iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** I **Month of publication:** January 2025

DOI: https://doi.org/10.22214/ijraset.2025.66505

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue I Jan 2025- Available at www.ijraset.com

Ethical Concerns in Large-Scale Data Labeling and Usage: Challenges and Solutions

Ammar Ali Ammar[1], Alhoussein I Houssein[2], Ali Mohammed Abusbaiha[3]
[1]*Higher Institute of Engineering Technology, Sebha.*
[2]*College of science and Technology, Qaminis*
[3]*Higher institute of science and Technical, AlQARAYT*

Abstract: ***The rapid adoption of artificial intelligence (AI) and machine learning (ML) has created an unprecedented demand for high-quality labeled data. Large-scale data labeling, a critical component of AI system development, often involves vast datasets sourced from diverse populations and annotated using a combination of automated processes and human labor. However, the ethical challenges associated with these practices have gained significant attention. This paper explores key ethical concerns in large-scale data labeling and usage, focusing on four critical areas: bias, privacy, labor practices, and transparency. Bias in labeled data, arising from the inherent subjectivity of annotators and the unrepresentative nature of many datasets, exacerbates the risk of unfair or discriminatory outcomes in AI applications. Privacy violations occur when sensitive information is collected or used without proper consent, often challenging the effectiveness of anonymization techniques. Furthermore, the reliance on crowdsourced labor for data annotation raises concerns about worker exploitation, low compensation, and the mental toll of labeling sensitive or explicit content. Lastly, the lack of transparency and accountability in data collection and labeling processes undermines public trust and ethical standards.***

Through a comprehensive review of existing practices, this paper highlights real-world case studies and controversies, including biased datasets and privacy violations. Current technological and policy-driven solutions—such as privacy-preserving techniques, labor reforms, and bias mitigation algorithms—are critically examined. Finally, the paper discusses the challenges of implementing these solutions at scale and identifies future research directions. By addressing these concerns, this work aims to promote more equitable, transparent, and ethical practices in the lifecycle of AI data management.

Keywords: Ethical AI, Data Labeling, Bias in Machine Learning ,Privacy in AI, Crowdsourced Data Annotation ,Data Transparency ,Algorithmic Fairness ,Responsible AI Practices, Federated Learning ,Differential Privacy ,Data Ethics ,Annotator Bias ,AI Governance, Data Privacy Regulations ,Sustainable AI Development.

I. INTRODUCTION

In the era of artificial intelligence (AI) and machine learning (ML), data has become a cornerstone for innovation and progress. High-quality labeled data fuels the development of predictive models, enabling applications ranging from healthcare diagnostics to autonomous vehicles. However, the rapid expansion of AI-driven technologies has amplified the demand for large-scale datasets, leading to complex ethical challenges surrounding their collection, labeling, and usage [1].

Data labeling, the process of annotating raw data to make it usable for AI training, often relies on a combination of automated techniques and human effort. Crowdsourced platforms, such as Amazon Mechanical Turk and Scale AI, have emerged as popular solutions for addressing the growing need for annotated data. Despite their utility, these practices raise significant ethical concerns. Issues such as biases introduced by annotators, privacy violations during data collection, exploitative labor practices, and the lack of transparency in labeling workflows have garnered attention from researchers, practitioners, and policymakers alike [2].

In summary, data labeling is a multifaceted process that demands precision, adaptability, and collaboration. By understanding these challenges and considering diverse viewpoints, organizations can unlock the true business value of labeled data, propelling AI innovations forward [3].

