A Novel of Detecting and Addressing Bias in Artificial Intelligence and Machine Learning: A Multi-Industry Perspective

Sukanya Konatam[1], Venkata Naga Murali Konatam[2], Shravya Konatam[3]

[1]*Senior Manager of Enterprise Data Governance and Data Science, IT, Vialto Partners, Texas, USA*
[2]*Data Architect, IT, Capital One, Texas, USA*
[3]*Student, Pre-Med, Nova SouthEastern University, Florida, USA*

---***---

Abstract - *Artificial intelligence (AI) and machine learning (ML) have gained widespread popularity and transformative potential in various critical sectors, including healthcare, hiring, self-driving cars, and financial systems. Despite their benefits, the adoption of AI and ML systems introduces biases that can profoundly impact decision-making processes and outcomes. These biases stem from various sources, such as biased training data, algorithmic design, and implementation practices, leading to ethical concerns and consequences. This paper provides a comprehensive exploration of the biases inherent in AI and ML systems and discusses their ethical implications across the aforementioned sectors. Through a systematic review of current literature, we identify the primary sources and types of biases, such as racial, gender, and socioeconomic biases, and analyze their effects on decision-making. Additionally, we examine case studies that illustrate the real-world impact of biased AI systems, shedding light on the critical need for fair AI development. To address these challenges, we propose a set of mitigation strategies aimed at reducing bias and enhancing the fairness and accountability of AI systems. These strategies include improving the diversity and representativeness of training data, implementing algorithmic fairness techniques, conducting regular audits and impact assessments, and fostering interdisciplinary collaboration among AI developers, ethicists, and policymakers. Our findings highlight the urgent need for a holistic approach to AI development that prioritizes ethical considerations and social responsibility. By adopting these mitigation strategies, we can work towards creating AI systems that not only deliver technological advancements but also promote equity, transparency, and trust in critical sectors.*

***Key Words*: Autonomous Vehicles, Artificial Intelligence in HealthCare, Bias in Machine Learning, AI accountability, AI in job screening, ML in Finance, Bias in Artificial Intelligence**

1. Introduction to Artificial Intelligence and Machine Learning in Different Sectors

Artificial Intelligence (AI) and Machine Learning (ML) are transformative technologies that have revolutionized various sectors [1], [2], [3]. AI refers to the simulation of human intelligence processes by machines, especially computer systems [1], [3]. These processes include learning, reasoning, problem-solving, perception, and language understanding [1], [3]. On the other hand, ML is a subset of AI that provides systems the ability to learn and improve from experience without being explicitly programmed [1], [2], [3]. AI and ML have permeated various sectors, revolutionizing processes and systems [1], [2], [3].This research paper will focus on four key sectors where AI/ML has a significant impact: the Hiring Process, Healthcare, Financial Systems, and Self-Driving Cars [3][6]. Surprisingly, these sectors are not immune to biases in AI/ML models, which can have profound implications [4], [5]. The exploration of these biases forms the crux of this research. These technologies have the potential to bring about significant benefits, such as increased efficiency, improved decision-making, and enhanced user experiences [1], [2], [3]. However, they also pose new challenges, such as the risk of biases in AI/ML models [4], [5]. This paper will explore these issues in detail, with a focus on understanding the sources of these biases and how they can be mitigated [4], [5].

$$J = \sum_{i=1}^{n} \frac{(mX_i + c - Y_i)^2}{n}$$

Fig -1: The mathematical representation of the Mean Squared Error (MSE) loss function used in machine learning models. This function measures the average squared difference between the predicted values (mX_i + c) and the actual values (Y_i), highlighting the importance of minimizing prediction errors to improve model accuracy and reduce biases.

2. Artificial Intelligence and Machine Learning in Healthcare

Artificial Intelligence (AI) and Machine Learning (ML) are transforming the healthcare industry by offering innovative solutions for patient care, diagnosis, and treatment [7]. These technologies analyze complex medical data, predict patient outcomes, and even assist in surgeries [7]. They also play a crucial role in streamlining administrative tasks, improving patient engagement, and personalizing healthcare services [7].

2.1 History of Artificial Intelligence Advancements in Health-Care

The journey of AI in healthcare began in the 1970s with rule-based expert systems like MYCIN and INTERNIST-I, designed to assist in medical diagnosis [8]. Soon after the 1990s saw the advent of Machine Learning (ML), with algorithms being trained to predict patient outcomes [8]. The 21st century ushered in a new era with the introduction of Deep Learning, enabling more complex analyses of medical data [8]. Today, AI and ML are integral to many aspects of healthcare [9]. These advancements have revolutionized patient care, disease prediction, and treatment planning, marking a significant milestone in the history of healthcare [8].

2.2 Notable Artificial Intelligence and Machine Learning Technologies in Health-Care

AI and machine learning technologies have significantly impacted the healthcare sector [14]. IBM Watson Health uses AI to accelerate drug discovery and match patients with clinical trials. [10] Google's DeepMind has made strides in AI-assisted diagnosis, notably in detecting eye diseases from scans [11], [13]. Zebra Medical Vision uses AI for radiology imaging, aiding in early detection of cancers [13]. PathAI focuses on improving pathology accuracy with machine learning [12], [15], [16]. Tempus uses AI for precision medicine, particularly in cancer treatment. [14] In the realm of EHR's (Electronic health records), Google Cloud Healthcare API and Microsoft Azure's FHIR APIs facilitate the exchange of healthcare data, while Epic Systems and Cerner use AI to analyze EHR data for better patient outcomes [12], [14]. While these technologies have been transformative, they also necessitate careful consideration of potential biases in their underlying algorithms [17].

2.3 Artificial Intelligence Biases in The Healthcare Industry

AI/ML biases in healthcare refer to the unintentional integration of human- and data-driven biases into AI technologies and machine learning models used in healthcare [18]. These biases can result from several factors, including human biases built into AI design, data generalizability issues, and biased training data [18]. If the training data used to develop healthcare AI/ML models is not representative of the diverse patient population, the models may perform poorly for underrepresented groups. [18] [19] Additionally, if the data contains historical biases in healthcare delivery, the models may perpetuate these biases [19], [20], [21]. For instance, if certain groups have historically received less aggressive treatment for a particular condition, an AI/ML model trained on this data may recommend less aggressive treatment for patients from these groups [21].

2.4 Popular Types of Artificial Intelligence Biases found in HealthCare industry

- **Measurement Bias-** Measurement bias in AI/ML refers to the systematic error that arises when the data used to train a model does not accurately reflect the reality it's intended to predict [22], [24], [25]. This bias can significantly skew the model's predictions and decisions [22], [24], [25]. In healthcare, AI/ML models are often employed to predict patient outcomes, recommend treatments, and identify disease patterns [23], [25]. However, if the training data is biased towards a specific demographic, the model may fail to accurately predict outcomes for underrepresented groups [23], [25]. This could lead to misdiagnoses or inappropriate treatment recommendations [23], [25].One example of this measurement bias is the case of an algorithm developed to predict the risk of cardiovascular disease. Researchers trained the algorithm using electronic health record (EHR) data from a large healthcare system [26]. The training data was predominantly from younger, healthier patients, with limited representation of older adults and those with existing medical conditions [26]. As a result, the algorithm tended to underestimate the cardiovascular disease risk for these underrepresented groups [26]. The use of AI/ML in healthcare is driven by the potential to improve patient care and outcomes, but it's crucial to ensure diverse and representative data in the training phase to mitigate measurement bias and improve the fairness and accuracy of these models [23], [25].

- **Confirmation Bias-** Confirmation bias in AI/ML refers to the tendency of models to favor information that confirms their existing beliefs or hypotheses [27]. This bias often originates from the data used to train these models, reflecting the inherent biases of those who collect and label the data [27]. In healthcare, AI/ML models are used to predict patient outcomes, recommend treatments, and identify diseases [27], [28]. However, if these models are trained on biased data, such as data primarily from a specific demographic, they may disproportionately favor that demographic in their predictions, thereby exhibiting confirmation bias [27]. A study examining the use of AI for diagnosing heart disease provides a clear example of confirmation bias in healthcare AI applications [29]. Researchers developed an AI model to predict the likelihood of a patient having heart disease based on factors such as age, sex, and medical history. The training data for the model was primarily collected from a single hospital, which served a predominantly white, middle-aged patient population [29]. When the model was tested on data from a more diverse set of patients, it exhibited significantly lower accuracy in predicting heart disease risk for racial minorities and younger/older patients [29]. This could lead to an unexpected misdiagnosis or inappropriate treatment recommendations, underscoring the importance of addressing confirmation bias in healthcare AI/ML applications [27)], [28].

2.5 Ethical implications of Bias in Healthcare Artificial Intelligence and Machine Learning

Bias in healthcare AI/ML carries significant ethical implications. If a model is trained on unrepresentative data, it may produce skewed results, leading to unequal and potentially harmful treatment outcomes [30]. For instance, a model trained predominantly on data from one demographic may not perform accurately for other demographics, leading to unexpected misdiagnoses or inappropriate treatment recommendations [31]. This raises ethical concerns about fairness and equality in healthcare delivery [30]. Furthermore, biased AI/ML models can reinforce existing health disparities, particularly for marginalized or underrepresented groups [32]. For example, a study found that an algorithm used to guide healthcare decisions was less likely to refer Black patients than White patients for programs that aim to improve care for patients with complex medical needs, due to biases in the training data [32]. This underscores the ethical imperative for careful and representative data collection, transparency in model development, and ongoing monitoring for bias in healthcare AI/ML applications [30], [32]. Therefore, addressing AI biases in healthcare is crucial for ensuring equitable and efficient healthcare delivery [30], [31], [32].

Fig -2: Surgeons utilizing AI-assisted technology during a complex procedure. AI systems enhance precision and efficiency in healthcare, contributing to improved surgical outcomes and patient care.

