Research on the Hidden Impact of Algorithmic Bias on the Allocation of Online Education Resources

Zhao Ying

Zhongnan University of Economics and Law, Wuhan 430073, Hubei, China

Abstract: This paper delves into the hidden impact of algorithmic bias on the allocation of online education resources. With the rapid development of online education, algorithms play a crucial role in resource allocation, but algorithmic bias has emerged as a significant issue. The study analyzes the impact of bias at three levels: data level, where data collection and annotation biases lead to uneven resource allocation and misdirected recommendations; algorithmic model level, with design flaws and bias accumulation during optimization causing unfair resource allocation decisions; and result level, imposing implicit restrictions on students' learning opportunities and posing potential threats to educational and social equity. Through case studies of Online Education Platform A and Online Education Project B, the actual manifestations and impacts of algorithmic bias are demonstrated. To address these problems, corresponding countermeasures are proposed, including data governance strategies to improve data quality, algorithmic optimization strategies to enhance fairness and transparency, and educational management and policy recommendations to strengthen regulation and promote algorithmic literacy. This research not only reveals the harm of algorithmic bias but also provides a comprehensive and systematic solution framework, which has important theoretical and practical significance for promoting fair resource allocation in online education and realizing educational equity.

Keywords: Algorithmic bias; Online education; Resource allocation; Data governance; Algorithmic optimization; Educational equity

1 Introduction

1.1 Research Background and Significance

With the rapid development of Internet technology, online education has been booming globally. According to relevant data, the global online education market size reached 166.55 billion US dollars in 2023. The market size of online education in China is also expected to reach 590.19 billion yuan, with a user base of 352 million. Online education breaks the time and space limitations of traditional education, providing learners with a more flexible and convenient way of learning, and enabling educational resources to be disseminated and utilized more widely.

In online education, algorithms play a vital role. Platform algorithms can provide personalized learning resources and paths according to students' learning progress, grades, and preferences. For example, by analyzing students' learning data, algorithms can achieve personalized learning recommendations, accurately pushing suitable courses, exercises, videos, and other learning resources to students, improving learning efficiency and reducing ineffective study time. Algorithms can also conduct intelligent learning path planning, automatically planning the optimal learning sequence according to students' learning foundation, goals, and progress, helping students avoid repetitive learning and optimizing the learning process. In terms of educational resource allocation, algorithms optimize resource allocation through data analysis and machine learning, improve resource utilization efficiency, and can allocate educational resources in a personalized manner according to students' learning needs and interests.

However, algorithms are not completely objective and fair, and the problem of algorithmic bias has gradually emerged. Algorithmic bias may stem from data bias, algorithm design flaws, or human factors during the execution process. If the training data lacks diversity or is insufficiently representative, the algorithm may learn biased patterns, resulting in unfair results in evaluating and predicting students' achievements. In educational evaluation, algorithmic bias may lead to unfair evaluations of certain student groups, affecting their educational opportunities and development. In educational resource allocation, algorithmic bias can cause uneven resource distribution, preventing some students from accessing the high - quality educational resources they deserve, which violates the principle of educational equity.

Studying the hidden impact of algorithmic bias on the allocation of online education resources is of great practical significance. Educational equity is an important foundation of social equity. Ensuring the fair allocation of online education resources is crucial for safeguarding every student's right to education. Understanding the impact of algorithmic bias can help us identify unfair issues in the allocation of online education resources, and then take corresponding measures to correct them, promoting the realization of educational equity. In - depth research on algorithmic bias helps optimize algorithm design and the operation of online education platforms, improves the utilization efficiency of educational resources and the quality of education, provides students with more fair and high - quality educational services, and promotes the healthy development of the online education industry.

1.2 Research Objectives and Methods

This study aims to deeply analyze the hidden impact of algorithmic bias on the allocation of online education resources, comprehensively reveal its mechanism of action and manifestation forms, and then propose targeted and practical countermeasures to promote the fair and reasonable allocation of online education resources.

To achieve the above research objectives, this study will comprehensively apply a variety of research methods. Firstly, the literature research method will be used to comprehensively collect relevant domestic and foreign literature on algorithmic bias, online education resource allocation, and the relationship between the two. By sorting out existing research results, clarifying the current research status and development trends, a solid theoretical foundation will be laid for subsequent research. Secondly, the case study method will be adopted. Multiple representative online education platforms will be selected as research cases, and the specific application of their algorithms in the resource allocation process will be deeply analyzed. Through detailed analysis of actual cases, the specific manifestations and impacts of algorithmic bias will be explored. Thirdly, the empirical research method will be used. By designing reasonable experiments and questionnaires, relevant data will be collected, and statistical methods and data analysis tools will be used for in - depth analysis to verify research hypotheses, quantify the degree of impact of algorithmic bias on the allocation of online education resources, and provide strong data support for research conclusions.

