International Journal of Future Generation Communication and Networking
Vol. 13, No. 4, (2020), pp. 3997–4006

Gender Discrimination-Aware Classification Model for Course Selection of Higher Studies using Real-World Data-Set

Kirti Soni
Research Scholar, Department of Computer Science &Application
RKDF University, Bhopal, India

Dr. Sharad Gangele*
Professor, Department of Computer Science & Application
RKDF University, Bhopal, India

**sharadgangele@gmail.com*

Abstract—The gender discrimination in the rural region of India for higher education context is still widely common in this modern era. Women are insufficiently represented in the academic system where gender discrimination is the major factors. This paper analyzes and maps the structure and evolution of the gender ratio and differences in higher education subject-wise course such as Mathematics, Biology, Arts, Commerce and Agriculture. This research investigates a gender discrimination-aware model for higher education in rural society of India. In this paper, the association rule classification model is proposed for effectively filtering out the discriminatory patterns to present predictive accuracy. The result of this paper presents psychological factor and thought for women weakness in the field of mechanical and heavy production engineering.

Keywords—Gender Discrimination, Association Rule Classification, Higher Education, Course Selection

****Corresponding Author:** Dr. Sharad Gangele, Professor, Department of Computer Science & Application, RKDF University, Bhopal (MP)*

1. Introduction

Participation of women in higher education and jobs in several fields (mechanical, automobile, industrial and production engineering etc.) are the sign of economical and social advancement has attracted appreciable attention to numerous researchers and world-wide organizations [1]. To study and synthesize gender equality as well as involvement of women in higher education and jobs, many initiatory reports have been presented[2]. Significant gender gaps and lack of women representations in various fields of industries is presented by many national research data as well as salary difference, general promotion, priority of employment [3–4]. The female academics are continuously publishing very less articles in various academic journals and conferences as com- pared to their male academics counterparts. The report of gender disparity in academic ranking and contribution in re- search organization represents the social and psychological impact to the women representation [5–9].

Gender discrimination in higher education, jobs, labor marketing possibilities, health, economic authorization, political empowerment and legal rights, have been a permeative characteristics around the globe in the recorded history. The gender discrimination level are extremely different across the world so there is no accurate and transparent

ISSN: 2233-7857 IJFGCN

chronically affected trends in gender inequality [10].

The persistent gender gap has prompted many studies seeking to identify different explanatory factors in various areas of science, across different time periods, and in diverse national settings. Much of this research has identified factors related to family formation and childrearing as being the most influential causes of women under-representation in academia. The surveyed data collected from different villages of MP state as presented in Table 1, these surveyed data is utilized by gender discrimination-aware classification model and it developed a prediction model [11–12].

2. Factors of Gender Equality Index

Research on gender in academical record shows that the most important factors producing gender inequality at universities related to the images of science, scientific practice and the ideal scientist. Data from many Universities revealed that engineering branches like mechanical, automobile, industrial production, civil engineering etc. have lack of female students.

Table 1: Higher Studies Survey Report Subject-Wise of Rural Region

S. No.	Village Name	Maths		Agriculture		Commerce		Biology		Arts		Total
		BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	
1	BADGONA JOSHI	0	0	0	0	1	1	0	0	0	1	3
2	BADNOOR	1	2	0	0	0	0	1	2	0	0	6
3	BAGWANI	0	4	0	4	0	3	0	0	0	6	17
4	BHANDKHAPA	1	0	0	1	0	0	1	4	0	0	7
5	BHUTAI	0	0	5	0	0	0	0	1	0	0	6
6	BISAPUR KALA	12	7	16	16	1	0	1	14	9	12	88
7	CHARGAON	2	2	0	0	0	0	3	4	1	0	12
8	DEV VARDHA	0	0	0	0	0	0	0	0	0	2	2
9	GOREGHAT	0	0	0	0	0	0	0	1	0	2	3
10	IMLIKHEDA	0	1	0	0	0	0	1	0	0	1	3
11	JHILMILI	4	0	4	0	5	0	0	0	5	1	19
12	JUNNARDEO	19	0	4	2	14	0	3	9	4	7	62
13	KAMTHI	4	15	30	26	16	9	3	20	0	0	123
14	KARABOH	0	0	0	0	0	1	3	1	2	3	10
15	KUKDA CHIMAN	1	0	2	2	0	0	0	0	0	1	6
16	LINGA	8	3	2	1	4	14	2	7	6	5	52
17	LONIYA KARBAL	1	0	0	0	0	1	1	1	0	1	5
18	MAINIKHAPA	1	0	3	1	1	0	1	2	0	0	9
19	MOHKHED	13	22	2	0	0	0	4	23	2	6	72
20	PARTALA	0	0	0	0	0	0	0	3	0	3	6
21	SAORI	6	4	0	0	0	3	3	14	4	11	45
22	SARANGBHIRI	7	0	6	2	1	1	3	2	0	1	23
23	SHIKARPUR	1	1	5	0	0	0	0	0	6	0	13
24	SHIVLAL DHANA	2	0	1	0	0	1	0	0	0	0	4
25	TIKADI	0	0	0	0	0	0	2	1	0	0	3