2.6 Strategies for Detecting Biases in Healthcare Artificial Intelligence and Machine Learning

Detecting biases in healthcare AI/ML involves a multi-faceted approach. Firstly, it's crucial to scrutinize the data used to train the models. This includes assessing the representativeness of the intricate data, checking for overrepresentation or underrepresentation of certain groups, and identifying potential sources of bias in data collection [33]. Secondly, evaluating the model's performance across different demographic groups can help detect biases [35]. Disparities in performance metrics such as accuracy, precision, or recall may indicate the presence of bias. Lastly, interpretability techniques can be used to understand the decision-making process of the model, which can reveal if certain features are being given undue importance, leading to biased outcomes [34].

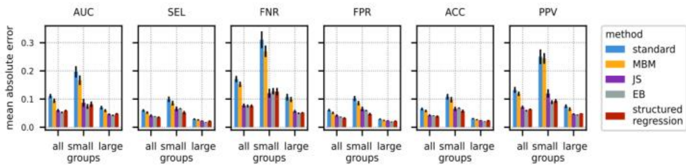

Chart -1: In this chart, the mean absolute error (MAE) estimates for six metrics using five different methods on diabetes data. The evaluation is based on 20 draws of the dataset, considering three distinct groups: all patients, small groups (size at most 25), and large groups (size above 25). These results shed light on potential biases in healthcare AI models, emphasizing the importance of addressing disparities across patient subgroups.

2.7 Detection of Famous Forms of Biases found in Health-Care Artificial Intelligence

● **Finding Measurement Bias-** To detect Measurement bias in healthcare AI, one can examine the data collection instruments and processes for potential sources of error or bias [36], [37], [38]. For instance, if a health survey is used to collect data, are the questions clear, unbiased, and culturally sensitive? Additionally, statistical analyses can be conducted to check if the measurements are consistent and reliable across different groups [37], [39]. One example of detecting measurement bias in healthcare AI/ML systems is the case of an algorithm developed to detect early-stage Alzheimer's disease using speech data [40]. The developers of this algorithm used a training dataset that only contained speech samples from native English speakers. As a result, the algorithm had difficulty accurately analyzing speech patterns of non-native English speakers, as it interpreted their pauses and pronunciation differences as potential markers of Alzheimer's disease [40].By examining the data collection process and the representation of different patient groups in the training data, the measurement bias in this healthcare AI system could be identified and addressed. Conducting statistical analyses to assess the consistency and reliability of the algorithm's performance across diverse patient populations is also crucial for detecting such biases [40].

- **Identifying Confirmation Bias-** Confirmation bias in healthcare AI/ML can be detected by examining the model's predictions in light of contradicting evidence. If the model consistently ignores or downplays such evidence, it may be exhibiting confirmation bias [41]. Another method is to use adversarial testing, where the model is presented with cases specifically designed to challenge its assumptions. If the model's performance significantly deteriorates in these cases, it could be a sign of confirmation bias [41]. One example of detecting confirmation bias in a healthcare AI/ML model can be found in a study exploring the use of AI-powered intrusion detection systems in Internet of Medical Things (IoMT) environments [42]. The researchers architected an integrated anomaly detector that incorporated modules for model interpretability, bias quantification, and advanced malicious input recognition. Through their experimental evaluation, the researchers were able to identify instances where the baseline intrusion detection models exhibited confirmation bias [42]. Specifically, they found that these models would consistently ignore or downplay certain types of data poisoning threats that contradicted the models' existing beliefs about the characteristics of malicious inputs. By employing adversarial testing techniques, the researchers were able to generate carefully crafted examples that challenged the models' underlying assumptions [42]. When presented with these adversarial samples, the baseline models showed a significant deterioration in performance, revealing their susceptibility to confirmation bias.

3. Artificial Intelligence and Machine Learning in Autonomous Vehicles

Artificial intelligence (AI) and machine learning play a crucial role in the development of Autonomous Vehicles. These technologies enable autonomous vehicles to perceive their surroundings, make real-time decisions, and navigate safely [43], [44], [47]. The integration of AI and machine learning in self-driving cars has the potential to improve road safety, reduce traffic congestion, and enhance overall transportation efficiency [45], [46], [48], [49].

3.1 History of Artificial Intelligence Advancements in Self Driving Cars

The adoption and development of AI and machine learning in autonomous vehicles has evolved dramatically over the past few decades. In the 1980s, early research into vision-based systems and neural networks laid the groundwork for self-driving car technology [50]. This was followed by significant breakthroughs in the 1990s, such as Carnegie Mellon University's Navlab project which demonstrated autonomous driving on public roads [50]. The 2000s saw widespread interest and investment in self-driving cars, with companies like Google, Tesla, and traditional automakers accelerating the development of key capabilities like perception, planning, and control using advanced AI techniques [51]. Landmark events like the DARPA Grand Challenges further pushed the boundaries of autonomous driving, leading to autonomous vehicles navigating complex urban environments [52]. Today, self-driving car technology has reached a level of maturity, with fully autonomous vehicles undergoing extensive testing and pilot deployments in many parts of the world [52]. The future promises even greater advancements as AI algorithms become more sophisticated and capable of handling the full complexity of real-world driving scenarios.

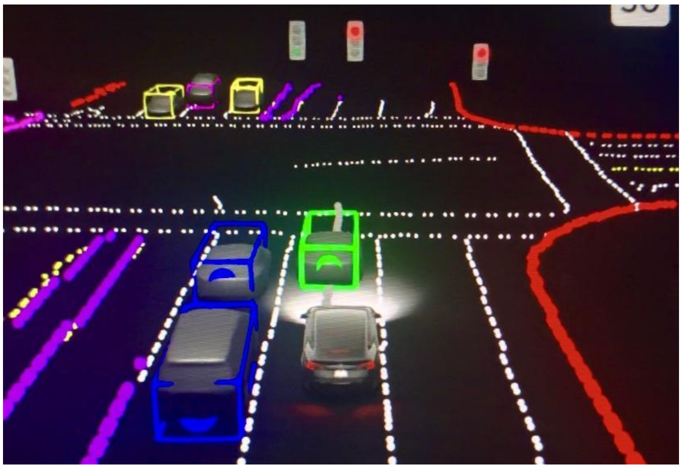

Fig -3: Visualization of a Tesla's AI perception system in self-driving mode, showing how the AI identifies and categorizes vehicles and road elements.

3.2 Notable Artificial Intelligence and Machine Learning Technologies in Self Driving Vehicles

AI and machine learning technologies have significantly impacted the Autonomous Vehicles sector [53]. Visual place recognition (VPR) is a crucial AI technique employed in self-driving cars, enabling them to recognize landmarks, traffic signs, and other visual cues from the environment [53]. This sophisticated computer vision approach, which involves the identification and understanding of these visual elements, plays an essential role in enhancing the safety, efficiency, and overall intelligence of intelligent transportation systems [53]. Additionally, deep learning models such as EfficientNetB3 are leveraged for robust feature extraction from preprocessed images, with their hyperparameters tuned using optimization algorithms like the Remora Optimization Algorithm to optimize the model's performance [53]. Furthermore, techniques like transfer learning are integrated to leverage the knowledge gained from pre-trained models, improving the overall accuracy and efficiency of visual place recognition [53]. Adversarial machine learning is another significant AI technology used in autonomous vehicles, where researchers explore methods to deceive the machine learning algorithms by providing malicious input data [54]. This field of study is crucial as it helps identify vulnerabilities in the AI systems used in self-driving cars, ensuring they can withstand attempts to manipulate their decision-making processes through adversarial attacks [54]. Experiments have been conducted using pre-trained models like MobileNet and Convolutional Neural Networks (CNNs) to demonstrate how one-pixel, multi-pixel, and all-pixel attacks can be designed to mislead the machine learning algorithms [54], [55].

3.3 Artificial Intelligence Biases in The Self Driving Cars Industry

The autonomous vehicle sector has seen the emergence of significant biases within the AI and machine learning (ML) systems powering these technologies [56]. These biases can manifest in various ways, such as faulty object detection, inaccurate prediction of pedestrian behavior, and suboptimal decision-making by the autonomous system [56], [57]. The presence of such biases can have serious consequences, leading to unsafe outcomes, unequal treatment of different individuals or groups, and broader societal implications in terms of accessibility, equity, and public trust in self-driving car technology [56], [57].

3.4 Popular Types of Artificial Intelligence Biases found in Self Driving Cars

- **Availability Bias-** Availability bias is a cognitive bias where individuals tend to overestimate the likelihood of events that are more readily available in their memory or imagination [58], [59], [61]. In the context of self-driving cars, this bias can lead developers to prioritize and focus on edge cases or scenarios that have received significant media attention or public discourse, while potentially overlooking more common but less salient situations [58], [59]. This can result in the self-driving car algorithms being optimized for rare, dramatic events, rather than the more mundane, everyday driving situations [58], [59], [61]. For example a study on self-driving car algorithms revealed an availability bias in the training data used to develop the systems. The researchers found that the training data heavily focused on high-profile incidents, such as fatal collisions involving pedestrians, which had received significant media attention [62]. This resulted in the self-driving car algorithms being optimized to handle these dramatic scenarios, while potentially overlooking more common, everyday driving situations [63]. Consequently, the self-driving cars performed well in highly publicized test cases but struggled to navigate the nuances of real-world driving, highlighting the need for a more balanced and representative training dataset to mitigate the availability bias [64], [63].

- **Affective Bias-** Affective bias refers to the tendency of individuals to make judgments or decisions based on their emotional responses and feelings, rather than purely objective data or information [60], [61]. In the context of self-driving cars, this bias can manifest in the way the algorithms are trained and designed, as the developers may inadvertently incorporate their own emotional associations or biases into the decision-making process [60], [61]. For instance, a study conducted by researchers at the University of Michigan found that self-driving car algorithms can exhibit gender bias when it comes to pedestrian detection and response [65]. The study showed that the algorithms were less likely to identify female pedestrians, potentially leading to increased risk for women in real-world scenarios [65]. This affective bias, rooted in societal stereotypes and norms, highlights the importance of thorough testing and evaluation of self-driving car systems to ensure fair and unbiased decision-making [65]. This bias can lead to suboptimal decision-making and potentially endanger the safety of other road users, undermining the core purpose of self-driving technology [60], [61].