©IJRASET: All Rights are Reserved

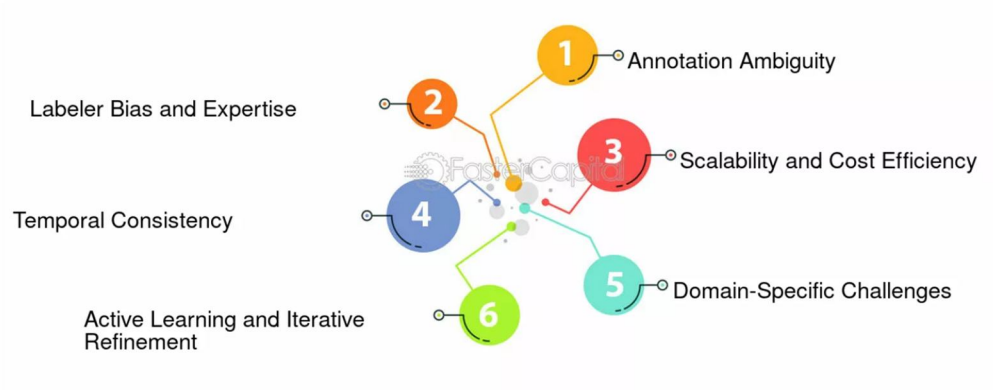


Figure 1: Challenges and Considerations in Data Labeling

One pressing concern is the propagation of bias through labeled datasets. Human annotators bring their perspectives, experiences, and cultural contexts to the labeling process, which can inadvertently lead to datasets that reinforce existing inequalities. Furthermore, the widespread use of personal and sensitive information in datasets poses critical privacy risks, particularly in the absence of robust consent mechanisms and anonymization practices. Labor exploitation is another major issue, as many data labeling tasks are outsourced to low-wage workers who may lack adequate support, training, or protection from exposure to harmful content [4].

This paper aims to comprehensively examine the ethical dimensions of large-scale data labeling and usage. By identifying key challenges and reviewing existing solutions—ranging from privacy-preserving technologies to policy-driven labor reforms—it seeks to provide a roadmap for fostering more ethical practices in data management. The discussion also includes real-world case studies and controversies that highlight the urgency of addressing these concerns [5].

As the reliance on labeled data continues to grow, ensuring the ethical integrity of data collection and labeling processes is paramount. By addressing these challenges, we can build AI systems that are not only innovative but also equitable, responsible, and aligned with societal values [6].

II. KEY SECTIONS

Background:

The field of artificial intelligence (AI) and machine learning (ML) has undergone significant advancements over the past decade, driven largely by the availability of massive datasets. These datasets, when properly labeled, serve as the foundation for training complex AI models. Data labeling is the process of annotating raw data—such as text, images, audio, or video—into a structured format that machines can interpret and learn from. For example, labeling objects in images, classifying text into categories, or transcribing speech into text are common data labeling tasks[7,8].

A. Evolution of Data Labeling Practices

Traditionally, small, manually annotated datasets were sufficient for AI research. However, with the rise of deep learning and the need for large-scale models, the demand for extensive and diverse labeled datasets has grown exponentially. This demand has led to the development of new data labeling methods, including:

1) Manual Labeling: Human annotators perform labeling tasks through dedicated platforms or in-house teams. Crowdsourcing platforms like Amazon Mechanical Turk and Appen are widely used for this purpose [9].

2) Automated and Semi-Automated Labeling: Advances in AI have enabled the use of pre-trained models or algorithms to perform initial labeling, which is later refined by humans (human-in-the-loop systems).

B. Role of Labeled Data in AI Systems

Labeled data is essential for supervised learning, the most common paradigm in AI. Models trained on labeled datasets are used in various applications, including:

1) Healthcare: Medical imaging diagnostics rely on labeled datasets to identify diseases.

2) Autonomous Vehicles: Object detection and traffic sign recognition depend on extensive labeled image data.

©IJRASET: All Rights are Reserved

3) Natural Language Processing (NLP**)**: Sentiment analysis, machine translation, and chatbots require annotated textual data [10].
The quality, diversity, and volume of labeled data directly influence the accuracy and generalizability of AI systems. However, the increased reliance on large-scale data labeling has exposed several ethical and operational challenges [11].

C. Emergence of Ethical Concerns

The process of data labeling, while critical, is not without ethical pitfalls. Bias, privacy risks, labor exploitation, and the lack of transparency are recurring issues. For instance, datasets labeled without careful oversight may reflect societal stereotypes, leading to biased AI outcomes. The reliance on crowdsourced labor often involves precarious working conditions for annotators, who are paid low wages and lack protections. Furthermore, the collection and usage of sensitive personal data, often without proper consent, have raised significant privacy concerns [12,13].

