Strathmore Law Review (SLR)
eISSN: 2414-8164 – Volume 9, Issue 1 (2024) p.13-36
https:doi.org/10.52907/slr.v9i1.480

Algorithmic Bias in Hiring Algorithms: A Kenyan Perspective

Kazungu Mrashui[*]

Abstract

The use of machine learning algorithms has permeated into nearly all aspects of life. With this steady integration, tasks previously handled by humans are increasingly falling into the 'hands' of machines. Ideally this would be celebrated as a great improvement for mankind. Tasks that were previously riddled with human bias such as hiring would now be performed by an 'omniscient algorithm' that could harbor no bias therefore resulting in fair outcomes for the previously oppressed. However, this is not the case. The integration of machine learning algorithms in the hiring process risks further exacerbating existing bias that was prevalent or introducing new data-driven bias. The question then is how to contend with this novel form of discrimination: algorithmic discrimination. The answer to combating algorithmic discrimination is algorithmic fairness. The goal should not be to create 'fair' algorithms but rather to detect and mitigate fairness-related harms as much as possible. By doing so, a balance can be struck between the competing interests of innovation and employee rights. This article demonstrates that algorithmic discrimination during hiring is a real threat to the Kenyan jobseeker. Although this form of discrimination can be addressed by Kenyan law, more needs to be done to detect and mitigate fairness-related harms as much as possible.

Key Words: *Machine Learning (ML), discrimination, algorithmic fairness, Artificial Intelligence (AI), Labour Law*

[*] The author holds a Bachelor of Laws (LLB) degree with Second Class Honours (Upper Division) from Strathmore Law School (Nairobi, Kenya), where he cultivated a strong interest in the intersection of law, technology, and policy. Currently serving as a Tax Associate at Deloitte, he specializes in tax compliance and transfer pricing. His professional and academic pursuits center on the convergence of law, technology, and policy, with a particular emphasis on tackling emerging legal challenges in the digital age. His research focuses on the critical issue of algorithmic bias in hiring processes within the Kenyan context, underscoring his dedication to promoting equitable solutions in the digital era.

I. Introduction

The use of technology, more specifically Machine Learning (ML) is permeating rather quickly into most if not all avenues of life.[1] ML has aptly been defined as a unified algorithmic framework designed to identify computational models that accurately describe empirical data and the phenomena underlying it, with little or no human involvement.[2] Simply, ML is a branch of AI that refers to computer algorithms that detect patterns in data to produce a desired output and automatically improve their own performance over time.[3]

The employment space is an area where ML algorithms are being quickly integrated into normal business operations.[4] This use of ML in labour relations has been termed by some scholars as *Algorithmic Management*.[5] This term was initially coined to illustrate the manner in which gig-working platforms made use of ML in their business operations, however, it has also been used to describe the incorporation of ML in traditional forms of employment.[6] Algorithmic Management in traditional forms of employment entails the use of various ML tools to assist in making autonomous or semi-autonomous decisions in the processes of *recruitment, monitoring and allocation of work* as well as *rating and dismissal*.[7] This study will focus particularly on Algorithmic Management in the recruitment process.

In recent years, simple algorithms have been used in the recruitment process.[8] The most basic and widespread being Curriculum Vitae (CV) checkers

1 Mihajlovic Ilija, 'How artificial intelligence is impacting our everyday lives' Towards Data Science, 13 June 2019 -<https://towardsdatascience.com/how-artificial-intelligence-is-impacting-our-everyday-lives-eae3b63379e1> on 1 January 2022.

2 Watt J, Borhani R, Katsaggelos A, *Machine learning redefined*, Cambridge University Press, Cambridge UK, 2020, 1.

3 Surden H, 'Machine learning and law: An overview' in Vogl R (ed) Research Handbook on Big Data Law, Edward Elgar Publishing, Cheltenham UK, 2021, 171 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3686055> on 1 January 2022.

4 Alexandra M and Nguyen A, 'Algorithmic management in the workplace' Data and Society, February 2019 -< https://datasociety.net/wp-content/uploads/2019/02/DS_Algorithmic_Management_Explainer.pdf> on 1 January 2022.

5 Alexandra M and Nguyen A, 'Algorithmic management in the workplace',3.

6 Gig work as the name suggests is work done as 'gigs' where the conventional structure of management is absent. In most gig working platforms, the 'manager' is the algorithm. Gig working platforms connect freelancers with short-term or task-based work hence the term gig. Some well known gig working platforms include Uber, Bolt and Fiverr.

7 Alexandra M and Nguyen A, 'Algorithmic management in the workplace',4.

8 Buranyi S, ''Dehumanising, Impenetrable, Frustrating' : The Grim Reality of Job Hunting in the Age of AI', The Guardian London, 4 March 2018 - <https://perma.cc/B94W-MPCC> on 1 January 2022.

which sift through resumes submitted by prospective employees to identify various keywords deemed advantageous by the employer.[9] The improvement of technology has resulted in the expansion of this concept beyond simple keywords. Algorithmic Management entails using ML tools to mimic human intelligence in the hiring process in order to improve performance and streamline hiring thus ensuring the selection of the best candidates.[10] The ML algorithms achieve this by sorting through data input into the system by either developers or end users.[11] The data fed into the ML tools on prospective employees is broad, ranging from traditional characteristics such as previous places of employment, residence, and alma maters, to more novel criteria such as social media activity and internet footprint.[12] Advocates of this algorithmic hiring process claim that it not only increases efficiency but also reduces unconscious bias present in human managers due to the algorithms use of objective data points.[13]

However, just as human managers are prone to unconscious bias there is the risk that these hiring algorithms could introduce data driven algorithmic bias into the workplace.[14] Data driven algorithmic bias can arise due to the nature of the training data that is input into the hiring algorithm.[15] This data is input by persons who could themselves be inadvertently training the algorithm on biased data.[16] This was observed when Amazon scrapped their hiring algorithm that was discriminating against women.[17] The rationale behind the discrimination was that the algorithm had been trained on a majority of resumes from male candidates.[18] Therefore, the algorithm learned to associate the best candidates as being male, hence inadvertently discriminating against women.

9 Buranyi S, "Dehumanising, Impenetrable, Frustrating": The Grim Reality of Job Hunting in the Age of AI', The Guardian London, 4 March 2018 - <https://perma.cc/B94W-MPCC> on 1 January 2022.

10 Kim P, 'Data-driven discrimination at work' 58 (3) *William & Mary Law Review*, 2017, 862–863<https://scholarship.law.wm.edu/wmlr/vol58/iss3/4/>.