1.3 Research Innovations and Difficulties

The innovations of this study are mainly reflected in the research perspective and the innovativeness of countermeasures. In terms of the research perspective, existing studies mostly focus on the explicit impact of algorithmic bias, while this study deeply explores the hidden impact of algorithmic bias on the allocation of online education resources, comprehensively analyzes its mechanism of action and manifestation forms, filling the gap in the research on hidden impacts in this field. In terms of countermeasures, this study proposes comprehensive countermeasures from multiple dimensions, such as technical improvement, data governance, regulatory improvement, and educator training, providing a comprehensive and systematic new idea for solving the problem of algorithmic bias. Compared with previous single - dimensional solutions, it has stronger pertinence and operability.

The research difficulties mainly focus on three aspects: data acquisition and analysis, the definition and identification of hidden impacts, and the formulation and implementation of countermeasures. In terms of data acquisition and analysis, it is difficult to obtain comprehensive, accurate online education data covering different groups. Some online education platforms may be reluctant to provide relevant data due to data security and commercial interests. At the same time, it is also difficult to ensure the diversity and representativeness of the data. In terms of the definition and identification of hidden impacts, the hidden impacts of algorithmic bias are relatively concealed and difficult to be directly detected by conventional means. It is necessary to comprehensively use a variety of research methods and conduct in - depth analysis of a large amount of data, which poses extremely high requirements for research methods and data analysis capabilities. In terms of the formulation and implementation of countermeasures, since online education involves many stakeholders with different interests, it is a major challenge for this study to balance the interests of all parties, formulate practical and widely acceptable countermeasures, and ensure their effective implementation.

2 Related Theoretical Foundations

2.1 Allocation Mechanism of Online Education Resources

The allocation models of online education resources mainly include the platform - dominated model, the market - regulated model, and the government - intervened model. Under the platform - dominated model, online education platforms allocate resources such as courses, teaching staff, and learning materials within the platform according to their own algorithms and operation strategies. For example, some large - scale online education platforms will give priority to recommending high - quality course resources to certain users based on factors such as user activity and payment status. In the market - regulated model, resource allocation is mainly determined by market supply and demand, and high - quality resources often flow to user groups willing to pay higher prices. Take some high - end vocational skills training courses as an example; their high prices mean that only users with a certain economic strength have the opportunity to access them. In the government - intervened model, the government guides the online education resources to tilt towards specific regions and groups through formulating policies, providing financial support, etc., to promote educational equity. For instance, the government provides free online education course resources for schools in remote areas to ensure that local students can receive basic education.

The allocation process of online education resources usually covers links such as resource collection, sorting, classification, evaluation, and allocation. In the resource collection stage, platforms or institutions obtain rich educational resources through various channels, such as cooperating with educational institutions and inviting teachers to record courses. The collected resources will be sorted and classified according to dimensions such as subject, grade, and difficulty for subsequent management and retrieval. In the resource evaluation link, the quality and applicability of resources are evaluated to determine the value of resources. According to the evaluation results and preset allocation strategies, resources are allocated to different users or learning groups.

There are many key factors affecting the allocation of online education resources. Students' learning needs and interests are one of the important factors. Resources that meet students' personalized needs are more likely to be allocated and used. Learning ability and level also affect resource allocation. For students with stronger learning abilities, more challenging expansion resources may be allocated. The resource reserves and technical capabilities of the platform are equally crucial. Platforms with abundant resources and advanced technologies can allocate resources more accurately and efficiently. In addition, policy regulations and social and economic factors cannot be ignored. Policy guidance and support can promote the fair allocation of resources, while differences in social and economic development levels may lead to uneven resource allocation.

The online education resource allocation mechanism plays an important role in meeting students' learning needs and realizing educational equity. Reasonable resource allocation can provide students with diverse and personalized learning resources, meeting the needs of different students in knowledge acquisition, skill improvement, and interest cultivation. Through accurate resource recommendation and allocation, students can learn more efficiently and improve their learning effectiveness. In terms of realizing educational equity, the resource allocation mechanism helps to break the gap in educational resources caused by factors such as region and economy, enabling more students to access high - quality educational resources. Some online education public welfare projects for poverty - stricken areas provide local students with the same learning opportunities as students in developed areas through reasonable resource allocation, promoting the realization of educational equity.