ISSN: 2233-7857 IJFGCN

26	TURKIKHAPA	5	4	16	1	0	0	2	0	0	1	29
27	UMRANALA	10	2	20	13	0	1	1	6	0	2	55
	Total	98	67	116	69	43	35	35	115	39	66	683
		165		185		78		150		105		

There are many other fields where gender discrimination can be found. Gender equality has recently developed a multidimensional indicator of gender equality that enables comparison between countries and an evaluation of evolution over time, according to different dimensions of gender equality. According to these indicators, many countries cover quite a wide spectrum of gender equality cases [13]. Clearly, some dimensions of the gender equality measures are more significant than others. Thus, many countries score similarly on health indicators, but they are characterised by large discrepancies in relation to knowledge, time and power, with intermediate levels of dispersion in the domains of work and money. Gender equality affects many factors as presented in Figure 1 [14].

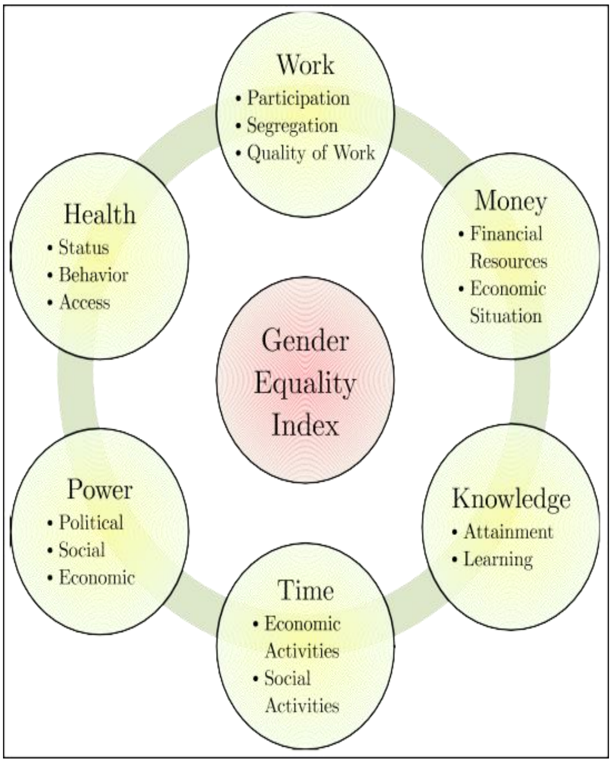

Figure 1: Gender Equality Index

Forms of discrimination such as occupational segregation by gender, harassment and discrimination related to pregnancy and family responsibilities can affect student's opportunities, treatment and ability to complete their studies. They also limit faculty member"s career satisfaction, advancement, and economic opportunities.

ISSN: 2233-7857 IJFGCN

Figure 2: Girls Academic Lifecycle

Gender discrimination may result in teachers leaving the educational institution or students not graduating and entering the health workforce. This has consequences for the quality and scale of health services, particularly since gender discrimination primarily affects female health workers, who constitute a large proportion of many countries health workforces, and who also face a continuation in the workplace of the types of gender discrimination experienced in preservice education [15–16]. Cultural beliefs as well as gender norms and stereotypes create environments in which sexual harassment and/or assault are normalized but not reported, and perpetrators are unpunished. It can be difficult for some students to concentrate on or complete their coursework while being threatened, harassed, or assaulted by other students or teachers. In addition, gender-blind institutional policies and practices prevent or limit female students from participating in classes, practical, and other curricular offerings by failing to consider student's family responsibilities or potential safety issues. The academic lifecycle of female students is illustrated in Figure 2 Gender discrimination is a hottest topic and issues not only in India, as a whole world. Gender inequality is mostly suffering by women and girls. In India time of baby girl birth discrimination is start and same environment on work place. Female faculty too much problems suffered like mentally and physically harassment, negative environment, lose confidence, job dissatisfaction and job attitudes. Female faculty discriminated against because of race, religious, sex and age factor. In academic sector working women who work under the male management or seniors to do the work under the pressure [17].