11 Kim P, 'Data-driven discrimination at work', 862 – 863.

12 Kim P, 'Data-driven discrimination at work', 861.

13 Kim P, 'Data-driven discrimination at work', 869.

14 Ajunwa I, 'The paradox of automation as anti-bias intervention' 41 *Cardozo Law Review,* (2020) 1 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2746078>

15 Barocas S and Selbst A, 'Big data's disparate impact' 104 *California Law Review*, 2016, 680-687 - <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899>

16 Barocas S and Selbst A, 'Big data's disparate impact', 680-687.

17 Dastin J, 'Amazon scraps secret AI recruiting tool that showed bias against women' Reuters, 11 October 2018 - <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> on 1 January 2022.

18 Dastin J, 'Amazon scraps secret AI recruiting tool that showed bias against women' Reuters, 11 October 2018 - <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> on 1 January 2022.

Bias can also occur during feature selection where the employer decides on the features to include in the algorithmic hiring tool.[19] When certain features are excluded from such an analysis, there is the possibility of creating a bias known as omitted variable bias.[20] ML models may also cause prejudicial outcomes to prospective employees by considering a person's protected characteristics or using proxies to arrive at those characteristics.[21]

It is evident that algorithmic hiring tools have the potential to introduce new biases into the hiring process or to further exacerbate the existing biases already present in the recruitment process, therefore leading to discrimination. The question is how labour laws and anti-discrimination laws respond to the integration of this technology into the employment space to prevent disadvantageous outcomes for both the prospective employees and employers.

Article 27 of the Constitution of Kenya expressly prohibits both direct and indirect forms of discrimination.[22] In addition, the Constitution also highlights protected characteristics that one cannot discriminate against.[23] Furthermore, the Employment Act prohibits discrimination against a prospective employee during recruitment.[24]

The law on discrimination in Kenya therefore appears broad enough to encompass any discrimination or bias that can arise from ML hiring tools. The difficulty for a prospective employee however lies in trying to prove that algorithmic discrimination has occurred.[25] To sustain the claim of discrimination before the courts, a prospective employee must be able to prove a *prima facie* case of discrimination, which would then shift the burden to the employer to prove that their conduct was not discriminatory.[26] Due to the complexity of ML hiring tools it would be extremely difficult for a prospective employee to realize that they have been discriminated against, let alone prove a *prima facie* case of discrimination.[27]

Currently, the Kenyan courts have not had to assess any issues pertaining to algorithmic bias in the hiring process. The use of algorithms in other facets of

19 Barocas S and Selbst A, 'Big data's disparate impact', 688-690.

20 Kim P, 'Data-driven discrimination at work', 878.

21 Kim P, 'Data-driven discrimination at work', 890.

22 Article 27, *Constitution of Kenya* (2010).

23 Article 27, *Constitution of Kenya* (2010).

24 Section 5, *Employment Act* (Act No. 11 of 2007).

25 Kelly-Lyth A, 'Challenging biased hiring algorithms' *Oxford Journal of Legal Studies*, 2021, 23, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3744248>

26 *GMV v Bank of Africa Kenya Limited* [2013] eKLR.

27 Kelly-Lyth A, 'Challenging biased hiring algorithms', 23.

labour law has however come under scrutiny before the courts. In a case decided by the Employment and Labour Relations Court, ABSA Bank was ordered to compensate a former employee for unfair dismissal after the bank's performance management system unfairly rated the employee as a poor performer for two years which resulted in her dismissal.[28] This is a clear indicator of the steady integration of technology within the labour market in Kenya. Furthermore, with multinational companies with a presence in Kenya such as Unilever, PwC and G4S already making use of hiring algorithm HireVue, it is important that the law and policy on proving and preventing algorithmic bias develops in Kenya.[29]

The major issue then arising from the use of ML algorithms in hiring is its opacity;[30] opacity in the sense that it is difficult to understand how the algorithm arrives at a certain hiring decision.[31] This is a direct challenge not only for prospective employees who have limited to no knowledge on ML but also for the developers of ML algorithms themselves.[32] The concern is in proving that discrimination has occurred and consequently preventing any further discrimination from materializing. The opaque nature of hiring algorithms also poses a significant challenge to the courts in adjudicating on matters that eventually come before them.[33] This article will demonstrate that algorithmic discrimination during hiring is a real threat to the Kenyan jobseeker and to that effect it will provide solutions to aid in mitigating algorithmic discrimination.

There is limited scholarly research on ML in Kenya, particularly concerning discrimination and labour law. However, Cecil Abungu has written on how algorithmic decision making and discrimination can be handled in developing countries so as not to leave protected groups prone to unlawful discrimination.[34] In his study he speaks to the fact that discrimination by machine learning algorithms is more concerning than human discrimination due to the larger scale

28 Muthoni K, 'Banker who was sacked 'by computer system' awarded Sh 6.5 million' The Standard, 28 July 2021 - <https://www.standardmedia.co.ke/business/business/article/2001419375/banker-who-was-sacked-by-computer- system-awarded-sh-65-million> on 22 August 2022.

29 Ribeiro J, '5 companies that are revolutionizing recruiting using AI', Medium, 14 November 2020, - <https://medium.com/tech-cult-heartbeat/5-companies-that-are-revolutionizing-recruiting-using-artificial-intelligence-9a70986c7a7e> on 1 January 2022.

30 Kelly-Lyth A, 'Challenging biased hiring algorithms', 21.

31 Raden N, 'The problem of algorithmic opacity, or "What the heck is the algorithm doing?"' Diginomica, 13 January 2021 - <https://diginomica.com/problem-algorithmic-opacity-or-what-heck-algorithm-doing> on 22 August 2022.

32 Kim P, 'Data-driven discrimination at work', 992.

33 Kelly-Lyth A, 'Challenging biased hiring algorithms', 15.

34 Abungu C, 'Algorithmic decision-making and discrimination in developing countries', Journal of Law, Technology & the Internet, 2021,11<https://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?article=1135&context=jolti>

that AI operates on.[35] He also notes that algorithmic opacity poses a challenge when it comes to proving that unlawful discrimination has occurred.[36] In his work he asserts that most of the proposed solutions to algorithmic bias emanating from developed nations would be difficult to replicate in developing nations because such countries lack the necessary foundation. He defines the necessary foundation as one with; a well rooted culture of transparency and statistical analysis of the disparities faced by protected groups, vigilant non-governmental actors alive to the risks of algorithmic discrimination and a reasonably robust and proactive executive branch or independent office policing of discrimination.[37]

Building from that, this study adds to the body of knowledge by drawing attention to the imminent threat that algorithmic discrimination poses to the Kenyan jobseeker as well as showcasing the challenges of proving and preventing algorithmic discrimination. This study also proposes recommendations that can be implemented both in law and policy to create a more transparent ML environment. This transparency will significantly reduce the difficulty of proving and preventing algorithmic bias.