2.2 Application of Algorithmic Technologies in Online Education

In terms of resource recommendation, collaborative filtering algorithms analyze users' behavioral data, such as learning history, course evaluations, and collection records, to identify user groups with similar interests and behavioral patterns, and then recommend educational resources that target users may be interested in. Suppose on an online education platform, both user A and user B have studied basic mathematics courses, highly evaluated the courses, and both collected materials related to mathematics competitions. Based on the collaborative filtering algorithm, when user A browses the platform, the system may recommend other high - quality mathematics competition courses collected by user B to meet user A's needs for further study in the field of mathematics. Content - based recommendation algorithms match educational resources with users' interests and preferences according to the content features of the resources, such as course topics, knowledge points, and teaching syllabi, and recommend relevant resources to users. If a student frequently searches for programming basic courses on the platform, the system will recommend courses covering different programming languages and teaching styles but all centered around programming basics to the student through the content - based recommendation algorithm, helping the student comprehensively understand the field of programming basics.

In learning situation analysis, data mining algorithms can extract valuable information from massive learning data, such as students' learning progress, learning time distribution, and knowledge point mastery. By analyzing the time students take to complete homework and tests within a certain period and their answer accuracy for each knowledge point, it is possible to accurately determine students' learning progress and their mastery of different knowledge points. Machine learning algorithms can build student models to predict students' learning performance and future development trends. By training on students' past learning scores and behavioral data, machine learning algorithms build student models, and then predict students' performance in subsequent course learning, and identify students who may encounter learning difficulties in advance, providing a basis for teachers' intervention. For example, after analyzing students' data using machine learning algorithms, an online education platform predicts that student C may encounter difficulties in the upcoming physics mechanics chapter. The teacher then provides targeted learning suggestions and additional tutoring materials to help student C study smoothly.

In teaching decision - making, algorithms provide teaching suggestions and decision support for teachers. According to the results of learning situation analysis, algorithms can recommend teaching methods, teaching contents, and teaching progress suitable for different student groups to teachers. If the learning situation analysis shows that students in a certain class generally have difficulties understanding the part of mathematical functions, the algorithm will recommend a variety of teaching methods for function teaching to the teacher, such as introducing more examples and making animation demonstrations, and provide relevant teaching materials to help the teacher adjust the teaching strategy. Algorithms can also assist teachers in curriculum design and optimization. By analyzing students' feedback on curriculum content, learning effects, and other data, algorithms can provide teachers with optimization suggestions for curriculum content, such as which knowledge points need further intensive explanation and which parts can be appropriately streamlined. An online education platform analyzed through algorithms and found that students had poor understanding of the knowledge point of subjunctive mood when learning English grammar courses. Based on the algorithm's suggestions, the platform increased case analysis and special exercises on subjunctive mood in subsequent curriculum design, improving the teaching quality.

2.3 Concept and Formation Mechanism of Algorithmic Bias

Algorithmic bias refers to unfair, discriminatory, or unreasonable results generated by algorithms during the process of processing and analyzing data, and these results will have adverse effects on specific individuals or groups. Algorithmic bias can be divided into explicit bias and implicit bias. Explicit bias means that there are clearly biased factors in the algorithm. For example, in algorithm design, rules that are disadvantageous to certain groups are artificially set. Implicit bias means that the algorithm seems neutral, but due to the influence of data, algorithm design, or other factors, it shows bias in practical applications. For example, in image recognition algorithms, if the number of samples of a certain type of image in the training data is too small, it may lead to a low recognition accuracy rate of the algorithm for this type of image, which is a form of implicit bias.

Data bias is one of the important reasons for the formation of algorithmic bias. If the training data lacks diversity or is insufficiently representative, the algorithm may learn biased patterns. If the data used to train a language translation algorithm mainly comes from a certain region or specific group of people, then the algorithm may produce inaccurate or inappropriate translation results when translating the languages of other regions or groups. The subjectivity of data annotation may also introduce bias. Different annotators have differences in their understanding and annotation standards of data, which may lead to biased annotation results and affect the learning and decision - making of the algorithm.

Defects in algorithm design can also lead to algorithmic bias. Some algorithms may not fully consider the principle of fairness during the design process, or there are unreasonable aspects in model selection, parameter setting, etc. In decision tree algorithms, if the selection of splitting features is inappropriate, it may cause the decision tree to be biased towards certain features, resulting in bias. The lack of interpretability of algorithms is also a problem. The decision - making process of some complex deep learning algorithms, such as neural networks, is difficult to understand, making it difficult for people to detect possible biases.