D. Current Efforts and Gaps

While technological and policy-driven solutions have been proposed to address these concerns, gaps remain in their implementation and scalability. Privacy-preserving techniques like differential privacy and federated learning show promise but require broader adoption. Similarly, initiatives to improve working conditions for data annotators are often limited to specific platforms or regions[14].
This background sets the stage for a deeper exploration of the ethical challenges in large-scale data labeling and usage, as well as potential pathways toward more responsible practices.

III. ETHICAL ISSUES

A. Bias in Data Labeling:

Bias in data labeling is a pervasive issue that can significantly impact the fairness, accuracy, and reliability of AI systems. It arises when datasets used for training machine learning models do not adequately represent the diversity of real-world scenarios or reflect prejudiced perspectives. This section delves into the sources, manifestations, and consequences of bias in data labeling [15].

1) Sources of Bias in Data Labeling

Annotator Bias:
Human annotators bring their subjective perspectives, cultural backgrounds, and experiences to the labeling process, which can inadvertently introduce bias. For example:

- Differences in interpretations of ambiguous labels (e.g., determining the tone of a text as “neutral” or “negative”).
- Cultural or linguistic misunderstandings that skew labels for diverse populations.

2) Selection Bias in Datasets:

The data being labeled may not accurately represent the target population. For instance:

- Facial recognition datasets might overrepresent certain demographic groups (e.g., Caucasian faces) while underrepresenting others (e.g., darker skin tones).
- Medical datasets may reflect data from specific geographic regions, ignoring global variations in health conditions.

3) Instructions and Guidelines:

Poorly designed or overly generalized labeling guidelines can lead to inconsistent or biased annotations. When instructions are unclear, annotators may rely on personal judgment, introducing variability[16].

4) Algorithmic Pre-labeling Bias:

In semi-automated labeling workflows, pre-labeling by AI models is common. If the pre-labeling model is already biased, it can propagate and even amplify these biases when human annotators validate the labels.

B. Manifestations of Bias in Labeled Data

Bias can take multiple forms, including:
1) Representation Bias: Underrepresentation or overrepresentation of certain groups or attributes.

2) Stereotyping: Reinforcement of harmful stereotypes, such as associating specific professions or behaviors with particular genders or ethnicities.
3) Annotation Drift: Variability in labeling due to subjective interpretations or fatigue among annotators, especially for repetitive tasks [17].

C. Consequences of Biased Data Labeling

1) Model Performance:

Biased labeled data leads to models that perform poorly for underrepresented groups or scenarios. For instance, biased datasets in facial recognition systems have resulted in disproportionately high error rates for darker-skinned individuals.

2) Ethical and Social Impacts:

Bias in AI systems can perpetuate or exacerbate societal inequities. For example, biased algorithms used in hiring, law enforcement, or credit scoring can unfairly disadvantage marginalized communities.

IV. CASE STUDIES

Real-world case studies provide valuable insights into the ethical challenges and consequences of large-scale data labeling and usage. These examples highlight how biases, privacy violations, labor issues, and lack of transparency manifest in practice, offering lessons for future improvements

While manual audits have provided invaluable insights into the contents of datasets, as datasets swell in size this technique is not scalable. Recent work has proposed algorithmic interventions that assist in the exploration and adjustment of datasets. Some methods leverage statistical properties of datasets to surface spurious cues and other possible issues with dataset contents. The AFLITE algorithm proposed by [18] provides a way to systematically identify dataset instances that are easily gamed by a model, but in ways that are not easily detected by humans. This algorithm is applied by [19] to a variety of NLP datasets, and they find that training models on adversarially filtered data leads to better generalization to outof-distribution data. In addition, recent work by Swayamdipta et al.76 proposes methods for performing exploratory data analyses based on training dynamics that reveal edge cases in the data, bringing to light labeling errors or ambiguous labels in datasets. [20] combine an algorithmic approach with human validation to surface labeling errors in the test set for ImageNet.