II. Biased Algorithms in the recruitment process

i. Understanding Machine Learning

Machine Learning has aptly been defined as a unified algorithmic framework designed to identify computational models that accurately describe empirical data and the phenomena underlying it, with little or no human involvement.[38] ML focuses on the question of how to get computers to program themselves from experience as well as initial software.[39] ML models are mainly used in automation and prediction.[40]

From the definitions, it is evident that ML involves the use of algorithms that process input data to produce a desired output. The algorithms are sets of instructions that enable the machine to sift through large volumes of data and

35 Abungu C, 'Algorithmic decision-making and discrimination in developing countries' 8.

36 Abungu C, 'Algorithmic decision-making and discrimination in developing countries' 15.

37 Abungu C, 'Algorithmic decision-making and discrimination in developing countries' 2.

38 Watt J, Borhani R, Katsaggelos A, *Machine learning redefined*, Cambridge University Press, Cambridge UK, 2020, 1.

39 Mohammed M, Khan M, Bashier E, *Machine learning: Algorithms and application,* CRC Press, London UK, 2017, 5.

40 Surden H, Machine learning and law, 171.

identify various correlations thus resulting in a desired outcome.[41] The algorithms are written with the aim of making machines 'learn' so that they can produce desired outcomes without further human interference.[42]

This form of 'learning' is different from learning displayed by humans.[43] In his work, Surden alludes to this when he opines that even the most advanced systems today lack the higher-order cognitive skills routinely displayed by humans, such as abstract reasoning, general learning or flexible problem-solving.[44] He then goes further to state that ML is narrowly constrained to well defined tasks.[45] A computer program is said to learn from an experience (E) with respect to some class of tasks (T) and performance measure (P), if its performance at tasks in T, as measured by P, improves with experience E.[46]

The performance measure P is a quantitative measure of a ML model's performance of a certain task T.[47] To test the performance of a ML model, the program is tested using data that it has not interacted with before, known as a 'test set'.[48] The purpose of this is to simulate real world scenarios.[49] From the output derived from the test set of data, software engineers create suitable metrics for measuring performance. Different tasks typically require different performance measures. Performance could therefore be measured in terms of accuracy, error rate or precision based on the task at hand.[50]

The experience E is based on the nature of data that the model is exposed to for it to undertake task T.[51] Data can either be labeled, unlabeled, a mixture of labeled and unlabeled and in some cases a ML model could learn without any input of data.[52] Labeled data is data that includes both the input and a corresponding output. In the hiring space, labeled data includes resumes with corresponding outcomes i.e. to hire or not to hire. Unlabeled data on the other hand is data without the corresponding outcome labels. This nature of

41 Goodfellow I, Yoshua Bengio Y, Courville A, *Deep learning*, MIT Press, Cambridge, 2016, 99.

42 Mohammed *et al*, *Machine learning*, 5.

43 Shetty S, Shetty S, Singh C, Rao A, 'Supervised Machine Learning: Algorithms and Applications' in Singh P (ed) *Fundamentals and Methods of Machine and Deep Learning,* Scrivener Publishing, Beverly MA, 2022, 5. Surden H, Machine Learning and Law, 174.

44 Surden H, Machine learning and law, 172.

45 Surden H, Machine learning and law, 172.

46 Goodfellow *et al*, Deep learning, 99.

47 Goodfellow *et al*, *Deep learning,* 103.

48 Watt *et al*, *Machine learning redefined,* 5.

49 Goodfellow *et al*, *Deep learning,* 104.

50 Goodfellow *et al*, *Deep learning,* 104.

51 Goodfellow *et al*, *Deep learning,* 104.

52 Mohammed *et al*, *Machine learning,* 7.

data categorization gives rise to the following taxonomy of ML techniques:[53] i) Supervised ML which deals with classified (labeled) data, ii) unsupervised ML which deals with unclassified (unlabeled) data, iii) semi-supervised ML which deals with a combination of labeled and unlabeled data, and iv) reinforcement learning which involves learning through interactions with the environment without relying on labeled datasets.

This research will focus on Supervised Machine Learning (SML) because majority of the algorithmic ML tools used in recruitment currently deal with labeled data. SML can further be classified into regression and classification based on whether the data is continuous or categorical.[54] Classification trains the ML model to distinguish between different classes based on the labeled data hence a categorical output of whether to hire or not.[55] Understanding the classification problem is pertinent as this is the problem posed to ML algorithmic hiring tools to distinguish between a 'good' employee that should be hired and those that will be rejected. A discussion on regression, however, is beyond the scope of this research.

The classification problem is then divided into four steps; data collection, feature design, model training and model validation.[56]

a. Data collection

Data collection is the process of gathering large volumes of data that are necessary inputs for creating a ML model.[57] A ML model requires substantial amounts of data during both training and testing phases.

Simple actions like responding to tweets or liking Meta posts could provide valuable data for ML model developers. The data collected must be extensive enough to be divided into training data and test data.[58] These two sets need to be different to properly assess the ML model's performance. Ideally, the larger the data set, the 'smarter' the ML model.[59]

---

53 Mohammed *et al*, *Machine learning,* 5.

54 Shetty *et al*, Supervised machine learning, 6.

55 Watt *et al*, *Machine learning redefined,* 10.

56 Watt *et al*, *Machine learning redefined,* 2.

57 Watt *et al*, *Machine learning redefined,* 2.

58 Goodfellow *et al*, *Deep learning,* 104.

59 Barocas S and Selbst A, 'Big data's disparate impact', 680-687.

b. Feature design

Features are important characteristics drawn from input data.[60] Feature design is the process of selecting and creating features that will be used when training the model.[61] For example, in a ML model tasked with predicting housing prices in different neighborhoods relevant features might include the size of the house, number of bedrooms, proximity to social amenities and general security of the neighborhood. On the other hand, in a hiring model tasked with identifying the most suitable candidate, the features could be the candidate's education level, past job experiences and relevant test scores.

c. Model training

Once the necessary features have been selected during feature design, the subsequent step is to train the ML model. Model training is the process of teaching a ML model to make decisions based on data.[62] During this process the engineers input *training data* into the model and observe the output produced. From this the engineers can adjust certain parameters in the algorithm to minimize the error between the predicted output and the observed output.[63]

d. Model validation

Model validation evaluates a model's performance using a separate set of data known as the validation set.[64] This is data that the model has previously not been exposed to during training. The purpose of this is to ensure that the model can make accurate decisions based on unseen data.[65]