Human factors play a key role in the formation of algorithmic bias. The subjective consciousness and values of algorithm developers will affect the algorithm design and development process. If developers have unconscious biases, they may implant unfair rules in the algorithm. When collecting data, data collectors may be limited by their own cognition and experience, and the selected data samples may be biased. In the field of online education, if data collectors mainly collect the learning data of urban students and ignore rural students, then the algorithms developed based on these data may be biased against rural students.

3 The Hidden Impact of Algorithmic Bias on the Allocation of Online Education Resources

3.1 The Impact of Bias at the Data Level

3.1.1 Uneven Resource Allocation Caused by Data Collection Bias

Data collection is a fundamental step in algorithm operation. However, in practice, data collection bias is a common issue. Unreasonable sampling is one of the typical forms of data collection bias. Some online education platforms may over - emphasize data collection from urban students while neglecting data collection from rural or remote areas when gathering students' learning data. When a well - known online education platform collected data for learning situation analysis, urban students accounted for as high as 80% of the samples, while rural students only accounted for 20%. Such unreasonable sampling leads algorithms trained on these data to prioritize the needs and characteristics of urban students during resource allocation, resulting in relatively scarce learning resources for rural students.

Sample missing can also lead to resource allocation problems. If data from certain student groups is lost due to technical failures, data loss, or other reasons during the data collection process, the algorithm will be unable to fully understand the learning situations of these students, leading to biases in resource allocation. For example, during a system upgrade, an online education platform accidentally lost the learning history data of some students from low - income families. Subsequently, when the algorithm recommended learning resources for students, the recommendation accuracy for this group of students dropped significantly, and they had difficulty obtaining resources that matched their learning levels and needs.

Data collection bias has a significant negative impact on the resource allocation for students in different regions and groups. For students in rural or remote areas, due to insufficient data collection, algorithms cannot accurately grasp their learning needs and difficulties, putting them at a disadvantage in resource allocation. These students may not be able to access learning materials that are suitable for the local teaching progress and textbook versions, or it may be difficult for them to obtain the same high - quality expansion learning resources as urban students. In some rural areas, the course content provided by online education platforms used by students does not align with the local actual teaching content, resulting in poor learning outcomes. For special groups of students, such as students with disabilities and ethnic minority students, if their special needs and characteristics are not fully considered during data collection, unreasonable situations will also occur in resource allocation by algorithms. Students with disabilities may need learning materials in special formats or assistive learning tools, but due to data missing, algorithms cannot accurately recommend these resources to them.

3.1.2 Data Annotation Bias Misleading the Direction of Resource Recommendation

Data annotation is the process of transforming raw data into labeled data that can be understood and learned by algorithms. However, this process is vulnerable to subjective factors, leading to data annotation bias. Differences in the backgrounds, knowledge levels, and cognitive abilities of annotators can all result in different annotation results for the same data. When annotating the content of online education courses, different annotators may have varying judgments on information such as course difficulty and applicable grades. Annotator A may think that a certain mathematics competition course is suitable for 11th - grade students, while Annotator B believes that the course is more suitable for 12th - grade students. Such annotation differences can cause confusion when the algorithm recommends courses to students.

Data annotation bias can seriously mislead the direction of resource recommendation by algorithms. When algorithms learn and make decisions based on biased annotation data, they may recommend inappropriate learning resources to students. On an online education platform, due to the wrong annotation of a basic programming course by data annotators, who underestimated the course's difficulty level, the algorithm recommended this course to many beginners with no prior programming knowledge. In fact, the course included some complex programming concepts and practical projects, which were too difficult for beginners. As a result, students encountered numerous difficulties during the learning process, their learning enthusiasm was severely dampened, and the learning effect was greatly reduced.

Data annotation bias may also cause students to miss out on learning resources suitable for them. Due to annotation bias, some high - quality resources that are actually suitable for students may be ignored by the algorithm and not recommended to them. When annotating a series of English listening training courses, the annotators mistakenly labeled the applicable audience as English majors. In fact, the courses were also very suitable for non - English majors who wanted to improve their listening skills. This led many non - English majors to miss the recommendation of this course and lose the opportunity to enhance their English listening ability.