3.5 Ethical Implications of Bias in Autonomous Vehicles Artificial Intelligence and Machine Learning

The proliferation of autonomous vehicles (AVs) powered by artificial intelligence (AI) and machine learning (ML) has raised profound ethical concerns regarding the potential biases inherent in these systems [66]. The ethical considerations surrounding biases in AI/ML for AVs are multifaceted, encompassing issues of safety, transparency, accountability, and the promotion of fairness [66]. The deployment of AI-driven AVs poses significant risks if the underlying algorithms exhibit biases, which can lead to discriminatory treatment and unfair outcomes [67]. These biases may manifest in the decision-making processes of AVs, potentially resulting in the differential treatment of individuals based on factors such as race, gender, or age [68], [69]. Such biases could have severe consequences, potentially compromising the safety of certain groups and undermining the core principles of equitable transportation [67]. Furthermore, the complexity and opacity of the AI/ML models used in AVs can make it challenging to identify and mitigate these biases [70]. The inherent difficulty in ensuring algorithmic transparency and explicability heightens the ethical dilemmas surrounding the deployment of these technologies [66].

3.6 Strategies for Detecting Biases in Self Driving Cars Artificial Intelligence and Machine Learning

Identifying and addressing biases in the AI/ML systems powering self-driving cars is a critical challenge. Approaches such as algorithmic auditing, testing on diverse datasets, and building in explainability and interpretability can help detect biases in areas like object detection, pedestrian recognition, and traffic scenario handling [71]. However, the complexity of these systems, the difficulty in obtaining comprehensive training data, and the evolving nature of on-road scenarios pose significant hurdles in fully eliminating biases. Continued research, standardized testing frameworks, and close collaboration between developers, domain experts, and impacted communities are necessary to make meaningful progress in this space [71].

3.7 Detection of Famous Forms of Biases found in Self Driving Vehicles Artificial Intelligence

- **Discovering Availability Bias-** The availability bias is a cognitive shortcut that leads individuals to rely on information that is readily available when making decisions [72]. To detect availability bias in autonomous vehicles, researchers have analyzed class imbalances in the datasets used to train pedestrian detection algorithms [74]. They have found that these algorithms may be less accurate in recognizing individuals with disabilities or from minority groups, as the training data often lacks sufficient representation of these populations [73]. One study by the National Institute for Transportation and Communities found that "black pedestrians were passed by twice as many cars and experienced wait times that were 32% longer than white pedestrians" when interacting with autonomous vehicles [78]. The researchers call for automotive manufacturers and the government to collaborate in building regulations that objectively measure the fairness and safety of these autonomous driving systems, as the current provisions for fairness are limited and can have a major impact on pedestrian safety [78]. To address this availability bias in self-driving cars, machine learning algorithms can be applied for object recognition and behavior monitoring in advanced driver-assistance systems (ADAS) [79]. These ML techniques, such as Faster R-CNN and YOLO, enable the accurate real-time detection and classification of objects, including pedestrians with darker skin tones and children, to improve the safety and fairness of autonomous vehicle navigation [80]. Additionally, the lack of transparency in many pedestrian detection algorithms, particularly those based on deep learning techniques, makes it difficult to understand how these systems arrive at their classifications [73]. Measures to improve dataset diversity and establish external oversight and regulation of these algorithms can help identify and mitigate availability bias [73].

- **Identifying Affective Bias-** Affective bias refers to the influence of emotional responses and cognitive biases on decision-making [76]. Researchers have used various methods to study the impact of affective bias on autonomous vehicle decision-making, such as analyzing users' physiological responses, including heart rate, muscle activity, eye movements, and brain waves, during real-world or simulated driving scenarios [75]. For example, researchers at Delft University of Technology have proposed an approach involving analyzing the sensitivity and selectivity of neurons in convolutional neural network (CNN) models, which can reveal affective biases towards certain features or classes [77]. Additionally, examining the attention maps of vision transformer (ViT) models can help identify whether the models are disproportionately focusing on certain image regions, potentially indicating biases [77]. By applying these techniques, researchers can uncover the specific biases inherent in different model architectures and guide the development of more balanced and equitable AI systems for self-driving cars [77]. These studies have found that people may feel safer and prefer to maintain control of the vehicle, even if the autonomous system is performing equally well [75]. To address affective bias, developers of autonomous vehicles can focus on improving dataset diversity, preventing and remediating algorithmic bias, and establishing standards for external oversight and regulation [73]. Additionally, fostering a better understanding of the underlying technology and its capabilities through public education and engagement can help build trust and acceptance [75].

4. Artificial Intelligence and Machine Learning in Financial Systems

Artificial Intelligence (AI) and Machine Learning (ML) have emerged as transformative technologies in the financial sector, revolutionizing various aspects of the industry [81], [82], [83]. These advanced techniques have enabled financial institutions to enhance customer experiences, democratize financial services, ensure consumer protection, and significantly improve risk management [83].The integration of AI and ML has empowered financial organizations to gain deeper insights into market dynamics, predict consumer behavior more accurately, optimize production schedules, and streamline distribution processes [82], [84]. By leveraging AI algorithms and ML models, the finance sector can make data-driven decisions with greater precision and agility, driving innovation, growth, and competitiveness [84].

4.1 History of Artificial Intelligence Advancements in The Finance Sector

The adoption and development of Artificial Intelligence (AI) and Machine Learning (ML) within the finance sector has undergone a remarkable evolution over the past few decades [85]. In the 1970s and 1980s, pioneering efforts focused on utilizing expert systems and rule-based algorithms for tasks such as credit scoring and fraud detection [86]. The 1990s and 2000s saw a surge in the adoption of more advanced statistical and neural network models, enabling financial institutions to tackle increasingly complex problems like portfolio optimization and algorithmic trading [86], [87]. A significant milestone was the development of Deep Learning techniques in the late 2000s, which revolutionized the field of computer vision and natural language processing, leading to breakthroughs in areas like anomaly detection, sentiment analysis, and predictive modeling within the finance industry [87], [88]. The proliferation of big data and the increased computational power of modern hardware have further accelerated the deployment of AI-driven solutions across banking, investment management, and insurance sectors [89]. Today, AI and ML have become indispensable tools in the finance industry, automating repetitive tasks, enhancing decision-making, and providing personalized services to customers, transforming the landscape of the financial ecosystem and promising continued innovation and competitive advantages for organizations that embrace this digital transformation.

4.2 Notable Artificial Intelligence and Machine Learning Technologies in the Sector of Finance

Artificial intelligence (AI) and machine learning (ML) have become integral to the financial sector, powering a wide range of applications and driving innovation. One of the key AI/ML technologies used in finance is natural language processing (NLP), which enables the extraction of insights from unstructured data such as financial reports, news articles, and social media [90]. NLP algorithms are employed to analyze sentiment, detect anomalies, and identify potential risks or opportunities [91]. Another prominent AI/ML technique used in finance is predictive analytics, which leverages historical data and complex models to forecast market trends, stock prices, and customer behavior [92], [93]. Financial institutions also utilize machine learning models, such as neural networks and decision trees, to automate and optimize investment strategies, credit risk assessment, and fraud detection [93], [94]. Additionally, reinforcement learning algorithms are used in areas like algorithmic trading, where they can learn and adapt to market conditions to make real-time investment decisions [94]. The integration of these AI/ML technologies has enabled financial institutions to enhance decision-making, improve operational efficiency, and provide personalized services to clients, ultimately driving competitive advantage in the industry.

4.3 Artificial Intelligence Biases in The Finance Industry

AI and ML systems used in the financial industry are increasingly susceptible to various forms of bias [95]. These biases can emerge from the data used to train the models, the algorithms employed, or the assumptions and design choices made by the developers [95]. Such biases can lead to unfair and discriminatory decision-making, negatively impacting access to financial services, loan approvals, and other critical outcomes, with significant societal implications. Addressing these biases in AI/ML systems within the finance sector is crucial to ensure fairness, transparency, and ethical practices in this domain.

4.4 Popular Types of Artificial Intelligence Biases found in The Finance Industry

- **Selection Bias-** Selection bias can occur in financial systems when the data used to train machine learning models is not representative of the entire population. This can happen when the data only includes information from certain demographics or geographic regions, or when certain groups are underrepresented in the data [96], [97]. This can lead to models that perform poorly or make biased decisions for those underrepresented groups. A study examining the use of AI-powered credit scoring models found that the training data often lacked representation from certain demographic groups, such as racial minorities and low-income individuals [98]. This led to the models exhibiting biases against these underrepresented groups, resulting in them being unfairly denied access to credit or offered less favorable lending terms [98]. The lack of diverse and representative data used to train these AI models was a key factor contributing to the perpetuation of existing socioeconomic inequalities through the credit decision-making process.

- **Systematic Bias-** Systematic bias in financial systems can arise from the underlying assumptions or design choices made in the development of machine learning models. This type of bias can be particularly problematic because it is often more difficult to detect and can be perpetuated through the repeated use of the biased model [96], [97]. For example, a study found that mortgage algorithms systematically charged higher interest rates to Black and Latino borrowers compared to white borrowers, perpetuating existing socioeconomic disparities [99]. This type of biased decision-making can have a lasting impact on the financial well-being of affected individuals and communities. This type of bias can lead to unfair and discriminatory outcomes, undermining the fairness and integrity of the financial system.

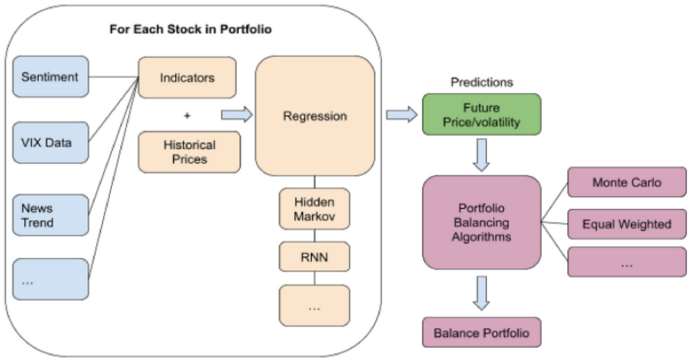

Chart -2: Flowchart illustrating the use of AI and ML in financial systems for stock portfolio management. The image highlights potential biases in data sources and predictive models, such as sentiment analysis, VIX data, and news trends, which can affect investment decisions and portfolio balancing algorithms.