After CLIP's initial success, ALIGN and BASIC improved contrastive multimodal learning by increasing the training set size and the batch size used for training [21, 22]. also increased training scale and experimented with a combination of pre-trained image representations and contrastive fine-tuning to connect frozen image representations to text [23]. Flamingo introduced the first large vision-language model with in-context learning [24]. Other papers have combined contrastive losses with image captioning to further improve performance [25]. Beyond image classification and retrieval, the community later adapted CLIP to further vision tasks such as object navigation and visual question answering [26].

A. Bias in Facial Recognition Systems

A study revealed significant biases in commercially available facial recognition systems from major companies[27]. The research found that:

- The error rate for identifying darker-skinned women was as high as 34.7%, compared to 0.8% for lighter-skinned men.
- The datasets used to train these systems were predominantly composed of lighter-skinned and male faces, leading to skewed model performance.

1) Ethical Implications:

- The biased performance highlighted the lack of diverse representation in training datasets.
- This case raised awareness about the societal impacts of deploying biased AI systems, particularly in law enforcement and public surveillance.

2) Lessons Learned:

- Incorporate diverse datasets during the data labeling phase to ensure equitable model performance.
- Regularly audit datasets for demographic biases.

B. Privacy Violations in Voice Data Collection

1) Case Study:

The stude revealed that contractors reviewing voice assistant recordings had access to sensitive user data. These recordings included private conversations inadvertently captured by the devices [28].

- Contractors reported hearing sensitive information, including medical details and personal conversations.
- Users were unaware that their interactions with voice assistants could be reviewed by humans.

2) Ethical Implications:

- Lack of transparency about data usage violated users' privacy expectations.
- Insufficient anonymization measures exposed individuals to potential privacy breaches.

3) Lessons Learned:

- Implement explicit consent mechanisms for data usage.
- Use privacy-preserving methods, such as differential privacy, during data labeling and review processes.

C. Labor Exploitation in Crowdsourced Data Labeling

1) Case Study:

Crowdsourcing platforms like Amazon Mechanical Turk (MTurk) rely on a global workforce to perform data labeling tasks [29].

- Reports indicate that many workers earn less than the U.S. federal minimum wage, with some tasks paying as little as $2 per hour.
- Workers lack benefits, job security, and protection from tasks involving harmful or explicit content.

2) Ethical Implications:

- Exploitation of labor undermines the ethical principles of fairness and dignity.
- Exposure to harmful content without mental health support can lead to psychological distress.

3) Lessons Learned:

- Establish fair wage policies and worker protections on crowdsourcing platforms.
- Offer mental health resources and limit exposure to explicit content.

D. Misinformation in Content Moderation

1) Case Study:

[30] content moderation system, heavily reliant on labeled datasets, struggled to identify and remove misinformation and hate speech during global elections and the COVID-19 pandemic.

- Reports highlighted that outsourced moderators faced high workloads, low pay, and inadequate training.
- Automated systems were unable to capture nuanced cultural or contextual variations in content.

2) Ethical Implications:

- Inadequate moderation contributed to the spread of misinformation and polarization.
- Exploitation of moderators raised concerns about labor ethics.

3) Lessons Learned:

- Invest in culturally sensitive data labeling and moderation practices.
- Balance automated systems with human oversight to improve accuracy.

E. Dataset Transparency Issues

1) Case Study: ImageNet Dataset Controversy

ImageNet, one of the most widely used datasets in AI research, came under scrutiny for including problematic labels and categories [31].

- Certain images were labeled with offensive or discriminatory terms.
- The dataset lacked transparency regarding its labeling processes and guidelines.

2) Ethical Implications:

- Offensive labels perpetuate harm and bias in downstream applications.
- Lack of transparency hindered accountability and trust in the dataset.

3) Lessons Learned:

- Adopt standardized documentation practices like "Datasheets for Datasets" to enhance transparency.
- Regularly review and refine labeling guidelines to align with ethical standards.

4) Current Solutions and Best Practices:

- AI techniques to detect and mitigate bias.
- Privacy-preserving data labeling methods (e.g., Federated Learning, Differential Privacy).
- Improving labor conditions in crowdsourced data labeling.
- Legal frameworks and policies (e.g., GDPR, CCPA).