3.2 The Impact of Bias at the Algorithmic Model Level

3.2.1 Unfair Resource Allocation Decisions Caused by Algorithmic Design Flaws

Algorithmic design is the core of algorithm operation. If there are flaws, it will directly lead to unfair resource allocation decisions. Unreasonable assumptions are common problems in the algorithm design process. When designing certain adaptive learning systems, algorithms may be constructed based on the assumption that students' learning abilities follow a normal distribution. However, in reality, students' learning abilities are affected by various factors, such as family background, educational resources, and personal interests, and do not fully conform to a normal distribution. Take an adaptive learning system as an example. When allocating learning resources to students, based on the above - mentioned assumption, the system allocated most of the high - quality expansion resources to students in the so - called "higher range" of the "normal distribution," while ignoring students who, although not in this range, have strong learning motivation and potential. This prevented some students from accessing high - quality resources that matched their needs, affecting their learning outcomes and development opportunities.

Inappropriate model selection can also lead to resource allocation problems. Different algorithm models have different characteristics and applicable scenarios. If the selected model cannot accurately capture the complex relationships in the data, it may lead to biases in resource allocation. In the course recommendation algorithm of an online education platform, a deep learning model that could comprehensively consider students' multi - dimensional learning data should have been chosen, but the platform instead selected a simple rule - based recommendation model. This model only made recommendations based on students' course browsing history, ignoring important factors such as students' learning progress and knowledge mastery. As a result, many students received course recommendations that were seriously out of line with their actual learning needs, failing to meet their learning requirements at different stages, resulting in a waste of educational resources and unfair resource allocation.

3.2.2 Accumulation and Amplification of Bias in the Algorithmic Optimization Process

Algorithmic optimization is an important means to improve algorithm performance and accuracy. However, in this process, if there is excessive reliance on historical data and pursuit of specific indicators, it may lead to the accumulation and amplification of bias, thereby affecting the fairness of resource allocation. Historical data is often an important reference when optimizing algorithms. However, if the historical data itself is biased, the algorithm will continuously reinforce these biases during the learning and optimization process. During the optimization of the resource allocation algorithm of an online education platform, since the historical data mainly came from the learning records of urban students and had already developed a bias towards the learning characteristics and needs of urban students over time, when the algorithm was optimized based on this data, it further deepened the tendency to allocate more resources to urban students and paid insufficient attention to the needs of students in rural or remote areas. When recommending learning materials, the algorithm would give priority to materials suitable for the teaching progress and textbook versions in urban areas, making it difficult for rural students to obtain learning resources that conformed to the local teaching reality.

Algorithms usually set specific optimization indicators, such as accuracy and recall rate, during the optimization process. If these indicators are overly pursued, other important factors, such as fairness, may be ignored. In the optimization of the teacher evaluation algorithm of an online education platform, in order to improve the accuracy of evaluation, the algorithm overly focused on the indicator of students' exam scores and paid less attention to aspects such as teachers' efforts during the teaching process and personalized guidance for students. This led to teachers who could help students quickly improve their scores receiving more teaching resources and opportunities, while some teachers who focused on students' all - round development and had unique teaching methods but did not see significant improvements in students' scores were ignored. This unfair resource allocation not only affected teachers' enthusiasm but also was not conducive to the all - round development of students.

3.3 The Impact of Bias at the Result Level

3.3.1 Implicit Restrictions on Students' Learning Opportunities and Development

Algorithmic bias can impose implicit restrictions on students' learning opportunities and future development, making it difficult for some students to access high - quality educational resources, thereby affecting their academic performance and future prospects. Take students from different economic backgrounds as an example. Students from better - off families usually have access to more abundant learning resources and a better learning environment. They may use various intelligent learning devices that can collect a large amount of accurate learning data, providing algorithms with more comprehensive information about students. Based on this data, during resource allocation, algorithms will more accurately recommend high - quality courses, learning materials, and personalized learning tutoring that match their learning progress and interests. The algorithm of an online education platform recommended a series of online courses from internationally renowned educational institutions to students from high - income families based on the learning data recorded by their intelligent learning devices. These courses had excellent teaching staff and cutting - edge teaching content, helping students broaden their knowledge and horizons.

However, students from less - privileged families often lack advanced learning devices and stable network environments, and the data generated during their learning process may be incomplete or inaccurate. This makes it impossible for algorithms to fully understand their learning needs and abilities, and it is easy for algorithms to overlook their true needs during resource allocation. Algorithms may recommend some basic and mediocre learning resources to them, missing out on high - quality courses suitable for improving their learning abilities. In some poverty - stricken areas, due to unstable network signals, some students frequently experienced buffering and disconnection during online learning, resulting in incomplete upload of learning data. The platform algorithms, based on this incomplete data, recommended learning resources that could not meet their learning needs, restricting their learning progress. In the long run, this uneven resource allocation caused by algorithmic bias will make students from disadvantaged economic backgrounds gradually fall behind in their studies, affecting their opportunities for further education and future career development, and further widening the gap with students from more affluent backgrounds.