Model validation also assists in detecting instances of overfitting and underfitting. Overfitting occurs when the model is too complex and as a result performs well on the training data but poorly on the validation set.[66] In a sense the model has learned to cram the training data similar to when a person memorizes a reading chart during an eye test at an ophthalmologist. Overfitting can be corrected by reducing the dimensionality of data as well as the features

---

60 Watt *et al*, *Machine learning redefined,* 3.

61 Barocas S and Selbst A, 'Big data's disparate impact', 688.

62 Watt *et al*, *Machine learning redefined,* 4.

63 Watt *et al*, *Machine learning redefined,* 3.

64 Watt *et al*, *Machine learning redefined,* 5

65 Goodfellow *et al*, *Deep learning,* 111.

66 Goodfellow *et al*, *Deep learning,* 111.

used.[67] Underfitting on the other hand occurs when a model is too simple and performs poorly on both sets of data.[68] This can be corrected by increasing the amount of data used in both training and validation.[69]

Model validation answers the question of whether the model has 'learnt' from the experience. Due to the different tasks that different models perform, the evaluation metrics vary. Common evaluation metrics include accuracy, precision and error rate.[70]

ii. Machine Learning and Employment

Technology has transformed how prospective employees apply for jobs and how they are subsequently hired.[71] ML hiring algorithms are being used to predict who will succeed in a given position by comparing applicants against a model employee, i.e. a composite profile created from the attributes of successful employees in similar positions.[72] These ML models automate the hiring process, reducing the workload on human resource departments that would traditionally have to sift through numerous job applications.[73] This automation saves numerous work hours, increases efficiency, reduces administrative costs and ultimately saves businesses money.[74]

There are several types of algorithmic hiring tools currently in use. Ifeoma Ajunwa classifies these tools into two main categories: *off the shelf* algorithms and *bespoke* algorithms. Off the shelf algorithms are pre-designed tools that employers can purchase or license from developers. Bespoke algorithms, on the other hand, are customized to meet the specific needs and specifications of an employer.[75] The use of such ML tools has been heralded by those in the ML space as being not only more efficient but fairer than traditional hiring practices.[76] However, it is essential to acknowledge potential challenges. Critics argue that these tools can perpetuate existing biases if the training data itself is biased. Additionally,

67 Goodfellow *et al*, *Deep learning,* 111.

68 Goodfellow *et al*, *Deep learning,* 111.

69 Goodfellow *et al*, *Deep learning,* 112.

70 Goodfellow *et al*, *Deep learning,* 105

71 Bornstein',Antidiscriminatory algorithms' 530.

72 Bornstein',Antidiscriminatory algorithms' 531.

73 O'Neil C',Personality tests are failing Americans at work' 18 January 2021 - <https://www.bloomberg.com/opinion/articles/2018-01-18/personality-tests-are-failing-american-workers> on 22 December 2022.

74 O'Neil C, *Weapons of Math Destruction,* Penguin Books, Crown Publishers, 2017, 103.

75 Ajunwa I, 'The paradox of automation as anti-bias intervention' 18.

76 Kim P, 'Data-driven discrimination at work' 869.

there are concerns about transparency and accountability in the decision-making processes of ML algorithms.

iii. How algorithmic bias occurs in hiring

Ideally, the use of ML hiring tools in the recruitment process should largely reduce bias present during recruitment.[77] However, there is the risk that these ML tools could further entrench existing bias or introduce data-driven discrimination. Ifeoma Ajunwa refers to this as a legal paradox.[78] In her work she mentions that what makes algorithmic bias particularly jarring is that whereas one biased human manager is constrained within a specific area, the potential adverse effect of a biased hiring algorithm could be faced by numerous people across various fields.[79] She states that this is due to the volume, velocity, and variety of data used in automated hiring that has the potential to magnify and multiply any bias held by one human manager.[80] Cathy O'Neil writing on the same issue posits that what is particularly alarming is the opacity of such hiring models. She terms their operations as invisible to all but mathematicians and computer scientists. She states that their verdicts, even when harmful or wrong, are beyond dispute or appeal, punishing the poor and oppressed in society while benefiting the rich.[81] Cathy O'Neil terms such models as 'Weapons of Math Destruction' due to their opaque and disruptive nature.[82]

As discussed, the form of ML incorporated in the hiring process is SML and more specifically classification. When deciding what constitutes the most suitable candidate for a job, the features to take into consideration are subjective.[83] Training a ML tool on subjective metrics is what creates an opportunity for data-driven discrimination to occur.[84] In their work, Barocas and Selbst provide an analysis of the ways in which algorithmic bias can present itself.[85] They expound on the following five ways in which discrimination can arise from data:[86] i) when defining the target variable and class labels, ii) from the training data itself, iii) during feature selection, iv) through the use of proxies and v) through masking.

77 Kim P, 'Data-driven discrimination at work', 869-874.

78 Ajunwa I, 'The paradox of automation as anti-bias intervention' 3.

79 Ajunwa I, 'The paradox of automation as anti-bias intervention' 8.

80 Ajunwa I, 'The paradox of automation as anti-bias intervention' 8.

81 O'Neil C, *Weapons of math destruction,* 11.

82 O'Neil C, *Weapons of math destruction,* 11.

83 Barocas S and Selbst A, 'Big data's disparate impact', 679.

84 Barocas S and Selbst A, 'Big data's disparate impact', 679.

85 Barocas S and Selbst A, 'Big data's disparate impact', 677-692.

86 Barocas S and Selbst A, 'Big data's disparate impact', 677-692.

a. Defining the target variable and class labels

Defining the target variable and class labels is the process by which data miners and computer scientists assign importance to various aspects of data obtained about prospective employees.[87] In essence this is the process of defining what makes a 'good' employee.[88] At this stage the employer defines a good employee on subjective metrics based on the needs of their business. These needs could be longer job tenure, higher sales, previous employment records or shorter production times.

The issue arises when the class labels replicate a human bias that was present in traditional hiring.[89] For example, using length of job tenure as the benchmark for a good employee. This criterion could make the ML model discriminate against woman if the data shows that women are more prone to shorter job tenure due to factors such as pregnancy. This could therefore result in a situation where the ML algorithmic hiring tools disregards women candidates entirely due to this correlation.