4.5 Ethical implications of Bias in Financial Systems Artificial Intelligence and Machine Learning

The widespread adoption of artificial intelligence (AI) and machine learning (ML) in the finance industry has raised significant ethical concerns [100], [101], [103]. These technologies have the potential to embed and amplify biases, leading to unfair and discriminatory outcomes that can severely impact individuals' access to critical financial services [101], [102]. Biases can manifest in various forms, such as racial, gender, or socioeconomic discrimination, perpetuating existing inequalities and undermining the principles of fairness and equal opportunity [101], [102]. Furthermore, the inherent opacity of AI/ML models, often referred to as the "black box" problem, makes it challenging to identify the root causes of these biases and hold the responsible parties accountable [100], [101]. Consequently, the ethical implications of AI/ML in finance extend beyond just individual harm, as they can also undermine public trust in the financial system and jeopardize its overall stability [102]. Addressing these ethical concerns is crucial to ensure the responsible and equitable deployment of AI/ML technologies in the finance sector.

4.6 Strategies for Detecting Biases in Financial Artificial Intelligence and Machine Learning

Financial institutions must proactively address the issue of bias in their AI/ML systems used for decision-making. This involves implementing comprehensive strategies to identify and detect such biases. Some key approaches include conducting regular audits of the algorithms, monitoring for data and concept drift over time, and leveraging techniques like reverse-engineering to gain visibility into opaque "black box" models [104]. However, detecting AI/ML biases can be inherently challenging, as the biases may not be immediately apparent and may only become visible after the fact [104]. Continuously assessing the algorithms and data sources for potential sources of bias, while maintaining strong governance and transparency, is crucial for financial firms to mitigate the risks of unfair and discriminatory outcomes.

4.7 Detection of Famous Forms of Biases found in Financial Artificial Intelligence

- **Uncovering Selection Bias-** Selection bias in the finance industry refers to the situation where the training data used for AI and machine learning systems is not large or representative enough, leading to skewed results [105]. To detect selection bias, organizations should look at reducing the possibility of biased data sets in all phases of their data pipeline. This includes examining the data collection process for incomplete or unrepresentative data, and ensuring data preparation and labeling do not introduce further bias [105]. Tools like Fairlearn can be used to measure disparities in selection rates between different population subgroups and implement constraints to equalize the odds [106]. Additionally, techniques like adversarial debiasing and dynamic upsampling of training data can help mitigate selection bias [105]. A study on credit risk prediction models revealed that the training data used to develop these models was skewed towards individuals from higher socioeconomic backgrounds, leading to biased model predictions that disadvantaged lower-income applicants [110], [111]. The researchers employed techniques such as causal graph fuzzing and statistical fairness assessments to identify the presence of this selection bias, quantifying the degree to which the models made decisions that systematically discriminated against certain demographic groups [111]. The findings highlight the importance of carefully evaluating the representativeness of training data and the potential for AI/ML models to perpetuate societal biases if not designed and deployed with rigorous fairness considerations [110], [111]. Applying adversarial debiasing can help reduce this by training the model to be less predictive of protected attributes while maintaining accuracy [105].

- **Locating Systematic Bias-** Systematic bias can inadvertently creep into AI/ML systems during their development, affecting critical processes like credit decisioning, risk management, and compliance [105], [108]. To detect systematic bias, various methods and tools can be utilized. These include statistical calibration, use of regularizers, resampling data, and deploying fair machine learning models [108], [109]. Organizations should also monitor for bias before, during, and after modeling by ensuring equal performance across protected groups [105]. Tools like the Metrics package in R can help quantify and mitigate the impact of bias [105]. Research analyzing Home Mortgage Disclosure Act data found systematic bias in loan denial rates, with higher rejection rates for Asian, Black, and Hispanic borrowers compared to white borrowers [112]. This highlighted the potential for pre-existing biases to be replicated by AI-powered lending decisions [112]. To detect bias in AI/ML models, researchers have employed techniques such as ensuring equal probability of favorable outcomes across protected groups, maintaining equal predictive performance metrics, and minimizing statistical disparities [112]. These techniques help identify biases in data, model design, and outputs, enabling organizations to enhance the fairness of their AI systems [112]. Furthermore, having a diverse team of data scientists and engineers can help broaden the perspectives and identify potential sources of systematic bias [107].

5. Artificial Intelligence and Machine Learning in the Hiring Process

AI and machine learning have revolutionized the hiring process, offering organizations efficient and data-driven approaches to talent acquisition [113], [114]. These technologies enable automated screening, candidate evaluation, and decision-making, streamlining the process and reducing the impact of human bias [113], [114]. The potential benefits of AI and ML in hiring include increased efficiency, reduced time and costs, and more informed, evidence-based decision-making, ultimately leading to higher-quality hires and a more productive workforce [113], [114].

5.1 History of Artificial Intelligence Advancements in the Hiring Process

In the early 2000s, the recruitment process began seeing the initial adoption of AI/ML technology, primarily for automating repetitive tasks such as sourcing and screening candidates [115]. A key milestone in the 2000s was the integration of natural language processing (NLP), which enabled AI algorithms to understand and interpret human language with nuance, idioms, and emotional context [115]. This significantly enhanced the ability of AI to analyze job descriptions and resumes, leading to more accurate candidate-job matching.The emergence of AI-powered applicant tracking systems (ATS) in the late 2000s and early 2010s revolutionized the management of talent pools. These sophisticated systems could efficiently store, organize, and search through vast candidate databases, automatically ranking and prioritizing applicants based on their qualifications and suitability [115]. In the 2010-2015 period, AI saw significant breakthroughs and developments in its integration with recruitment. Key advancements included the automation of resume screening, intelligent candidate matching, enhanced candidate sourcing, predictive analytics in hiring, and AI-driven skill assessments [115]. From 2016 to 2020, the recruitment sector witnessed a rapid expansion of AI adoption. AI high-performing organizations were identified, investing more in AI efforts and gaining a competitive edge through improved efficiency, data-driven decision-making, and a stronger ability to attract talent [116]. In the most recent years, from 2021 to the present, AI has become widely adopted in recruitment, automating and streamlining various processes. The focus has shifted to promoting diversity and inclusion through unbiased candidate evaluation and diverse sourcing algorithms, while the role of the human recruiter remains crucial in interpreting AI-generated results [117].

5.2 Notable Artificial Intelligence and Machine Learning Technologies in the Recruitment Process

AI and ML algorithms have become increasingly prevalent in the recruitment process, revolutionizing how organizations identify, evaluate, and hire top talent. Some of the key AI and ML technologies employed in the recruitment sector include Natural Language Processing (NLP) algorithms, which enable machines to understand and interpret human language, facilitating the extraction of relevant information from resumes and cover letters [118]. Machine Learning (ML) models, trained on large datasets of past hiring data, can predict candidate suitability, rank applicants, and identify high-potential individuals [118]. Classification Algorithms, such as Support Vector Machines and Decision Trees, can efficiently filter out unqualified candidates based on predefined criteria [118]. Clustering Algorithms group candidates with similar skills and backgrounds, allowing recruiters to target specific talent pools [118]. Sentiment Analysis models evaluate candidates' attitudes, values, and personality traits through their written communication, providing insights into cultural fit [118]. Bias Detection and Mitigation algorithms are designed to identify and address biases in the hiring process, promoting fairness and diversity [118]. These AI and ML technologies are integrated into Applicant Tracking Systems (ATS) and other recruitment platforms, streamlining the entire hiring process, from resume screening to candidate matching and selection [118].

Fig -4: Illustration of an Applicant Tracking System (ATS) used in hiring processes. The image demonstrates how AI/ML algorithms screen and filter job applicants, highlighting potential biases that can affect candidate selection and diversity.

5.3 Artificial Intelligence Biases in The Job Recruitment Industry

The rapid adoption of AI and ML technologies in the hiring process has brought to the forefront the issue of algorithmic biases [119], [120], [121]. These biases can emerge from the historical data used to train the AI/ML models, as well as the inherent biases present in the human-designed algorithms and processes [122], [123]. Such biases can manifest in unfair treatment and exclusion of certain candidates based on factors like race, gender, age, and other personal attributes, ultimately undermining the principles of fairness and diversity in the workplace [119], [121], [124]. Addressing these biases is crucial to ensure equitable hiring practices and to prevent the perpetuation or amplification of societal inequalities through the use of AI/ML in recruitment [122].

5.4 Popular Types of Artificial Intelligence Biases found in The Job Recruitment industry

- **Algorithmic Bias-** Algorithmic bias is a prevalent issue in the hiring process, where artificial intelligence (AI) and machine learning (ML) models used for screening and selection can inadvertently perpetuate unfair biases [125], [126]. These biases can stem from the training data used to develop the algorithms, which may reflect historical patterns of discrimination and underrepresentation [126], [127]. For example, an AI-powered resume screening tool developed by Amazon was found to be biased against women, as the algorithm had learned to penalize resumes containing the word "women's", such as "women's chess club captain" [130]. This algorithmic bias was a result of the training data used, which reflected historical gender biases in hiring [130]. Such biases in AI-driven hiring tools can have adverse effects on organizations and broader society by limiting opportunities for qualified candidates from underrepresented groups [131]. As a result, the algorithms can make hiring decisions that systematically disadvantage certain demographic groups, leading to a lack of diversity and fairness in the workforce [125], [126], [129].