V. CHALLENGES AND FUTURE DIRECTIONS

The ethical challenges surrounding large-scale data labeling are multifaceted and require ongoing attention from researchers, practitioners, policymakers, and organizations. As AI technologies continue to evolve and expand, addressing these issues effectively will be crucial for ensuring that AI systems are equitable, transparent, and responsible. This section explores the key challenges in improving data labeling practices and identifies promising future directions for research and practice.

A. Scalability of Ethical Solutions

1. Challenge: Many of the ethical solutions, such as bias audits, privacy-preserving techniques, and fair labor practices, are often implemented at a small scale or in pilot projects. Scaling these solutions to accommodate the growing volume of data and the complexity of AI systems remains a significant challenge.
2. Impact: Without scalable solutions, AI development could continue to perpetuate biases, privacy violations, and exploitation, especially in resource-constrained environments.

B. Ensuring Diversity in Data Labeling

1) Challenge: Achieving sufficient diversity in data labeling processes to prevent bias is difficult due to the global nature of AI development. Annotators from different demographic and cultural backgrounds may be needed, but attracting a diverse workforce that reflects all segments of society remains a challenge.

2) Impact: Inadequate diversity in labeling can lead to skewed datasets that fail to generalize well to underrepresented populations, leading to discriminatory outcomes.

C. Maintaining Annotator Well-Being

1) Challenge: Many data labeling tasks involve repetitive or emotionally taxing work, such as labeling explicit content, which can lead to burnout or psychological harm among workers.

2) Impact: Poor working conditions may discourage workers from participating, leading to a lack of workforce availability, while also raising ethical concerns about labor exploitation.

D. Privacy Protection and Data Security

1) Challenge: Collecting, storing, and processing large datasets often involves sensitive personal information. Despite advances in privacy-preserving technologies like differential privacy and federated learning, fully ensuring the privacy and security of individuals' data during labeling remains challenging.

2) Impact: Privacy violations could undermine public trust in AI systems, particularly in sectors like healthcare, finance, and criminal justice where sensitive data is commonly used.

E. Lack of Transparency in Labeling Processes

1) Challenge: Transparency in data labeling practices is often inadequate. In many cases, the origins of the data, the guidelines followed by annotators, and the decision-making processes in model development are not fully disclosed.

2) Impact: Lack of transparency can hinder accountability and perpetuate unethical practices, making it difficult to identify and address errors or biases in AI systems [32].

©IJRASET: All Rights are Reserved

VI. FUTURE DIRECTIONS

A. *Improved Ethical Guidelines and Standards*

1) Direction: The development of comprehensive ethical guidelines and industry standards for data labeling is crucial. These standards should address issues like fairness, transparency, privacy, and labor rights and provide clear frameworks for responsible data labeling practices.
2) Impact: Well-defined guidelines can promote consistency and accountability across the AI development process, ensuring that ethical concerns are systematically addressed.

B. *5.7.Advances in Bias Mitigation Techniques*

1) Direction: Research into advanced bias detection and mitigation algorithms is needed to identify and correct biases in datasets during the labeling process. Techniques such as adversarial debiasing, fairness constraints in optimization, and data augmentation could be employed to create more balanced datasets.
2) Impact: More equitable datasets would lead to AI models that perform better across all demographic groups, reducing the risk of discriminatory outcomes.

C. *5.8.Leveraging Decentralized and Federated Learning*

1) Direction: Federated learning, a decentralized AI training method, offers the potential to improve data privacy by enabling AI models to be trained on local data without it ever leaving the user's device. This can be especially important for sensitive data, such as medical or financial records.
2) Impact: By reducing the need for centralized data storage, federated learning can mitigate privacy risks and empower users to retain control over their personal information.

D. *5.9.Crowdsourcing and Workforce Empowerment*

1) Direction: Future efforts should focus on improving the working conditions of data annotators by ensuring fair compensation, job security, and psychological support. Additionally, annotator training programs that promote cultural competency and reduce subjectivity in labeling can enhance data quality.
2) Impact: A well-supported workforce will lead to better quality labels, while ethical treatment of workers will address labor exploitation concerns.