Another example could be where the shorter production times are systemically linked to male employees. This could once again result in discriminatory outcomes to women candidates who could perform equally as well as their male counterparts.

b. Training data

In SML the ML model learns through the training data.[90] If the training data is biased then the ML model could replicate biased outcomes based on the nature of data.[91] The nature of data in this case could be data obtained through situations where human discrimination played a part in the hiring decision. If such data is used to train the model, there is a high probability that the ML model will learn from the data and continue to replicate the bias present in the training data thus resulting in discriminatory outcomes for protected classes of people.[92] A great (or not so great) example of skewed training data in operation was observed in England at the St. Georges Hospital. The hospital developed an automated system of sorting through medical school applications that used its previous admission decisions as training data. It later turned out that this

87 Barocas S and Selbst A, 'Big data's disparate impact', 677.

88 Barocas S and Selbst A, 'Big data's disparate impact', 677.

89 Ajunwa I, 'The paradox of automation as anti-bias intervention' 1.

90 Mohammed *et al*, *Machine learning,* 9.

91 O'Neil C, *Weapons of math destruction,* 129.

92 Ajunwa I, 'The paradox of automation as anti-bias intervention' 14.

automated model was biased against women and racial minorities. This is because the model relied on historical data from previous admissions as a benchmark, during which the hospital had predominantly admitted white male candidates into its program. It therefore associated the best candidates as being white males hence replicating an existing bias in the hospital's admissions process.[93]

The training data could also result in discriminatory outcomes if the data under-represents or over-represents a certain class of people. Under-representation in training data mainly occurs for classes of people who have systemically been neglected in data collection due to poverty, geography or lifestyle.[94] Professor Kate Crawford defines this lack of data on certain classes of people as 'dark zones' in big-data sets.[95] She gives the example of Street Bump, a mobile application based in Boston that uses smart phones to detect when a driver drives over a pothole and subsequently reports the same to the city.[96] She found that reporting would be significantly hampered by the ownership of smartphones across the city. In more affluent areas, where nearly everyone owns a smartphone, reporting would be significantly higher than in poorer areas, where the rate of smartphone ownership was considerably lower. As a result, the same potholes would go unreported and the roads in poorer areas of the city would be worse off than their more affluent counterparts; a result of under-representation in the data.[97]

Over-representation on the other hand creates a situation like that of St George's hospital where the training data is heavily populated with a specific class of people such that other classes find it nearly impossible to break through. Biased training data leads to discriminatory models.[98]

c. Feature selection

During feature selection the developer decides which features to assign within each of the class labels that have been previously identified as constituting a good employee.[99] Barocas and Selbst explain that this choice of attributes can

93 O'Neil C, *Weapons of math destruction,* 101-104.

94 Barocas S and Selbst A, 'Big data's disparate impact', 684.

95 Crawford K, 'Think again: Big data' Foreign Policy 10 May 2013 - <https://foreignpolicy.com/2013/05/10/think- again-big-data/> on 22 December 2022.

96 Crawford K, 'Think again: Big data' Foreign Policy 10 May 2013 - <https://foreignpolicy.com/2013/05/10/think- again-big-data/> on 22 December 2022.

97 Crawford K, 'Think again: Big data' Foreign Policy 10 May 2013 - < https://foreignpolicy.com/2013/05/10/think- again-big-data/> on 22 December 2022.

98 Barocas S and Selbst A, 'Big data's disparate impact', 687.

99 Barocas S and Selbst A, 'Big data's disparate impact', 688.

have serious implications for the treatment of protected classes. If the attributes that explain variation within a protected class are not incorporated, the model may be unable to distinguish among members of the group, leading it to rely on broad generalizations that disadvantage individual members of the group.[100] Bias also arises if certain protected characteristics are intentionally left out during feature selection.[101] Pauline Kim terms this type of bias as *omitted variable bias*.[102]

d. Proxies

Bias can occur when the data that is used in making algorithmic hiring decisions serves as a proxy for membership in a protected class such as gender or race.[103] The data used at arriving at hiring decisions could appear neutral in the sense that it does not include candidate's protected class as a metric for hiring. However, the protected characteristic could still be arrived at by the ML model and therefore result in a discriminatory result. This phenomenon is defined by experts in the ML field as *redundant encoding* where information about a protected characteristic is encoded in other facially neutral data.[104] For instance, a ML model could use data such as schools attended by candidates in arriving at a candidate's gender despite not being privy to data on the same, if for example a candidate attended an all-girls or boys school. Additionally, a ML model could also arrive at a candidate's race based on prevailing data on their neighborhood.

The problem arising from these two examples is that the ML model could then use the protected characteristic it has arrived at in making hiring decisions.

e. Masking

This form of discrimination is more direct in that it occurs when employers intentionally incorporate their bias within an automated hiring practice by deliberately manipulating a ML model to arrive at a predetermined outcome.[105]

Thus far, the article has provided an insight into ML algorithms and the incorporation of the same during the recruitment phase of employment. In addition, it has showcased that the danger of algorithmic bias is imminent.

---

100 Kim P, 'Data-driven discrimination at work' 878.

101 Kim P, 'Data-driven discrimination at work' 878.

102 Kim P, 'Data-driven discrimination at work' 878.

103 Barocas S and Selbst A, 'Big data's disparate impact', 691.

104 Barocas S and Selbst A, 'Big data's disparate impact', 691.

105 Barocas S and Selbst A, 'Big data's disparate impact', 692.

III. The Challenges of Proving and Preventing Algorithmic Discrimination

To expound on the challenges present in proving and preventing algorithmic discrimination, this section will analyze the current anti-discrimination laws in Kenya. This analysis aims to bring light to the challenges of proving and preventing algorithmic discrimination and question whether Kenya's current laws are adequate in addressing this novel form of discrimination.

As mentioned in section II, most of the ways in which algorithmic discrimination occurs are obscured by big data or algorithms making unwarranted correlations thus using a person's protected characteristic as a basis for a hiring decision. Direct discrimination by ML models against protected classes is rare.[106] This is because ML models are programmed to eradicate existing bias by creating an ostensibly neutral approach to hiring decisions that is not centered on protected characteristics.[107] Therefore, ML models represent a facially neutral employment practice. Consequently, any discrimination resulting from the use of an ML algorithmic hiring tool would be considered indirect discrimination.

i. Kenya's anti-discrimination law

Kenya's anti-discrimination laws are well rooted in the Constitution, which expressly prohibits both direct and in-direct forms of discrimination on various grounds by both the state and individuals.[108] In doing so, the Constitution lists protected characteristics such as race, sex, pregnancy, marital status, health status, ethnic or social origin, colour, age, disability, religion, conscience, belief, culture, dress, language or birth.[109]

The Employment Act also speaks to discrimination. The Act prohibits both direct and in-direct forms of discrimination.[110] It states that no employer shall discriminate against either employees or prospective employees with respect to recruitment, training, promotion, terms and conditions of employment, termination of employment and any other employment-related matters.[111] Similar to the Constitution, the Act also provides for a list of protected characteristics and places the burden of proving non-discrimination on the employer.[112]

106 Kim P, 'Data-driven discrimination at work' 874.

107 Kim P, 'Data-driven discrimination at work' 874.

108 Article 27, *Constitution of Kenya* (2010).

109 Article 27, *Constitution of Kenya* (2010).

110 Section 5, *Employment Act* (Act No. 11 of 2007).

111 Section 5, *Employment Act* (Act No. 11 of 2007).

112 Section 5, *Employment Act* (Act No. 11 of 2007).

However, neither the Constitution nor the Employment Act defines what constitutes direct and in-direct discrimination, thus leaving it to the Courts to build jurisprudence around discrimination.