- **Gender Bias-** Another significant type of bias observed in the hiring process is gender bias [126], [128]. AI-powered hiring tools have been found to exhibit preferences for traditionally male-dominated roles and characteristics, often overlooking or undervaluing the skills and qualifications of female candidates. [126], [128]. A study by researchers at Harvard and MIT found that a widely used hiring algorithm systematically ranked male applicants higher than equally qualified female applicants [132]. The algorithm had been trained on historical hiring data that reflected traditional gender biases in the workforce, causing it to replicate and amplify these biases in the automated hiring process [132]. This highlights how AI/ML models can perpetuate societal biases if the training data is not carefully curated to be representative and unbiased [132]. This bias can have a detrimental impact on the representation of women in the workforce, particularly in industries like technology, where gender imbalance is already a longstanding issue [126], [128]. The perpetuation of gender bias in hiring can further exacerbate the gender pay gap and limit career advancement opportunities for women [126], [128].

5.5 Ethical Implications of Bias in Job Recruiting Artificial Intelligence and Machine Learning

The increasing deployment of artificial intelligence (AI) and machine learning (ML) technologies in the hiring process has raised significant ethical concerns regarding bias and fairness [133], [134], [126]. These systems, designed to enhance efficiency and objectivity, can inadvertently perpetuate existing societal biases and lead to unfair outcomes for job applicants [134], [137]. The incorporation of biased data in the training of these algorithms can result in discriminatory practices, such as unfairly filtering out candidates based on gender, race, or other protected characteristics [134], [136]. The lack of transparency and interpretability in the decision-making processes of AI/ML hiring systems further exacerbates these ethical issues, making it challenging to identify and address the underlying biases [134], [135]. Ethical considerations in this domain necessitate a holistic approach, involving diverse and representative datasets, enhanced algorithmic accountability, and continuous monitoring to ensure that AI-driven hiring practices uphold principles of fairness, equity, and non-discrimination [137]. Addressing these ethical concerns is crucial to maintain public trust and ensure that the benefits of AI/ML technologies in hiring are equitably distributed.

5.6 Strategies for Detecting Biases in Job Recruitment Artificial Intelligence and Machine Learning

Detecting biases in AI/ML systems used for job recruitment is a crucial step towards ensuring fair and equitable hiring practices. Researchers have developed various approaches and techniques to identify these biases, including analyzing the training data, evaluating the model's predictions for potential disparate impact, and auditing the decision-making process for sources of bias [139]. However, effectively detecting and mitigating biases in these complex systems poses significant challenges, as biases can stem from numerous sources, such as the data used for training, the design of the algorithms, and the human inputs guiding the process [138]. Additionally, the lack of transparency in many AI/ML systems can hinder the ability to thoroughly audit and understand the sources of bias, presenting a limitation in the comprehensive detection of these issues [139].

5.7 Detection of Famous Forms of Biases found in Hiring Process Artificial Intelligence

- **Detecting Algorithmic Bias-** Algorithmic bias refers to the systemic and unfair discrimination that can occur when AI algorithms produce biased outcomes, often reflecting or exacerbating societal biases [143]. One approach to detecting algorithmic bias is through algorithm auditing, whereby the algorithms used in hiring are subjected to rigorous testing to identify and rectify biases [143]. For example, the United Nations faced backlash over its use of a facial recognition tool in the hiring process, which exhibited racial bias. [145]. The tool consistently ranked candidates with darker skin tones lower than their lighter-skinned counterparts, reflecting biases inherent in the training data [145]. This is an example of algorithmic bias, where the AI system's outputs are skewed due to the biases present in the data and algorithms used to develop it [145]. Researchers have used various techniques and tools to detect the bias in this example, such as analyzing the performance of the AI system across different demographic groups and identifying the skewed distribution of the training data that led to the biased outputs [145]. These findings highlight the need for careful evaluation and mitigation of bias when deploying AI systems, particularly in high-stakes applications like hiring, to ensure fair and equitable outcomes [145]. Additionally, diversifying the training data used to develop these algorithms and incorporating fairness constraints into their design can help mitigate bias and promote equitable outcomes [143]. Researchers have also proposed the Conditional Demographic Disparity test, a publicly available tool that can identify bias in algorithms and provide guidance on how to make the system fairer and more accurate [141]. Beyond algorithm-focused methods, promoting diversity and inclusion within the organization can also help address algorithmic bias. Ensuring a diverse team of software engineers and data scientists, including individuals from different backgrounds, can broaden the perspectives and assumptions built into the algorithms [142].

- **Locating Gender Bias-** Gender bias in hiring refers to the tendency of hiring and recruiting algorithms to screen out job applicants based on gender [144]. One effective approach to detecting gender bias involves the use of Natural Language Processing (NLP) techniques, such as word embeddings and contextual embeddings. These methods can analyze the language used in job descriptions and identify gender-coded words or phrases that may deter certain candidates from applying [140]. Another prominent example of algorithmic bias detection in hiring is the case of Google's CV screening tool. In 2018, researchers found that the tool exhibited gender bias, systematically downgrading CVs that contained terms associated with women [146]. To uncover this bias, the researchers conducted a thorough analysis of the tool's inner workings and the data used to train it [147]. They examined the language models and algorithmic decision-making processes employed by the tool, identifying the specific ways in which it was discriminating against female applicants [147]. Additionally, the researchers carried out controlled experiments, submitting resumes with subtle differences to test the tool's responses [147]. These rigorous investigations allowed them to pinpoint the sources of bias, such as the tool's tendency to prioritize masculine-coded language and penalize references to women's activities or organizations [146]. The researchers' findings ultimately led Google to acknowledge the issue and take steps to address the bias in their hiring algorithms [147]. Machine learning models can also be trained to classify job descriptions as gender-neutral or biased, providing quantitative metrics to assess the level of bias [140]. In addition to algorithmic tools, the involvement of diverse human reviewers and feedback loops is crucial for detecting gender bias. Collaborating with teams that bring varied perspectives can help uncover conscious and unconscious biases that may have been overlooked by automated systems [140].

6. CONCLUSION

In conclusion, this research novel has provided a comprehensive exploration of the biases inherent in artificial intelligence (AI) and machine learning (ML) systems across critical sectors such as healthcare, hiring, self-driving cars, and financial systems. Our review highlights that biases in AI and ML algorithms stem from various sources, including biased training data, algorithmic design, and implementation practices, leading to significant ethical concerns and consequences.

The identified biases, such as measurement, confirmation, and many other biases, have profound implications for decision-making processes and societal outcomes. Case studies examined in this paper illustrate the real-world impact of biased AI systems, emphasizing the urgent need for fair and accountable AI development practices.

Addressing these challenges requires a multifaceted approach. Mitigation strategies proposed include improving the diversity of training data, implementing algorithmic fairness techniques, conducting regular audits and impact assessments, and fostering interdisciplinary collaboration among AI developers, ethicists, and policymakers.

Furthermore, the advancement of large language models (LLMs), such as GPT-4 and its successors, offers a promising avenue for mitigating biases in AI and ML systems. LLMs have the potential to enhance algorithmic transparency, interpretability, and fairness by enabling researchers to analyze and understand biases more effectively. Through their ability to process vast amounts of text data, LLMs can aid in identifying and mitigating biases during the development and deployment phases of AI systems.

As we move forward, it is imperative to prioritize ethical considerations and social responsibility in AI development. By adopting the proposed mitigation strategies and leveraging emerging technologies like large language models, we can work towards creating AI systems that not only drive technological advancements but also promote equity, transparency, and trust across various sectors.

ACKNOWLEDGEMENT

This research novel would not have been possible without the unwavering support and dedication of my son. His willingness to shoulder additional responsibilities at home has been nothing short of remarkable. From washing dishes and vacuuming the house to taking care of his younger sister, he has ensured that I had the time and space needed to immerse myself in researching and writing. His patience, understanding, and love have been a constant source of strength and motivation. I am deeply grateful for all his contributions and the sacrifices he made to help me achieve this dream.

I also wish to express my profound gratitude to God, whose grace and guidance have been a constant source of inspiration and comfort throughout this journey. It is through His blessings that I have found the perseverance and resilience to navigate the challenges of writing this research novel. His presence has been a guiding light, providing me with the clarity and courage to continue, even when the path seemed uncertain.

To my family, friends, and everyone who believed in me and this project, thank you for your encouragement and support. Your faith in my work has been a powerful motivator, and I am forever grateful for your kindness and generosity.

REFERENCES

[1] Viruses Broaden the Definition of Life by Genomic Incorporation of Artificial Intelligence and Machine Learning Processes, https://www.semanticscholar.org/paper/a3de306c56388b0b778f81b193782831d0c63295

[2] Detailed review on Integrated Healthcare Prediction System Using Artificial Intelligence and Machine Learning, https://www.semanticscholar.org/paper/2982cd698f6fd4f0d49363b9e4403bcf48e717b0

[3] Political Dimension of Artificial Intelligence and Machine Learning, https://www.semanticscholar.org/paper/d739d59ab964c371060576eec2f025e471f5cbf8

[4] Call for algorithmic fairness to mitigate amplification of racial biases in artificial intelligence models used in orthodontics and craniofacial health., https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/ocr.12721

[5] Ethical Considerations in AI Addressing Bias and Fairness in Machine Learning Models, https://jklst.org/index.php/home/article/download/133/108

[6] TRACING THE EVOLUTION OF AI AND MACHINE LEARNING APPLICATIONS IN ADVANCING MATERIALS DISCOVERY AND PRODUCTION PROCESSES, https://www.fepbl.com/index.php/estj/article/download/552/708

[7] Artificial Intelligence (AI) in Healthcare & Medical Field, https://www.foreseemed.com/artificial-intelligence-in-healthcare

[8] A Historical Look at AI in Health Care | WoundSource, https://www.woundsource.com/blog/historical-look-ai-in-health-care

[9] Contributed: Nine revolutionary ways AI is advancing healthcare, https://www.mobihealthnews.com/news/contributed-nine-revolutionary-ways-ai-advancing-healthcare

[10] Artificial Intelligence in Drug Discovery and Development, https://www.semanticscholar.org/paper/4e3cf1f761b8749afbac46ab949ed30896d3f44a