E. *5.10.Explainability and Transparency in AI Systems*

1) Direction: Enhancing transparency in AI systems requires better documentation practices, such as "Datasheets for Datasets" and "Model Cards," which provide detailed information about the datasets used, the data labeling process, and the performance characteristics of AI models.
2) Impact: Increased transparency will foster trust and allow stakeholders to identify biases or errors in models more easily, leading to more responsible AI deployment.

F. *5.11.Ethical AI Governance and Regulation*

1) Direction: Governments and international organizations must establish regulations for data collection, labeling, and AI development to ensure that ethical standards are met. These regulations should focus on issues like data privacy, labor rights, and algorithmic accountability, with specific guidelines for industries like healthcare, finance, and criminal justice.
2) Impact: Clear regulations will ensure that AI systems are developed with societal well-being in mind, promoting ethical practices on a global scale.

G. *5.12.Public Awareness and Engagement*

1) Direction: Increasing public awareness of the ethical challenges in data labeling and AI can drive greater demand for responsible AI practices. Engaging the public in discussions about data privacy, bias, and transparency will foster a more informed society that can hold both corporations and governments accountable.
2) Impact: A well-informed public will be more likely to advocate for ethical data labeling practices and demand accountability from AI developers.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VII. CONCLUSION

Addressing the challenges of large-scale data labeling is vital for the future of AI systems. By pursuing these future directions—through the development of ethical standards, advances in bias mitigation, and improvements in labor conditions—society can ensure that AI technologies evolve in a way that benefits everyone, without compromising fairness, privacy, or transparency. The path forward requires collaboration across academia, industry, government, and civil society to create AI systems that are responsible, ethical, and socially just.

The ethical challenges of large-scale data labeling are complex, but they are not insurmountable. Through collaborative effort, technological innovation, and unwavering commitment to fairness, privacy, and labor rights, we can shape the future of AI in a way that is both technically sophisticated and socially responsible. Addressing these challenges is crucial for building AI systems that are trustworthy, equitable, and aligned with the values of a diverse and interconnected global society