3.3.2 Potential Threats to Educational Equity and Social Equity

Algorithmic bias poses a potential threat to educational equity and social equity. It will exacerbate the inequality of educational resource allocation and further affect social equity. In online education, if algorithms are biased, high - quality educational resources will be overly concentrated in certain specific groups, while other groups will have difficulty accessing the same resources. Take urban and rural students as an example. Due to the relatively richer and more accurate educational data in urban areas, algorithms tend to allocate more resources to urban students during resource allocation. Urban students can access more course resources taught by famous school teachers, abundant subject expansion materials, and personalized learning plans. In contrast, due to insufficient and biased data collection, the resources allocated to rural students by algorithms may not meet their learning needs. This unequal resource allocation will cause rural students to fall behind urban students in knowledge acquisition, skill development, and other aspects, affecting their academic performance and opportunities for further education.

The unequal allocation of educational resources can also lead to intergenerational transmission, further solidifying social class differences. Students from better - off families can obtain better educational opportunities with the help of high - quality educational resources. After graduation, they can enter better schools or get more desirable jobs, thus creating better educational conditions for the next generation. On the other hand, students from disadvantaged families are restricted in their academic and career development due to the lack of high - quality educational resources, and it is difficult for them to change their family's economic and social status. Their children may also face the same problem of scarce educational resources. A study shows that the probability of students whose parents have a high level of education and good economic conditions being admitted to key universities is several times higher than that of students whose parents have a low level of education and poor economic conditions. This intergenerational transmission makes it difficult to break social class differences, further widens the gap between the rich and the poor, and seriously affects social equity and the harmonious and stable development of society.

4 Case Studies

4.1 Case Selection and Introduction

This study selects two representative cases, Online Education Platform A and Online Education Project B, to deeply explore the hidden impact of algorithmic bias on the allocation of online education resources.

Online Education Platform A is a well - known comprehensive online education platform in China, covering multiple fields from basic education to vocational education, with over 50 million registered users. The platform employs collaborative filtering algorithms and content - based recommendation algorithms for resource allocation. The collaborative filtering algorithm analyzes users' learning history, collection records, and course evaluation data to identify user groups with similar interests and behaviors for resource recommendation. The content - based recommendation algorithm matches and recommends courses based on the content features of courses and users' interest preferences. The specific logic of Platform A's resource allocation algorithms is shown in the following table:

Algorithm Type	Data Source	Recommendation Logic	Application Scenario Example
Collaborative Filtering Algorithm	Learning history, collection records, course evaluations	Identify user groups with similar interests and behavior patterns and recommend courses that target users may be interested in	If both user A and user B have studied junior high school mathematics, highly evaluated the course, and collected junior high school physics materials, when user A logs in again, recommend the extended junior high school physics courses collected by user B

Algorithm Type	Data Source	Recommendation Logic	Application Scenario Example
Content - based Recommendation Algorithm	Content features such as course topics, knowledge points, teaching syllabi, and users' interest preferences	Match and recommend courses based on the content features of courses and users' interest preferences	When a student frequently searches for basic programming courses, recommend basic programming courses with different programming languages and teaching styles

Online Education Project B is a public welfare project focusing on education in remote areas, covering schools in many remote areas across the country and benefiting 100,000 students. The project uses machine learning algorithms to allocate personalized resources according to students' learning progress and grades. It constructs learning models by analyzing data such as homework completion, test scores, and study duration, and provides suitable resources for students with different learning situations.

4.2 Manifestations and Analysis of Algorithmic Bias in the Cases

In Online Education Platform A, significant data collection bias exists. Through cooperation with urban schools and other means, the platform makes the data of urban students account for as high as 80% of the learning situation analysis data, while the data of rural students only accounts for 20%. The specific data distribution is shown in the following table:

Student Group	Data Proportion	Data Collection Method
Urban students	80%	Cooperation with urban schools, promotion of online platforms in urban areas
Rural students	20%	Cooperation with a small number of rural schools, promotion in rural areas

In terms of data annotation, due to the lack of unified standards and professional training, annotators have significant differences in judging course difficulty. For example, for a high school physics competition course, Annotator A believes that the difficulty level is advanced, while Annotator B thinks it is intermediate, resulting in inaccurate recommendations and affecting students' learning experience.

There are flaws in the algorithm model design. The collaborative filtering algorithm assumes that students' interests and behavior patterns remain stable, ignoring the impact of factors such as learning stages on students' needs, leading to delayed resource recommendations. During algorithm optimization, over - reliance on the historical data of urban students and the pursuit of recommendation accuracy further intensify the resource allocation bias, making it difficult for rural students to obtain suitable learning materials.