The High Court in *James Nyasora Nyarangi & 3 others v Attorney General* defined direct discrimination as treating someone less favorably because of their possession of a protected attribute compared to someone without that attribute in similar circumstances.[113] Drawing inspiration from the US case of *Griggs v Duke Power Company* the Court defined in-direct discrimination as 'setting a condition or requirement which a smaller proportion of those with the attribute are able to comply with without reasonable justification'.[114]

In the case of *Peter K Waweru v Republic*, the Court defined discrimination as treating people differently based on their protected characteristics. This occurs when individuals with one characteristic face disadvantages or restrictions that others do not, or when they receive privileges or benefits that are not available to those with different characteristics.[115]

The Employment and Labour Relations Court in *GMV v Bank of Africa Kenya Limited* provided the process for instituting a claim on grounds of discrimination.[116] The process outlined by the Court was that an aggrieved party must establish that they hold membership in a protected class, demonstrate qualification for the job, show adverse effects from an employment action, and provide a nexus between the adverse decision and their protected characteristic. Once a *prima facie* case is established, the burden shifts to the employer to justify their actions.[117]

The Supreme Court in *Samson Gwer & 5 others v Kenya Medical Research Institute & 3 others* further built on the jurisprudence around anti-discrimination law by providing for the level of proof required in matters pertaining to constitutional rights, such as discrimination. The Court determined that such matters require a higher standard of proof than the balance of probability test that is required in other civil claims.[118]

The case of *Methodist Church in Kenya v Mohamed Fugicha & 3 others* at the Court of Appeal additionally provided insight into how Kenyan courts handle

113 *James Nyasora Nyarangi & 3 others v Attorney General* (2008) eKLR.
114 *James Nyasora Nyarangi & 3 others v Attorney General* (2008) eKLR.
115 *Peter K Waweru v Republic* (2006) eKLR.
116 *GMV v Bank of Africa Kenya Limited* (2013) eKLR.
117 *GMV v Bank of Africa Kenya Limited* (2013) eKLR.
118 *Samson Gwer & 5 others v Kenya Medical Research Institute & 3 others* (2020) eKLR.

cases of indirect discrimination.[119] In this case, the Court held that indirect discrimination occurs when a person, policy, measure or criteria though neutral, places a person at a disadvantage compared to others because of their protected characteristic.[120] In arriving at decision the Court applied Justice Silber's test from *Sarika Angel Walkins Singh v the Governing Body of Aberdare Girls High School, & Another* which involves:[121] i) identifying the relevant provision, criterion or purpose (PCP) which is applicable, ii) determining the issue of disparate impact which entails identifying a pool for the purpose of making a comparison of the relevant disadvantage, iii) ascertaining if the PCP also disadvantages the claimant personally, and iv) assessing whether the policy is objectively justified by a legitimate aim and to consider whether this is proportionate.[122]

The laws and cases highlighted establish the four themes that make up the theory of indirect discrimination i.e, membership in a protected class, facially neutral PCP, the particular disadvantage and the justification.

ii. The challenges of proving and preventing algorithmic discrimination

a. Proving algorithmic discrimination

A claim of indirect discrimination begins with an individual realizing that they have faced a form of discrimination based on their membership in a protected class. Herein lies the biggest challenge of proving algorithmic discrimination. How does a rejected applicant realize that they have suffered discrimination?

A company could be using either off the shelf ML models or bespoke models to make their hiring decisions.[123] A prospective employee who has submitted their application or attended an interview would not be aware of the specific type of ML hiring algorithm an employer is using, or even if the employer is using an ML model at all. Furthermore, once a recruitment decision has been made, rejected applicants are rarely provided with reasons for their rejection per the generic rejection email (if any).[124] It would be exceedingly difficult for a prospective employee to recognize that they have been discriminated against, thereby halting the process of establishing a *prima facie* case.

119 *Methodist Church in Kenya v Mohamed Fugicha & 3 others* (2019) eKLR.

120 *Methodist Church in Kenya v Mohamed Fugicha & 3 others* (2019) eKLR.

121 *Sarika Angel Walkins Singh v the Governing Body of Aberdare Girls High School, & Another* (2008) EWHC 1865.

122 *Methodist Church in Kenya v Mohamed Fugicha & 3 others* (2019) eKLR and *Gichuru v Package Insurance Brokers Ltd* (Petition 36 of 2019) (2021) KESC 12 (KLR).

123 Ajunwa I, 'The paradox of automation as anti-bias intervention' 18.

124 O'Neil C, *Weapons of math destruction,* 101-104.

In the extremely rare case that a rejected applicant is able to prove a *prima facie* case of indirect discrimination, the employer would have numerous ways of justifying the use of a discriminatory algorithm.[125] An employer could argue that compared to human recruitment methods, the algorithmic hiring model significantly reduced previous bias by diversifying the hiring pool.[126] In addition to that, they would be backed by rafts of statistical evidence from experts involved in the programming of the ML model that would point to the fairness of its outcomes.[127] It is highly unlikely that an employer would fail to provide a justification for using an ML tool.[128]

At this point the burden would shift back to the rejected applicant to prove that a less discriminatory employment practice was available to the employee.[129] Most rejected applicants will neither have the resources nor the access to test an algorithm's discriminatory tendencies and a generalized argument that ML algorithms have been shown to exhibit bias is unlikely to succeed.[130]