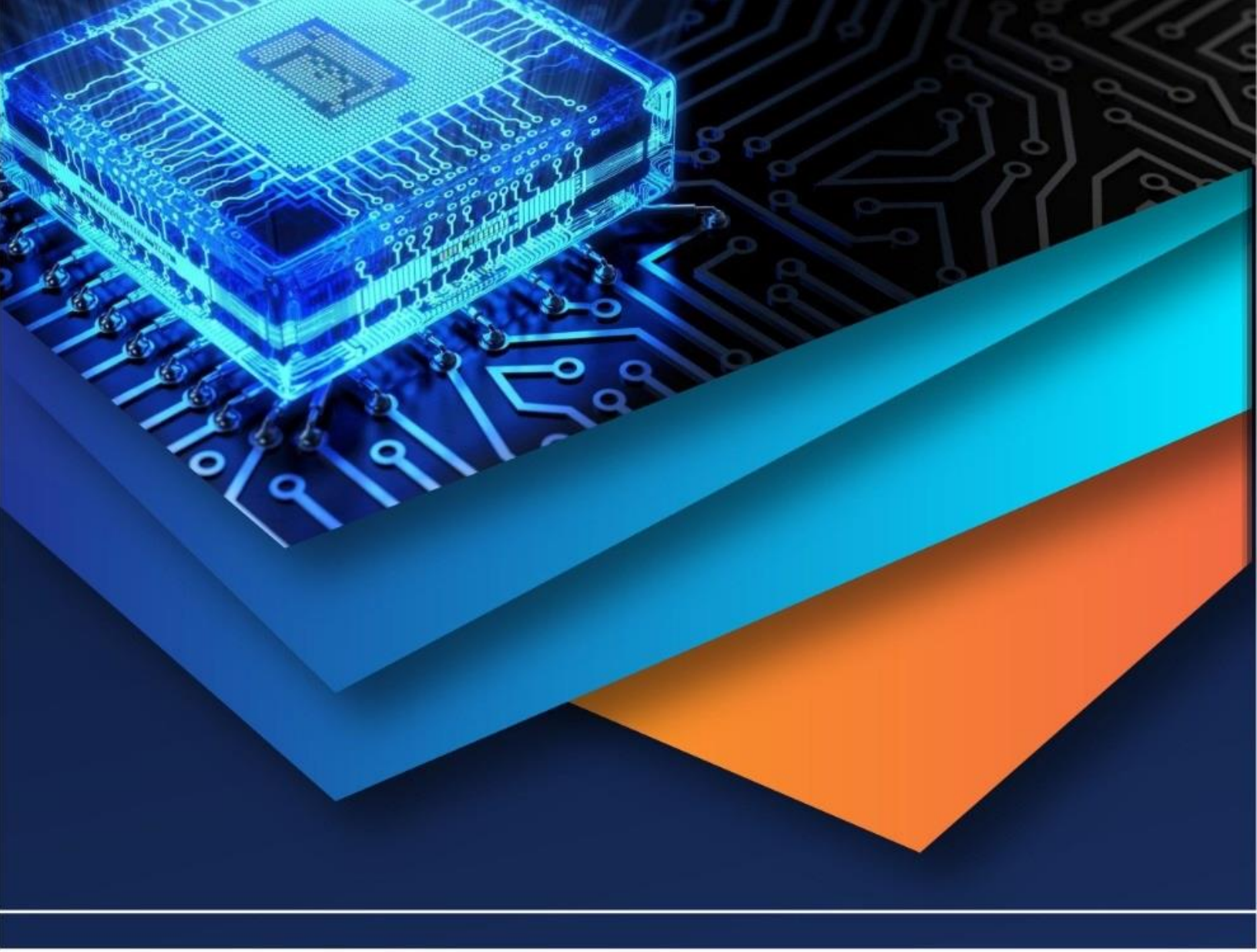
REFERENCES

[1] Kron J. Are there legal risks for doctors when they prescribe drugs off-label? Australian Doctor. 2005 13 January.
[2] Gazarian M, Kelly M, McPhee JR, Graudins LV, Ward RL, Campbell TJ. Off-label use of medicines: consensus recommendations for evaluating appropriateness. Med J Aust. 2006;185(10):544.
[3] Medicines, Poisons and Therapeutic Goods Act, ACT (2008).
[4] Kelly M, Gazarian M, McPhee J. Off-lable Prescribing. Australian Prescriber 2005;28:5-7.
[5] Therapeutic Advisory Group. Off-Label Use of Registered Medicines and Use of Medicines under the Personal Importation Scheme in NSW Public Hospitals: A Discussion Paper. NSW Department of Health; 2003.
[6] O'Donnell CPF, Stone RJ, Morley CJ. Unlicensed and Off-Label Drug Use in an Australian Neonatal Intensive Care Unit. Pediatrics. 2002;110(5):e52.
[7] Woodcock J. Implementation of the Pediatric Exclusivitiy Provisions. 2001 [cited 2012 February 26]; Available from: http://www.fda.gov/NewsEvents/Testimony/ucm115220.
[8] Jeffrey S Anastasi and Matthew G Rhodes. An own-age bias in face recognition for children and older adults. Psychonomic bulletin & review, 12(6):1043–1047, 2005.
[9] Jeffrey S Anastasi and Matthew G Rhodes. Evidence for an own-age bias in face recognition. North American Journal of Psychology, 8(2), 2006.
[10] Nazanin Andalibi and Justin Buss. The human in emotion recognition on social media: Attitudes, outcomes, risks. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pages 1–16, 2020.
[11] Jerone T. A. Andrews. The hidden fingerprint inside your photos. https://www.bbc.com/ future/article/20210324-the-hidden-fingerprint-inside-your-photos, 2021. [Accessed June 30, 2022].
[12] Jerone T A Andrews, Przemyslaw Joniak, and Alice Xiang. A view from somewhere: Humancentric face representations. In International Conference on Learning Representations (ICLR), 2023.
[13] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele. 2d human pose estimation: New benchmark and state of the art analysis. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 3686–3693, 2014.
[14] McKane Andrus, Elena Spitzer, and Alice Xiang. Working to address algorithmic bias? don't overlook the role of demographic data. Partnership on AI, 2020.
[15] McKane Andrus, Elena Spitzer, Jeffrey Brown, and Alice Xiang. What we can't measure, we can't understand: Challenges to demographic data procurement in the pursuit of fairness. In ACM Conference on Fairness, Accountability, and Transparency (FAccT), pages 249–260, 2021.
[16] Sweet M. Gabapentin documents raise concerns about off-label promotion and prescribing. Australia Prescriber. 2003;26(1):18-9.
[17] Kaiser Foundation Health Plan Inc v Pfizer Inc. United States District Court of Massachusetts; 2010.
[18] Medicines Australia. Code of Conduct. Deakin: Medicines Australia; 2011 [cited 2011 Sep 30]; Available from: http://medicinesaustralia.com.au/code-of-conduct/.