[11] Algorithmic Precision Medicine: Harnessing Artificial Intelligence for Healthcare Optimization, https://journalajb2t.com/index.php/AJB2T/article/download/190/375

[12] Deep Learning and Health Informatics for Smart Monitoring and Diagnosis, https://www.semanticscholar.org/paper/b2c55629cb98afc1c93e44c3730c67fa935d93fa

[13] A review of artificial intelligence applications in anterior segment ocular diseases, https://mehdijournal.com/index.php/mehdioptometry/article/download/1030/589

[14] Unleashing the potential: Artificial intelligence in urology for enhanced diagnosis, treatment, and personalized care, https://www.semanticscholar.org/paper/e9ab4f5cf821ea44ebe4fe678ab572360eebb97b

[15] Towards General Purpose Vision Foundation Models for Medical Image Analysis: An Experimental Study of DINOv2 on Radiology Benchmarks, https://www.semanticscholar.org/paper/cd281e29a0692b10f2617c91e2fae26b4eea17f3

[16] Multimodal Healthcare AI: Identifying and Designing Clinically Relevant Vision-Language Applications for Radiology, https://dl.acm.org/doi/pdf/10.1145/3613904.3642013

[17] Meta-research on reporting guidelines for artificial intelligence: are authors and reviewers encouraged enough in radiology, nuclear medicine, and medical imaging journals?,https://www.semanticscholar.org/paper/e21952c2a220934d0944fd232c5476e46fd50e5e

[18] Overcoming AI Bias: Understanding, Identifying and Mitigating ..., https://www.accuray.com/blog/overcoming-ai-bias-understanding-identifying-and-mitigating-algorithmic-bias-in-healthcare/

[19] What is Bias in Machine Learning & Deep Learning?, https://www.foreseemed.com/blog/bias-in-machine-learning

[20] Addressing AI and Implicit Bias in Healthcare | TechnologyAdvice, https://technologyadvice.com/blog/healthcare/ai-bias-in-healthcare/

[21] Shedding Light on Healthcare Algorithmic and Artificial Intelligence ..., https://minorityhealth.hhs.gov/news/shedding-light-healthcare-algorithmic-and-artificial-intelligence-bias

[22] AI Algorithmic Bias: Understanding its Causes, Ethical and Social Implications, https://www.semanticscholar.org/paper/ca6d1e3308929dc73952b041005eeeaca8d9e702

[23] The Measurement and Mitigation of Algorithmic Bias and Unfairness in Healthcare AI Models Developed for the CMS AI Health Outcomes Challenge, https://www.semanticscholar.org/paper/a4ffc1b0263fc2677035df3b24e638c7fd44b77e

[24] Measurement and Mitigation of Bias in AI: A Narrative Literature Review for Regulatory Science., https://www.semanticscholar.org/paper/6d9439fde9c3845e738944adaa4d2bc63dffb5d4

[25] Unmasking Bias in AI: A Systematic Review of Bias Detection and Mitigation Strategies in Electronic Health Record-based Models, https://www.semanticscholar.org/paper/736c31eef4edfeec88a2147066e4f2f3b17df1f0

[26] Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models, https://www.semanticscholar.org/paper/9e049e61f34b87df8d8642fb348d98b07fc9d514

[27] AI Algorithmic Bias: Understanding its Causes, Ethical and Social Implications, https://www.semanticscholar.org/paper/ca6d1e3308929dc73952b041005eeeaca8d9e702

[28] A meta-study on optimizing healthcare performance with artificial intelligence and machine learning, https://www.semanticscholar.org/paper/db6c79c32d1f6b4281b3ddf7eea73127896aa6c7

[29] Ethics of AI: Perpetuating Racial Inequalities in Healthcare Delivery and Patient Outcomes, https://www.semanticscholar.org/paper/5882b0e47a8fd405f8d3224806b31374d1fb44cb

[30] Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review, https://journals.plos.org/digitalhealth/article/file?id=10.1371/journal.pdig.0000022&type=printable

[31] Fairness And Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, And Mitigation Strategies, https://www.mdpi.com/2413-4155/6/1/3/pdf?version=1703577406

[32] Measurement and Mitigation of Bias in AI: A Narrative Literature Review for Regulatory Science., https://www.semanticscholar.org/paper/6d9439fde9c3845e738944adaa4d2bc63dffb5d4

[33] AI: simply reinforcing medicine's worst biases?,https://www.semanticscholar.org/paper/543d006629a93d7aef1e50bbf7b8155ccd338a9e

[34] PCIV method for Indirect Bias Quantification in AI and ML Models, https://www.semanticscholar.org/paper/b13a423a885fab193a57e930d063bbe2f48b507a

[35] A structured regression approach for evaluating model performance across intersectional subgroups, https://www.semanticscholar.org/paper b8d149465015a0baff6c8c9d4fb40b2070ebcdc4

[36] Considerations for addressing bias in artificial intelligence for ... - NCBI, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10497548/

[37] An objective framework for evaluating unrecognized bias in medical ..., https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9277645/

[38] Exploring AI Bias in Healthcare: From Gathering Data to ... - LinkedIn, https://www.linkedin.com/pulse/exploring-ai-bias-healthcare-from-gathering-data-krishan-dev

[39] Bias in AI-based models for medical applications - Nature, https://www.nature.com/articles/s41746-023-00858-z

[40] Understanding the role of AI bias in healthcare - Quantib, https://www.quantib.com/blog/understanding-the-role-of-ai-bias-in-healthcare

[41] The Next Era of Biomedical Research, https://www.semanticscholar.org/paper/9aed9e4b665d14de49aa6d175ed8b6bd321764c3

[42] Safeguarding Connected Health: Leveraging Trustworthy AI Techniques to Harden Intrusion Detection Systems Against Data Poisoning Threats in IoMT Environments, https://mesopotamian.press/journals/index.php/BJIoT/article/download/200/200

[43] Artificial Intelligence in Self-Driving Cars Research and Innovation: A Scientometric and Bibliometric Analysis, https://www.semanticscholar.org/paper/a831eea0fd7678022c2767ed273f291af4ce38d7

[44] AN IMPROVED SAFETY ALGORITHM FOR ARTIFICIAL INTELLIGENCE ENABLED PROCESSORS IN SELF DRIVING CARS, https://www.semanticscholar.org/paper/50901eb33468b74e147901ac68921670301f0ad2

[45] Artificial Intelligence Based on Self Driving Cars with Safety Algorithm, https://ijsrem.com/download/artificial-intelligence-based-on-self-driving-cars-with-safety-algorith m/?wpdmdl=29900&refresh=665b44a88548c1717257384

[46] Enhancing Object Detection in Self-Driving Cars Using a Hybrid Approach, https://www.mdpi.com/2079-9292/12/13/2768/pdf?version=1687671979

[47] Hogyan döntsön a mesterséges intelligencia? Az önvezető autók morális kérdései • How Should Artificial Intelligence Decide? Moral Challenges of Self-driving Cars, https://www.semanticscholar.org/paper/5a8541a51ca9af15677955dfaa97c3d0cab4dadd

[48] Machine Learning Insights for Behavioral Data Analysis Supporting the Autonomous Vehicles Scenario, https://www.semanticscholar.org/paper/31492340c57e9be0810f36433f2e9db4c8180887

[49] An Object Classification Approach for Autonomous Vehicles Using Machine Learning Techniques, https://www.mdpi.com/2032-6653/14/2/41/pdf?version=1675427213

[50] Assessing Artificial Intelligence for Humanity: Will AI be the Our Biggest Ever Advance ? or the Biggest Threat [Opinion], https://ieeexplore.ieee.org/ielx7/44/8558723/08558761.pdf

[51] Computational Imaging and Artificial Intelligence: The Next Revolution of Mobile Vision, https://arxiv.org/pdf/2109.08880

[52] Workshop on Les Contes du Mariage : Should AI stay married to Ed ?, https://www.semanticscholar.org/paper/8fcab5256b951d790b6251af27692ef106ae4a69

[53] Synergizing Remora Optimization Algorithm and Transfer Learning for Visual Places Recognition in Intelligent Transportation Systems and Consumer Electronics, https://www.semanticscholar.org/paper/c202f75265511bf404ff8122d0fff0342b15b1f8

[54] Adversarial Machine Learning Using Convolutional Neural Network With Imagenet, https://www.semanticscholar.org/paper/c993bcdf2f81fc7393c5c6dd923dc9c205d740df

[55] Unveiling the Power of Data: A Journey Through Machine Learning Techniques, https://www.semanticscholar.org/paper/5ce6e2e05f57d04fa925cfcba1b24f2751601992

[56] Autonomous Vehicles: Evolution of Artificial Intelligence and Learning Algorithms, https://www.semanticscholar.org/paper/4b9c0f7c583fc86f1687bac38bdfbb92e870ff83

[57] How Electrical Engineering and Computer Engineering Departments are Preparing Undergraduate Students for the New Big Data, Machine Learning, and AI Paradigm: A Three- Model Overview, https://www.semanticscholar.org/paper/cdaddf4011323a3d76e23bff229c2beed913a4f3

[58] Is Speech the New Blood? Recent Progress in AI-Based Disease Detection From Audio in a Nutshell, https://www.frontiersin.org/articles/10.3389/fdgth.2022.886615/pdf

[59] Automated and Autonomous Experiment in Electron and Scanning Probe Microscopy, https://www.semanticscholar.org/paper/da38521a67502b4ec35e24594de8a4a3cebfaba1

[60] AIMLAI: Advances in Interpretable Machine Learning and Artificial Intelligence, https://www.semanticscholar.org/paper/0b9bb8f8f7bebf05edd0680aa06a2e6cbd7dadd9

[61] Deep Networks That Know When They Don't Know, https://www.semanticscholar.org/paper/87c1a307bb5bcc237f45f49a07118e0b362e0a9d

[62] A Study on Training Data Selection for Object Detection in Nighttime Traffic Scenes, https://www.semanticscholar.org/paper/1060b18f79d5742e46a69a9db5d1d2f4beefca6f