It is evident that identifying algorithmic discrimination alone is an uphill task for prospective employees. Aislinn Kelly-Lyth avers that absent intervention from a coordinating body with significant technical expertise, an individual is unlikely to realize that they may have been disadvantaged by an employer's use of an algorithm and even if they do, they will struggle to find any evidence to prove it.[131]

b. Preventing algorithmic discrimination

Preventing algorithmic discrimination is challenging mainly because most people do not anticipate that algorithms will discriminate against protected classes. The opaque nature of ML algorithms creates predictability issues, making it difficult to foresee discriminatory outcomes from seemingly neutral algorithms once they are deployed. Since hiring algorithm are tailored to an employer's specific needs, their performance in a real-world environment may differ significantly from the training conditions.[132]

125 Khaitan T, *A Theory of Discrimination Law,* 75.

126 Kim P, 'Data-driven discrimination at work', 871.

127 Kim P, 'Data-driven discrimination at work' 869.

128 Barocas S and Selbst A, 'Big data's disparate impact', 708-710.

129 Barocas S and Selbst A, 'Big data's disparate impact', 708-710.

130 Kelly-Lyth A, 'Challenging biased hiring algorithms' 21.

131 Kelly-Lyth A, 'Challenging biased hiring algorithms' 23.

132 Kim P, 'Data-driven discrimination at work' 889.

Additionally, algorithms makes use of numerous data points, making it extremely difficult to deduce the features it has used in arriving at a particular decision.[133] The complex nature of ML makes it burdensome for programmers to identify instances of discrimination, as they may be unable to determine whether the model has used proxies to infer certain protected characteristics and whether the use of those characteristics constitutes discrimination.[134] The opacity of the models complicates the identification and correction of instances of discrimination especially after the models have been rolled out into the market.[135]

In addition to that, one of the reasons for using hiring algorithms is to reduce systemic bias that has historically been present in the recruitment process.[136] Therefore, by the time ML models are released into the market, the algorithms have undergone testing to ensure their outcomes comply with all necessary laws and regulations.[137] However, once a company begins implementing the hiring algorithm into their recruitment process they may not realize that the algorithm could be replicating existing bias or creating new data driven bias.[138]

Another significant issue is balancing the threat of discrimination against innovation within the ML space. Preventing algorithmic discrimination entails striking a balance between the likelihood of discrimination and the benefits of innovation. Legislators must determine whether the threat of instances of discrimination is enough to curtail the use of hiring algorithms by prohibiting the use of ML models that could result in discriminatory outcomes.

The primary challenge of preventing algorithmic discrimination then is the opacity of ML that makes identifying discrimination difficult before deployment into the real-world. Even post-deployment, it is difficult to identify instances of discrimination without intervention from a body with significant technical expertise in the area.

IV. The Way Forward in Combating Algorithmic Discrimination in Kenya

The current anti-discrimination laws in Kenya are broad enough to encompass the novel challenge of algorithmic discrimination. Indeed, as has been

133 Kim P, 'Data-driven discrimination at work' 881.

134 Kim P, 'Data-driven discrimination at work' 881.

135 O'Neil C, *Weapons of math destruction,* 6-15.

136 Ajunwa I, 'The paradox of automation as anti-bias intervention' 13 -15.

137 Ajunwa I, 'The paradox of automation as anti-bias intervention' 18.

138 Kim P, 'Data-driven discrimination at work' 881.

demonstrated in the previous section, most if not all discriminatory outcomes arising from the use of ML hiring algorithms will fall within the realms of indirect discrimination. Kenyan law expressly provides for indirect discrimination therefore placing the concept of algorithmic discrimination within the established laws will not be a challenge. What will be particularly difficult however would be the process of proving that discrimination has occurred, and how to prevent or reduce the likelihood of such discrimination. This then begs the question on how the current anti-discrimination laws in Kenya can evolve to better respond to this form of discrimination.

Genevieve writing on algorithmic fairness proposes that instead of trying to make a ML system completely fair (or 'de-biasing' it), the goal should be to detect and mitigate fairness-related harms as much as possible. She contends that fairness does not stop once the ML tool has been developed and that it is important to ensure users and stakeholders can observe, understand and appeal choices made by algorithms.[139]

In his work on algorithmic bias, Cecil Abungu asserts that most of the proposed solutions to algorithmic bias emanating from developed nations would be difficult to replicate in developing nations because such countries lack the necessary foundation.[140] He defines the necessary foundation as one with; a well rooted culture of transparency and statistical analysis of the disparities faced by protected groups, vigilant non-governmental actors alive to the risks of algorithmic discrimination and a reasonably robust and proactive executive branch or independent office policing of discrimination.[141]

Therefore, in order to achieve a sense of algorithmic fairness the necessary foundation must be created within the Kenyan legal system.

i. Recommendations

a. Recommendations regarding the law

The anti-discrimination framework in Kenya is robust enough to encompass algorithmic discrimination in recruitment. Where the law could develop, however, would be through regulations that would enable a jobseeker to easily identify instances of discrimination and subsequently prove the same to the courts.

139 Smith G, 'What does "fairness" mean for machine learning systems?' https://haas.berkeley.edu/wp-content/uploads/What-is-fairness_-EGAL2.pdf on 20 February 2022.

140 Abungu C, 'Algorithmic decision-making and discrimination in developing countries' 11.

141 Abungu C, 'Algorithmic decision-making and discrimination in developing countries' 8.

With this in mind, employers should have a legal duty to ensure that the algorithmic hiring systems that are in use are transparent, accountable and free from bias and discrimination. This should include legal provisions requiring employers to disclose the algorithms and data used in the hiring process, as well as to provide job seekers with information on why they were rejected for a particular job.

Employers should also be obligated to conduct regular evaluations of their hiring algorithms to ensure that they are free from bias. This could include provisions for independent audits as well as any other assessments that the government deems reasonable and necessary in order to ensure fair outcomes.

Another useful solution to prevent algorithmic discrimination would be auditing of ML tools ex ante through the use of a regulatory sandbox.[142] This regulatory sandbox would operate in a similar manner to the capital markets regulatory sandboxes currently in use in the financial sector.[143] Only after a ML algorithmic hiring tool has been tested in the sandbox and satisfied government requirements would it be released into the larger business market.

A reasonably robust and proactive executive branch interested and willing to implement the regulations discussed would be essential in ensuring that algorithmic discrimination is dealt with expeditiously.[144] This could be through establishment of a government agency with the technical expertise on matters pertaining to ML that could assist aggrieved job seekers in the process of proving algorithmic bias. This is particularly important because most prospective employees when faced with the issue of algorithmic bias neither have the expertise nor the finances to challenge the discriminatory practice.[145] The presence of a government agency with such a mandate would empower job seekers to seek recourse through the courts and would aid in building jurisprudence in relation to algorithmic discrimination. Ideally this body would operate like the Equality and Human Rights Commission in the UK.[146] The Commission has the power to investigate compliance with the Equality Act and take any necessary action against non-compliance.[147] The Commission also has

142 Think Tank, 'Artifical intelligence act and regulatory sandboxes', European Parliament, 17 June 2022 -<https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2022)733544> on 29 October 2022.