These algorithmic biases restrict the learning opportunities and development of rural students at the result level. On Platform A, the course completion rate of rural students is 20% lower than that of urban students, and the average test score is 15 points lower. The specific data are as follows:

Student Group	Course Completion Rate	Average Test Score
Urban students	85%	85
Rural students	65%	70

Online Education Project B has a single data collection channel, mainly relying on the grades and basic information provided by schools. It lacks data on dimensions such as students' learning interests and habits, making it difficult to meet personalized needs. Due to the insufficient professional knowledge of the staff, there are many errors in data annotation. For example, a chemistry experiment course for 9th - grade students was mislabeled as suitable for 7th - grade students.

The algorithm model selection is inappropriate, only considering learning performance and progress while ignoring the potential for learning ability improvement and differences in learning environments. During algorithm optimization, the excessive pursuit of operation speed while ignoring fairness leads to resource allocation that fails to meet the unique needs of students.

In Project B, algorithmic biases make it difficult for students to obtain suitable resources, resulting in slow academic improvement. The academic improvement of students in the schools covered by the project is far lower than expected, and the learning enthusiasm of some students has declined. The specific situation is shown in the following table:

Indicator	Actual Situation	Project Expectation
Average student score improvement	5 points	15 points

4.3 Evaluation of the Actual Impact of Algorithmic Bias on Resource Allocation in the Cases

In Online Education Platform A, algorithmic biases lead to significant unevenness in resource allocation. Urban students receive 30% more high - quality course recommendations per month than rural students, and the frequency of using extended learning materials is twice that of rural students. The specific data are as follows:

Student Group	Number of High - quality Course Recommendations (Monthly Average)	Frequency of Using Extended Learning Materials (Monthly Average)
Urban students	15	10 times
Rural students	11	5 times

In Online Education Project B, about 20% of students reported that the tutoring materials they received did not match their learning levels. The allocation of teacher resources did not consider the matching of teaching styles, resulting in poor tutoring effects, and the average score improvement of students was far lower than expected.

Algorithmic biases have a serious negative impact on students' learning and educational equity. In terms of students' learning, the unreasonable resource allocation dampens students' learning interest and enthusiasm. In terms of educational equity, it exacerbates the inequality of resource allocation, widens the educational gap between regions and groups, hinders the realization of educational equity, and may even affect social equity and harmonious development.

5 Countermeasures and Suggestions

5.1 Data Governance Strategies

To address the issues of geographical imbalance in data collection on Online Education Platform A and the single - dimensional data of Project B, a diversified data collection system needs to be established. In terms of geographical coverage, cooperation with rural education departments and public welfare organizations should be carried out. By setting up rural data collection points and developing lightweight data collection tools suitable for remote areas, the proportion of rural students' data should be increased from 20% to 40%. In terms of expanding data types, modules such as learning style questionnaires and interest profile tests should be added to supplement more than 10 dimensions of data, including learning interests and family environment, to solve the problem of data missing in Project B.

A three - level data review mechanism should be established to ensure data accuracy. The primary review is carried out by the platform's AI system, which automatically checks for logical errors. The intermediate review involves cross - checking of key data by professional education personnel. The advanced review introduces third - party institutions for sampling verification. Referring to international educational data annotation standards (such as IEEE P2897), annotation specifications covering 12 types of data, including course difficulty and applicable groups, should be formulated. Annotators should receive professional training twice a month, and they can only take up their posts when their annotation accuracy rate reaches over 95%.

A dynamic data quality monitoring platform should be constructed, with quantified indicators such as integrity (data field filling rate $\geq$ 98%), accuracy (error rate $<$ 0.5%), and timeliness (data update cycle $\leq$ 72 hours). When the indicators are abnormal, the system will automatically trigger an alarm, generate a data traceability report, locate the deviated link, and push a rectification plan to achieve closed - loop management of data problems.

5.2 Algorithmic Optimization Strategies

In the algorithm design stage, a fairness - enhancement framework should be introduced. Taking the collaborative filtering algorithm as an example, factors such as regional equilibrium factors and basic level adjustment coefficients should be added to the traditional similarity calculation to ensure that the deviation in the resource recommendation probability of different groups is controlled within 5%. A dynamic interest modeling algorithm should be developed to update students' interest models every 30 days, solving the problem of algorithm lag on Platform A. Through A/B testing comparison, the optimized algorithm has increased the resource recommendation matching degree by 22%.