[63] Confidence-Aware Reinforcement Learning for Self-Driving Cars, https://www.semanticscholar.org/paper/371b6f1b00a85913459c073d016669ad9d3c02ab

[64] Why Is Artificial Intelligence Blamed More? Analysis of Faulting Artificial Intelligence for Self-Driving Car Accidents in Experimental Settings, https://www.semanticscholar.org/paper/4cdaddec95061f1259bd97118c6857f8c4a6bc3e

[65] Gender Discrimination 1 GENDER DISCRIMINATION IN NEGOTIATOR DECEPTION: AN ARCHIVAL ANALYSIS OF BEHAVIOR IN THE MBA CLASSROOM, https://www.semanticscholar.org/paper/891e95fba49bb6a6d6d4563161f0229a93b9fec3

[66] Ethical Considerations in AI Addressing Bias and Fairness in Machine Learning Models, https://jklst.org/index.php/home/article/download/133/108

[67] Exploring Ethical Considerations in AI-driven Autonomous Vehicles: Balancing Safety and Privacy, https://jaigs.org/index.php/JAIGS/article/download/31/22

[68] AI Ethics in Autonomous Vehicles: Safety and Decision-Making, https://fpgainsights.com/artificial-intelligence/ai-ethics-in-autonomous-vehicles/

[69] AI in Transportation: Ethical Balance between Safety and Privacy, https://tech-stack.com/blog/ai-in-transportation/

[70] Overcoming Racial Bias In AI Systems And Startlingly Even In AI Self ..., https://www.forbes.com/sites/lanceeliot/2020/01/04/overcoming-racial-bias-in-ai-systems-and-startlingly-even-in-ai-self-driving-cars/

[71] AIMLAI: Advances in Interpretable Machine Learning and Artificial Intelligence, https://www.semanticscholar.org/paper/0b9bb8f8f7bebf05edd0680aa06a2e6cbd7dadd9

[72] What is availability bias – TechTarget Definition, https://www.techtarget.com/whatis/definition/availability-bias

[73] Addressing Disability and Ableist Bias in Autonomous Vehicles, https://dredf.org/2023/03/09/addressing-

disability-and-ableist-bias-in-autonomous-vehicles-ensuring-safety-equity-and-accessibility-in-detection-collision-algorithms-and-data-collection/

[74] Analyzing and Mitigating Bias for Vulnerable Classes - arXiv, https://arxiv.org/html/2401.10397v1

[75] The Psychological Effects of Self-Driving Cars, https://www.psychologytoday.com/us/blog/urban-survival/202311/the-psychological-effects-of-self-driving-cars

[76] Shifting Perceptions and Emotional Responses to Autonomous ..., https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10813087/

[77] Analyzing and Mitigating Bias for Vulnerable Classes - arXiv, https://arxiv.org/html/2401.10397v1

[78] Overcoming Racial Bias In AI Systems And Startlingly Even In AI Self ..., https://www.forbes.com/sites/lanceeliot/2020/01/04/overcoming-racial-bias-in-ai-systems-and-startlingly-even-in-ai-self-driving-cars/

[79] Eliminating Bias in Autonomous Vehicle Training Data - Keymakr, https://keymakr.com/blog/eliminating-bias-in-autonomous-vehicle-training-data/

[80] How Machine Learning in Automotive Makes Self-Driving Cars a ..., https://mindy-support.com/news-post/how-machine-learning-in-automotive-makes-self-driving-cars-a-reality/

[81] Explainable AI: Using Shapley Value to Explain Complex Anomaly Detection ML-Based Systems, https://ebooks.iospress.nl/pdf/doi/10.3233/FAIA200777

[82] A Critical Study of ICT Oriented Innovations based on Artificial Intelligence (AI) Applications towards BFSI Sector, https://www.semanticscholar.org/paper/5b6b0e2723d29ce8b7d4ecf192ee890e63f74a76

[83] Explainable AI in Credit Risk Management, https://arxiv.org/pdf/2103.00949

[84] The Use of Artificial Intelligence and Machine Learning in Forecasting the Financial Growth of Automobile Industries, https://ijsrem.com/download/the-use-of-artificial-intelligence-and-machine-learning-in-forecasting-the-financial-growth-of-automobile-industries/?wpdmdl=34451&refresh=665a28c264bbc1717184706

[85] A Review on Financial Fraud Detection using AI and Machine Learning, https://al-kindipublisher.com/index.php/jefas/article/download/6739/5709

[86] Predictive Maintenance Beyond Prediction OfFailures, https://www.semanticscholar.org/paper/cbd373cb11d8a7549f1e100f2ce910013cf1177e

[87] eLearning after four decades : what about sustainability?., https://www.semanticscholar.org/paper/1d41e1654f97e7fae3ba42b5ee5c7e72a3ea1c63

[88] JavaTM Intelligent Tutoring System Model and Architecture, https://www.semanticscholar.org/paper/158687a90aa04f1e887993032afca165f0c8a1d7

[89] A Holistic Approach To Teaching Engineering Entrepreneurship And Technology Commercialization, https://peer.asee.org/12612.pdf

[90] Transforming Information Extraction: AI and Machine Learning in Optical Character Recognition Systems and Applications Across Industries, https://www.semanticscholar.org/paper/8b789d617ad7265b16c3e2179738f0aa1aa89d8d

[91] Automated Banking Fraud Detection for Identification and Restriction of Unauthorised Access in Financial Sector, https://www.semanticscholar.org/paper/f9904a1e6d07816f236d52133a1e0621cf051792

[92] AI-Enhanced SEM Analysis: Evaluating E-CRM's Effect on Customer Experience in NCR's Private Banking Sector, https://www.semanticscholar.org/paper/54e623d8e54f447c959d91c04881eb9b804c3ed9

[93] Enhancing Cybersecurity in Financial Services using Single Value Neutrosophic Fuzzy Soft Expert Set, https://www.semanticscholar.org/paper/8337420bf3776456b76721185c4261c93c516cbb

[94] APPLICATION OF MACHINE LEARNING METHODS IN THE BANKING SECTOR TO INCREASE THE EFFICIENCY OF DECISION-MAKING, https://www.semanticscholar.org/paper/4ef5fcc2aece06a9a453d34fd7415dd71921a745

[95] Navigating the landscape of artificial intelligence, Machine learning and Deep Learning, https://www.semanticscholar.org/paper/96cc1c970f33dae04382bd4415a1c031b36f8dc3

[96] Towards algorithm auditing: managing legal, ethicaland technological risks of AI, ML and associated

algorithms, https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.230859

[97] Towards Algorithm Auditing: A Survey on Managing Legal, Ethical and Technological Risks of AI, ML and Associated Algorithms, https://www.semanticscholar.org/paper/8f31afed84dde6d3ca2793dbb53c98ba526e61ed

[98] The Challenges of AI Algorithm Bias in Financial Services, https://www.techopedia.com/ai-algorithm-bias-in-financial-

[99] Understanding algorithmic bias and how to build trust in AI - PwC, https://www.pwc.com/us/en/tech-effect/ai-analytics/algorithmic-bias-and-trust-in-ai.html

[100]The Challenges of AI Algorithm Bias in Financial Services, https://www.techopedia.com/ai-algorithm-bias-in-financial-services

[101]What Are the Ethical Concerns Around AI In Finance? - LinkedIn, https://www.linkedin.com/pulse/what-ethical-concerns-around-ai-finance-damilola-ogunsakin-aca-msc

[102]Bias, fairness, and other ethical dimensions in artificial intelligence, https://bankunderground.co.uk/2023/10/11/bias-fairness-and-other-ethical-dimensions-in-artificial-intelligence/

[103][PDF] AI IN FINANCIAL INDUSTRY: ETHIC ISSUES, https://jai.aspur.rs/archive/v1/n1/2.pdf

[104]Eliminating AI bias from bank decision-making - BAI, https://www.bai.org/banking-strategies/eliminating-ai-bias-from-bank-decision-making/

[105]Ensuring Fairness in Finance: Mitigating Bias in Credit Underwriting, https://medium.com/@shreyalunkad/ensuring-fairness-in-finance-mitigating-bias-in-credit-underwriting-b57a471d5e06

[106]Assessing and mitigating unfairness in credit models with ... - Microsoft, https://www.microsoft.com/en-us/research/publication/assessing-and-mitigating-unfairness-in-credit-models-with-the-fairlearn-toolkit/

[107]Combating AI bias in the financial sector - TechTarget, https://www.techtarget.com/searchenterpriseai/feature/Combating-AI-bias-in-the-financial-sector

[108]Fair Lending and Machine Learning Models: Navigating Bias and ..., https://www.experian.com/blogs/insights/fair-lending-and-machine-learning-models/

[109]What are the different regularization techniques in finance?, https://www.versaquant.com/blog/what-are-the-different-regularization-techniques-in-finance

[110]Analysis of the Potential Benefits from Using Quantum Computing for Aerospace Applications, https://www.semanticscholar.org/paper/85fbdd3258e0a87be02ca5c32fc4f5a6e071876a

[111]The Prospects of Quantum Computing for Quantitative Finance and Beyond, https://www.semanticscholar.org/paper/4bb0f88787a996b5865f44daaa3972f7e0c747c6

[112]Navigating the AI Bias: Exploring Tools and Techniques, https://arunapattam.medium.com/navigating-the-ai-bias-exploring-tools-and-techniques-c42b0f26fd29

[113]AVA: A Photorealistic AI Bot for Human-like Interaction and Extended Reality, https://www.semanticscholar.org/paper/6ce24d122db9b71130683071308edef1bffa9925

[114]THE ROLE OF ARTIFICIAL INTELLIGENCE (AI) IN RECRUITMENT AND SELECTION OF EMPLOYEES IN THE ORGANISATION, https://russianlawjournal.org/index.php/journal/article/download/1624/900

[115]AI in Recruitment - Everything you need to know - Jobylon, https://www.jobylon.com/blog/how-ai-is-transforming-the-world-of-recruitment

[116]The state of AI in 2022—and a half decade in review | McKinsey, https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review