[19] Department of Health, Education, and Welfare. The National Commission for the Protection of Human Subjects of Biomedical and Behavioural Research. 1979 [cited 2011 Oct 1]; Available from: http://ohsr.od.nih.gov/guidelines/belmont.html#goa.
[20] Australian Government. National Statement on Ethical Conduct in Human Research. In: National Health and Medical Research Council, editor. Canberra: Australian Government; 2007.
[21] Miner J, Hoffhines A. The discovery of aspirin's antithrombotic effects. Tex Heart Inst J 2007;34(2):179-86. [22] Craven LL. Acetylsalicylic acid, possible preventive of coronary thrombosis. Ann West Med Surg 1950;4(2):95.
[22] Alistair Barr. Google mistakenly tags Black people as 'gorillas,' showing limits of algorithms. The Wall Street Journal, 2015.
[23] Teanna Barrett, Quan Ze Chen, and Amy X Zhang. Skin deep: Investigating subjectivity in skin tone annotations for computer vision benchmark datasets. arXiv preprint arXiv:2305.09072, 2023.
[24] Tom Beauchamp and James Childress. Principles of biomedical ethics: marking its fortieth anniversary, 2019.
[25] Fabiola Becerra-Riera, Annette Morales-González, and Heydi Méndez-Vázquez. A survey on facial soft biometrics for video surveillance and forensic applications. Artificial Intelligence Review, 52(2):1155–1187, 2019.
[26] Sara Beery, Grant Van Horn, and Pietro Perona. Recognition in terra incognita. In European Conference on Computer Vision (ECCV), pages 456–473, 2018.
[27] Emily M. Bender and Batya Friedman. Data statements for natural language processing: Toward mitigating system bias and enabling better science. Transactions of the Association for Computational Linguistics, 6:587–604, December 2018. doi: 10.1162/tacl_a_00041.
[28] M. Minsky and S. Papert, "An introduction to computational geometry," Cambridge tiass., HIT, 1969.
[29] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," Int. J. Artif. Intell. Expert Syst., vol. 1, no. 4, pp. 111–122, 2011.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

[30] P. Sibi, S. A. Jones, and P. Siddarth, "Analysis of different activation functions using back propagation neural networks," J. Theor. Appl. Inf. Technol., vol. 47, no. 3, pp. 1264–1268, 2013.
[31] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," arXiv Prepr. arXiv1505.00853, 2015.
[32] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv Prepr. arXiv1609.04747, 2016.
[33] M. W. Gardner and S. R. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," Atmos. Environ., vol. 32, no. 14–15, pp. 2627–2636, 1998.
[34] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," J. Mach. Learn. Res., vol. 13, no. Feb, pp. 281–305, 2012.
[35] D. Maclaurin, D. Duvenaud, and R. Adams, "Gradient-based hyperparameter optimization through reversible learning," in International Conference on Machine Learning, 2015, pp. 2113–2122.
[36] J. W. Tweedale, "An Application of Transfer Learning for Maritime Vision Processing Using Machine Learning," in International Conference on Intelligent Decision Technologies, 2018, pp. 87–97.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)