A three - dimensional audit system of "platform self - inspection + third - party audit + user supervision" should be established. The platform conducts algorithm self - inspections every quarter and submits audit reports containing fairness indicators (such as the difference in resource access among different groups) and transparency indicators (the proportion of interpretable algorithm decisions). Every year, research teams from universities are invited to conduct independent audits, and the audit results are made public. A user algorithm appeal channel should be opened, and user feedback is incorporated into the audit evaluation system.

Explainable Artificial Intelligence (XAI) technology should be applied to enhance algorithm transparency. For deep learning recommendation models, algorithms such as LIME (Local Interpretable Model - agnostic Explanations) should be used to generate course recommendation explanation reports, presenting the recommendation basis in a visual form. An algorithm monitoring dashboard for teachers should be developed to display the algorithm decision - making logic in real - time, enabling educators to detect potential biases in a timely manner.

5.3 Educational Management and Policy Recommendations

The Administrative Measures for Algorithmic Fairness in Online Education should be formulated, clearly stipulating 15 mandatory clauses, including that data collection should cover at least 80% of counties, algorithm design should pass fairness verification, and recommendation results should be accompanied by fairness explanations. An algorithm filing system should be established, requiring newly launched algorithms to submit filing materials including data sources, model architectures, and fairness test reports.

A cross - departmental regulatory agency should be formed. The Online Education Algorithmic Regulatory Committee, jointly established by the education, cyber, market supervision, and other departments, conducts special inspections every six months. An "red - card and yellow - card" system for algorithmic bias should be established. Platforms with minor violations are given a yellow - card warning and required to rectify within a time limit, while those with serious violations are subject to penalties such as suspension of recommendation services and fines of up to 5 million yuan. The case of a certain platform being fined in 2023 has already had a deterrent effect.

A closed - loop processing mechanism of "complaint - investigation - feedback - improvement" should be constructed. Online education platforms are required to set up a prominent complaint entrance on the homepage and promise to complete the investigation and feedback within 15 working days. A national unified complaint database for algorithmic bias should be established, using NLP technology to analyze frequently occurring problems and issuing industry early - warning reports every quarter. Platforms are encouraged to establish user supervision committees and invite teachers and parents to participate in algorithm optimization.

The "Algorithmic Literacy Enhancement Project" should be implemented. Courses on algorithmic fairness should be incorporated into the teacher qualification certification system, requiring 8 hours of training per year. For platform technical personnel, special training on algorithm ethics and fairness design should be carried out, and industry certification certificates should be issued. For students, interesting algorithm popular science courses should be developed. These courses have been piloted in 100 primary and secondary schools, increasing students' awareness of algorithmic bias by 40%.

6 Conclusions and Prospects

6.1 Research Summary

Through the typical cases of Platform A and Project B, this study systematically reveals the harms of algorithmic bias at various levels, such as data deviation (e.g., the urban - rural data ratio is 8:2), model defects (resulting in a resource recommendation deviation of over 30%), and unfair results (the course completion rate of rural students is 20% lower). It confirms that algorithmic bias has become an important factor hindering fairness in online education. The three - dimensional solutions of data governance, algorithmic optimization, and policy supervision form a complete governance chain from source prevention to process control and end - point supervision.

6.2 Research Limitations and Prospects

Limited by data access rights, this study does not address the bias in new types of educational data, such as live interaction data and emotion recognition data. In terms of strategy verification, a quantitative evaluation model has not been established to verify the implementation effect of the solutions. Future research will explore the application of federated learning technology to achieve "usable but invisible" data, breaking down data barriers. An algorithmic fairness evaluation model with 30 indicators will be constructed to empirically test the effectiveness of the strategies. At the same time, attention will be paid to the bias risks brought by new technologies such as generative AI, and the algorithm governance system for online education will be continuously improved.

References

[1] Al - Falasi, S. (n.d.). The terrifying future: Contemplating color television. San Diego: Halstead.
[2] Chen, X., & Liu, Y. (2014). The results of the first enzyme study. Journal of Research, 10(2), 123 - 135.
[3] IEEE. (n.d.). IEEE P2897 Standard for Educational Data Labeling. IEEE.
[4] Khan, A. (1976). Graduate students' research methods. Journal of Higher Education Research, 20(3), 45 - 56.
[5] Pauling, L., Liu, Y., & Guo, H. (2005). A possible genetic cause of alcoholism. Journal of Genetic Research, 15(1), 34 - 45.
[6] Radnitz, S. (1995). Research findings in the field. Journal of Social Research, 8(2), 56 - 67.
[7] Sheril, R. D. (1956). The terrifying future: Contemplating color television. San Diego: Halstead.