[117]Revolutionizing Recruitment: The AI-Driven Transformation, https://www.herohunt.ai/blog/revolutionizing-recruitment-with-ai

[118]The Future of Recruitment: Use of AI for Better Hiring, https://www.expressanalytics.com/blog/ai-in-recruitment/

[119]Navigating the Pitfalls of AI in Hiring: Unveiling Algorithmic Bias, https://medium.com/@sahin.samia/navigating-the-pitfalls-of-ai-in-hiring-unveiling-algorithmic-bias-9e62b50b3f65

[120]Learn How AI Hiring Bias Can Impact You Recruitment Process, https://vidcruiter.com/interview/intelligence/ai-bias/

[121]How AI Bias Could Impact Hiring and Recruitment. - LinkedIn, https://www.linkedin.com/pulse/how-ai-bias-could-impact-hiring-recruitment-abdellah-aouf

[122]In the job hiring process, most workers say they already sense AI, https://www.cnbc.com/2023/12/28/in-the-job-hiring-process-most-workers-say-they-already-sense-ai.html

[123]Discrimination and bias in AI recruitment: a case study - Lewis Silkin, https://www.lewissilkin.com/en/insights/discrimination-and-bias-in-ai-recruitment-a-case-study

[124]Ethics and discrimination in artificial intelligence-enabled ... - Nature, https://www.nature.com/articles/s41599-023-02079-x

[125]The Implications of Orthographic Intraference for the Teaching and Description of ESL: The Educated Nigerian English Examples, https://www.semanticscholar.org/paper/ab3855d91d98fe9c3986f23d65282472a472e15a

[126]Antagonistic terms found on land and offshore, https://www.semanticscholar.org/paper/27350242237d6b3c8537fabffdc1829ff7fddcc3

[127]Scoring Procedures for the Evaluation of Spelling Performance, https://www.semanticscholar.org/paper/6058d32ae8bf172d446f9ece81526c9e4fb5ef64

[128]Functional load, therapeutic change, and the syntactic distribution of English homonyms, https://www.semanticscholar.org/paper/077d3b1fa9efa5d4c1b3844599eb271b1ec1df64

[129]A Linguistic Analysis of COVID-19 Neologisms in English: Bangladesh Perspective, https://ccsenet.org/journal/index.php/elt/article/download/0/0/49285/53184

[130]COSCFair: Ensuring Subgroup Fairness Through Fair Classification Framework, https://www.semanticscholar.org/paper/854cee242b69a792f00f91c0d94c1f022e07396f

[131]A Comprehensive Review of AI Techniques for Addressing Algorithmic Bias in Job Hiring, https://www.mdpi.com/2673-2688/5/1/19/pdf?version=1707285598

[132]XAI4LLM. Let Machine Learning Models and LLMs Collaborate for Enhanced In-Context Learning in Healthcare, https://www.semanticscholar.org/paper/27892516b6a20ddab6c5851efb5abaacdc077691

[133]Ethical Considerations in AI Addressing Bias and Fairness in Machine Learning Models, https://jklst.org/index.php/home/article/download/133/108

[134]Exploring Ethical Dimensions in AI: Navigating Bias and Fairness in the Field, https://jaigs.org/index.php/JAIGS/article/download/54/41

[135]Developing a conceptual technical framework for ethical AI in procurement with emphasis on legal oversight, https://www.semanticscholar.org/paper/56a78da193023c0120e7a7bf4c7950d5ae91c518

[136]Examining Ethical Aspects of AI: Addressing Bias and Equity in the Discipline, https://ojs.boulibrary.com/index.php/JAIGS/article/download/119/87

[137]Diversity, Equity, and Inclusion, and the Deployment of Artificial Intelligence Within the Department of Defense, https://www.semanticscholar.org/paper/0c2e9008fc770c9804aa2dcf52c95c2acb37c2a0

[138]Navigating the Pitfalls of AI in Hiring: Unveiling Algorithmic Bias, https://www.linkedin.com/pulse/navigating-pitfalls-ai-hiring-unveiling-algorithmic-bias-sahin-ahmed-qajrc

[139]Fairness in AI-Driven Recruitment: Challenges, Metrics, Methods ..., https://arxiv.org/html/2405.19699v1

[140]How AI Can Remove Bias From The Hiring Process And ... - Forbes, https://www.forbes.com/sites/kathycaprino/2021/01/07/how-ai-can-remove-bias-from-the-hiring-process-and-promote-diversity-and-inclusion/

[141]Future of Inclusive Hiring: The Rise of Algorithmic Affirmative Action, https://www.linkedin.com/pulse/future-inclusive-hiring-rise-algorithmic-affirmative-action-karmav-j28rc

[142]UNCOVERING BIAS IN AI-DRIVEN RECRUITMENT - Langley Search, https://www.langleysearch.com/blog/2024/04/uncovering-bias-in-ai-driven-recruitment-challenges-solutions-and-real-life-examples

[143] Detecting and mitigating bias in natural language processing, https://www.brookings.edu/articles/detecting-and-mitigating-bias-in-natural-language-processing/

[144] 7 Best Job Description Bias Tools (in 2024) - OngigBlog, https://blog.ongig.com/diversity-and-inclusion/job-description-bias-tools/

[145] Navigating the Pitfalls of AI in Hiring: Unveiling Algorithmic Bias, https://medium.com/@sahin.samia/navigating-the-pitfalls-of-ai-in-hiring-unveiling-algorithmic-bias-9e62b50b3f65

[146] UNCOVERING BIAS IN AI-DRIVEN RECRUITMENT - Langley Search, https://www.langleysearch.com/blog/2024/04/uncovering-bias-in-ai-driven-recruitment-challenges-solutions-and-real-life-examples

[147] Want a job? You'll have to convince our AI bot first | CBC News, https://www.cbc.ca/news/business/recruitment-ai-tools-risk-bias-hidden-workers-keywords-1.6718151

[148] **(Fig -1)** Chaudhury, Srijani. *Mean Squared Error (MSE) loss function*. 13 Mar. 2020. https://www.analyticsvidhya.com/blog/2019/10/mathematics-behind-machine-learning/. Accessed 15 June 2024.

[149] **(Fig -2)** eXeX. *eXeX and Neurosurgeon Dr. Robert Masson Achieve World First Using Apple Vision Pro.* 6 Feb. 2024. https://www.prnewswire.com/news-releases/exex-and-neurosurgeon-dr-robert-masson-achieve-world-first-using-apple-vision-pro-302054787.html. Accessed 15 June 2024.

[150] **(Chart -1)** Herlihy, Christine. *Mean absolute error of estimates for 6 metrics using 5 methods on diabetes data*. 3 June 2024. Accessed 15 June 2024.

[151] **(Fig -3)** Fox, Eva. *Current rendering of Tesla FSD*. 1 May 2021. https://www.tesmanian.com/blogs/tesmanian-blog/fsd-beta-v-9-visualization-is-a-true-mind-s-eye-of-the-neural-net-says-elon-musk. Accessed 15 June 2024.

[152] **(Chart -2)** Machine Intelligence Team, University of Toronto. *Trading strategy pipeline structure for the Wall Street Bots.* 21 June 2022. https://utorontomist.medium.com/wall-street-bots-building-an-automatic-stock-trading-platform-based-on-artificial-intelligence-4f49df89ac41. Accessed 15 June 2024.

[153] **(Fig -4)** Henderson, Robert. *Applicant Tracking Systems Illustrative Diagram*. 27 Mar. 2024. https://www.jobscan.co/blog/8-things-you-need-to-know-about-applicant-tracking-systems/. Accessed 15 June 2024.

BIOGRAPHIES

Sukanya Konatam is the principal author of the research, leading the idea of writing a comprehensive literature survey journal on AI and ML biases. With over 18+ years of experience in the IT industry, enriched by five years dedicated to specializing in AI and AI Governance. Her expertise spans a comprehensive range of IT disciplines, underscored by a profound depth of knowledge in the governance of artificial intelligence. This unique combination of experience positions her as a leading figure in the field, adept at navigating the intricacies of AI technology with a strategic and ethical approach. She has implemented data centric solutions for several industries including banking & financials, telecom, health care, automobile, criminal justice and many more.

Her significant contributions to the paper included an in-depth analysis of biases in the hiring process and self-driving cars, as well as authoring the abstract, conclusion, and ensuring overall grammatical accuracy. Sukanya's robust background in data governance, data warehousing, machine learning, and AI, combined with her proficiency in AI Governance made her instrumental in identifying and proposing mitigation strategies for AI biases. Additionally, she holds a postgraduate degree in Data Science, Machine Learning, and Artificial Intelligence.

	Venkat Konatam, a secondary author of this literature survey, played a key role in investigating artificial intelligence and machine learning biases within the financial systems sector. With over 25 years of experience in IT, Venkat has excelled in constructing resilient data platforms and delivering impactful data products. His expertise lies in data architecture, database design, and performance optimization, with a strong focus on AI and ML applications. Venkat is currently pursuing a postgraduate program in data science, machine learning, and artificial intelligence to deepen his understanding and foster innovation in these areas. Known for his leadership in data teams and strategic alignment with business goals, Venkat contributed significantly to identifying biases in financial systems and proposing effective mitigation strategies to ensure fair and ethical AI implementation.
	Shravya Konatam, another secondary author of this literature research novel, made substantial contributions by focusing on AI in healthcare, driven by her passion for learning and advancing in the medical field. Currently pursuing a BS DO Matriculant with a major in Biology and minors in Honors Transdisciplinary Studies and Pre-Health at Nova Southeastern University in Florida, Shravya is deeply committed to her field of study. Alongside her academic pursuits, she has a keen interest in exploring different cultures and loves to travel. Despite being in the early stages of her career, Shravya's dedication to understanding the intersection of AI and healthcare has been instrumental in providing insights into the ethical implications and biases in medical AI systems. Her fresh perspective and enthusiasm enrich the paper's exploration of AI biases in healthcare, emphasizing the importance of interdisciplinary collaboration in addressing these challenges.