143 Capital Markets Authority, 'Regulatory Sandbox' -<https://sandbox.cma.or.ke/> on 29 October 2022.

144 Abungu C, 'Algorithmic decision-making and discrimination in developing countries' 58.

145 Kelly-Lyth A, 'Challenging biased hiring algorithms' 23.

146 Section 28 and 29 *Equality Act* 2006.

147 *Equality Act* 2006.

the power to provide financial support to individual claimants in bringing their cases before the courts.[148]

The final limb in combating this novel challenge is the inclusion of penalties for non-compliance with the above regulations and laws as a means of serving as a deterrent to non-compliance.

b. Non-legal recommendations

Indeed, a culture of transparency and statistical analysis of disparities faced by protected groups is necessary to aid in identifying instances of algorithmic discrimination. Building a culture of transparency must begin with providing education and training for employers, job seekers, and other stakeholders on the use of algorithmic hiring systems and the potential challenges and risks associated with them. This could include information on how to identify and avoid bias and discrimination in algorithmic hiring systems, as well as how to use such systems effectively and ethically. Disclosure of the use of algorithmic hiring tools by employers could also lend a hand in making the recruitment process more transparent to job-seekers and vigilant observers within the ML space. By doing this, prospective employees would be equipped with necessary information needed in proving a prima facie case of discrimination and would be incentivized to challenge instances in which they feel they have faced discrimination. In Kenya, although transparency is enshrined in the constitution as well as in various statutes such information is either non-existent or difficult to obtain.[149]

This culture of transparency should also be coupled with statistical analysis of disparities faced by protected groups. This then shifts the conversation on proving algorithmic discrimination away from the jobseeker towards an independent policing of discrimination. This is necessary because most rejected applicants will not have the capacity to collate sufficient evidence to adequately challenge a biased hiring algorithm beyond proving a prima facie case.[150] A wealth of statistical evidence of disparities collected by vigilant non-governmental actors alive to the risks of algorithmic bias would aid in building a case against biased ML algorithmic tools. The importance of an independent actor with the capacity to collect and analyze data in a bid to uncover algorithmic discrimination was seen when the COMPAS, 'Correctional Offender Management Profiling

---

148 Section 28 and 29 *Equality Act* 2006.

149 Article 10, *Constitution of Kenya* (2010).

150 O'Neil C, 'Personality tests are failing Americans at work' 18 January 2021 - <https://www.bloomberg.com/opinion/articles/2018-01-18/personality-tests-are-failing-american-workers> on 22 December 2022.

for Alternative Sanctions', algorithm on recidivism was discovered to be biased against African Americans.[151]

COMPAS was a software developed by Northpointe (now Equivant) that was used by US courts and parole boards to predict recidivism rates among defendants.[152] The investigation into the potential bias engrained in the algorithm was conducted by ProPublica, an independent non-profit news organization within the US.[153] Their investigation revealed that the COMPAS algorithm was more likely to incorrectly label black defendants as having a higher recidivism rate as compared to their white counterparts therefore resulting in a higher conviction rate among blacks.[154] In arriving at this conclusion ProPublica evaluated the algorithms output in a pool of over ten thousand criminal defendants over a two year period in order to provide statistically sound evidence of algorithmic bias.[155] This is a feat that would be nearly impossible for an independent defendant aggrieved by the use of the COMPAS algorithm. The availability of the data in an open access manner also played a large part in identifying the algorithmic bias.[156] This goes to show the importance of independent bodies within the ML space with the capacity to collect and interpret data obtained from the use of ML models. It also illustrates the importance of open access data as a tool for aiding in proving and preventing discrimination.[157] In that respect it would be important for employers to be willing to provide access to data pertaining to their hiring decisions.

151 Larson *et al.*, 'How We Analyzed the COMPAS Recidivism Algorithm', Pro Publica, 23 May 2016 - <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm> Kleinberg *et al.*, 'Inherent trade-offs in the fair determination of risk scores', Working Paper (2016), 5-6 - <https://arxiv.org/abs/1609.05807>.

152 Larson et al., 'How We Analyzed the COMPAS Recidivism Algorithm', Pro Publica, 23 May 2016 - <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Kleinberg et al., 'Inherent trade-offs in the fair determination of risk scores', Working Paper (2016), 5-6 - <https://arxiv.org/abs/1609.05807>.

153 Borgesius F, 'Strengthening legal protection against discrimination by algorithms and artificial intelligence', 24(10) *The International Journal of Human Rights*, 2020, 1572-1593.

154 Angwin J, 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks', ProPublica, 23 May 2016, <https://www.ProPublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> on 13 February 2023.

155 Borgesius F, 'Strengthening legal protection against discrimination by algorithms and artificial intelligence', 1572-1593.

156 Larson et al., 'How We Analyzed the COMPAS Recidivism Algorithm', Pro Publica, 23 May 2016 - <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Kleinberg et al., 'Inherent trade-offs in the fair determination of risk scores', Working Paper (2016), 5-6 <https://arxiv.org/abs/1609.05807>.

157 Angwin J, 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks', ProPublica, 23 May 2016, - <https://www.ProPublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> on 13 February 2023.

The final consideration that would aid in proving and preventing algorithmic discrimination would be investing in research and development to improve the accuracy, fairness and transparency of algorithmic hiring tools. This research and development could prove pivotal in influencing how regulations are crafted regarding the use of hiring algorithms.

V. Conclusion

This article has demonstrated that algorithmic discrimination during hiring is indeed a real threat to the Kenyan jobseeker and to that effect solutions have been proposed to aid in mitigating algorithmic discrimination.

The author is cognizant that proving and preventing algorithmic discrimination is indeed an uphill task for an individual, however, moving from an individual centered approach to a more institutional approach where individuals and non-governmental organizations as well as government agencies work together could prove beneficial. An aggrieved jobseeker will find it difficult to prove algorithmic discrimination working independently, however when paired with organizations with the necessary expertise and funding the process would be manageable. Preventing instances of algorithmic bias from occurring is an equally daunting task that would require frequent auditing of algorithms to flag any discriminatory outcomes before they are replicated into the employment space. Indeed, this could prove an expensive venture but the same would be a necessary concession as the government seeks to balance the interests of employers against the rights of prospective employees.

The recommendations provided seek to strike a balance between protecting the rights of the citizenry and encouraging innovation.