savethechildren.net/node/18477/pdf/mdi_digital_final_rgb_rev_091220.pdf

Vigilanza

- File della BBC su 4 Podcast (2020). “Facial recognition is being rolled out across Britain. But is it a vital tool against crime - or a threat to the innocent?”
<https://www.bbc.co.uk/programmes/m000dpkt>
- Uso del riconoscimento facciale nelle scuole:
<https://www.biometricupdate.com/201910/french-privacy-regulator-finds-facial-recognition-gates-in-schools-illegal>
<https://www.politico.eu/article/french-privacy-watchdog-says-facial-recognition-trial-in-high-schools-is-illegal-privacy/>
(Vedere anche Automating Society Report 2020)
- In Svezia si stanno portando avanti due progetti pilota che concentrano la loro attenzione sull’uso dell’IA e sul riconoscimento facciale degli studenti a scopi amministrativi, per consentire agli insegnanti di dedicare il loro tempo all’insegnamento

<https://www.electronicsspecifier.com/products/artificial-intelligence/facial-recognition-tested-in-swedish-high-school>

- Cambiamenti strutturali e culturali delle scuole e della scolarizzazione, si veda Andrejevic and Selwyn, Facial Recognition technology in schools: critical questions and concerns, <https://www.tandfonline.com/doi/full/10.1080/17439884.2020.1686014>

Ripercussioni ambientali ed economiche dei sistemi basati sull'IA

- BBC The Documentary Podcast (2019). "Will AI kill development? Will robotisation prevent poorer countries taking the traditional route to prosperity?" <https://www.bbc.co.uk/programmes/p075nqyy>
- Sahota, Neil. "Beyond AI: The Need For Solutionists In Creating A Sustainable World". Forbes. 24 aprile 2019. <https://www.forbes.com/sites/cognitiveworld/2019/04/24/beyond-a-i-the-need-for-solutionists-in-creating-a-sustainable-world/>
- Jones, Phil. "Refugees help power machine learning advances at Microsoft, Facebook, and Amazon". Restofworld. 22 settembre 2021 <https://restofworld.org/2021/refugees-machine-learning-big-tech/>
- Deutsche Welle Documentary (2021) "Invisibles – exploitation in the digital world of work" <https://www.youtube.com/watch?v=o-HphMeZR3k>
- Strubell et al. 2019. "Energy and Policy Considerations for Deep Learning in NLP." Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. <https://www.aclweb.org/anthology/P19-1355>. DOI:10.18653/v1/P19-1355
- Schwartz et al. 2019. "Green AI." <https://arxiv.org/abs/1907.10597>
- Why we should care about the environmental impact of AI <https://www.forbes.com/sites/forbestechcouncil/2020/08/17/why-we-should-care-about-the-environmental-impact-of-ai/?sh=766825ef56ee>
- [in italiano] Roberto Cingolani nuovo ministro della transizione ecologica. Ecco il manifesto per la sobrietà digitale che ha scritto sull'Espresso
Il fisico neo titolare del dicastero: "Non dobbiamo rinunciare alla tecnologia. Ma è bene sapere che ogni azione digitale ha una conseguenza sull'ambiente. Un esempio? Dispositivi, server e reti producono il doppio di Co2 del traffico aereo" https://espresso.repubblica.it/idee/2021/02/09/news/mentre_leggi_questo_articolo_stai_inquinando_anche_tu-286796530/