3. Literature Review

A new classification model for learning unbiased models on biased training data was introduced by Kamiran *et al.* [18]. This method is based on massaging the dataset by making the least intrusive modifications which lead to an unbiased dataset. On this modified dataset they then learn a non- discriminating classifier. Classification models are trained on the historical data for the prediction of the class labels of unknown data samples. Often, however, the historical data is biased towards certain groups or classes of objects.

Sankay *et al.* [19] investigated the effect of gender in- equality on economic growth in Nigeria. Using panel data regression from 35 states in Nigeria from 2008 to 2016, They discovered that gender inequality in the civil service impacts negatively on state generated revenue and economic growth at large. Although the study shows that women participation in paid employment impacts positively on economic growth, it is not a sufficient condition to escape poverty as women still make-up for a higher share of low-wage work-force. Their result suggests that closing the gender gap in the workforce and promoting equi-gender representation that is proportionate to the population distribution of male and female in all sectors of the economy is a key prerequisite for empowering women, reducing their dependency, improving their socio economic status and achieving economic growth.

Krishnan *et al.* [20] created qualitative classification models by applying feature selection algorithm on various set of input attributes. It was observed that in as per feature selection algorithm, gender is taken as the highest priority while filtering the input variables to get the accurate results. Various data models were created as per different classification algorithms and their precision and accuracy rate is com- pared. The analysis has been conducted using data mining concept with three data sets of 1227 students from four different HEIs (Higher Education Institutions) in Sultanate of Oman. The Collected dataset was undergone a pre- processing process and later data mining tasks such as training and testing was applied to it to generate classification model. The case study reveals that female students were showing high academic performance than male students in all three data sets. This experiment can be extended to find other attributes which has high impact other than gender in the academic performance of students of HEIs.

Luo *et al.* [21] proposed DAAR (discrimination-aware association rule) classification algorithm that provides unbiased decision making support in data analytics processes. They showed that DAAR is able to address the discrimination issues occurred on sensitive attributes, while having a minimal impact on the classification accuracy. DAAR uses DCI, a new discrimination measure, to prune rules that discriminate based on sensitive attributes, such as race and gender. The rules that pass the confidence support-DCI filter will form the final DAAR rule set. To classify new instances, DAAR uses majority voting and a sum of DCI scores. They empirically evaluated the performance of DAAR on three real datasets from traffic management and finance domains, and compared it with two non-discrimination-aware methods (a standard AR classifier and the state of the art AR classifier SPARCCC), and also with the discrimination- aware decision tree DADT. The experimental results on all datasets consistently showed that DAAR is capable of providing a good trade-off between discrimination score and accuracy – it obtained low discrimination score while its accuracy was comparable with AR and SPARCCC, and higher than DADT. An additional advantage of DAAR is that it generates a smaller set of rules than the standard AR; these rules are easy to use by the users, in helping them make discrimination-free decisions.

4. Problem Formulation

The gender discrimination problem can be labelled by en example of job recruitment for Mechanical engineers in an industry. The *N* candidate dataset *D* having *n* attributes $A=\{A_1,A_2,...A_n\}$, from this attributes, the requirement attribute is identified as $R=\{R_1, R_2,...R_n\}$. The discretionary dataset can be acquired by putting correlation between „candidate" attributes and „requirement" attributes. For a recruitment process, the probability of selection of male candidate P_m and female candidate P_f are estimated as

$$P_m = P(\text{marks} = \text{good} \mid \text{gender} = \text{male}) \quad ...(1)$$

$$P_f = P(\text{marks} = \text{good} \mid \text{gender} = \text{female}) \quad ...(2)$$

P_m is given higher preference over P_f, so it is considered that this dataset is biased against the female candidate for mechanical engineer recruitment process.

The discrimination classifier DC measurement is calculated as

$$\text{DC} = | P(A = A_t \mid R = R_t) - P(A = A_t \mid R = R_f) | \quad ...(3)$$

This score is computed on the testing set using the predicted class labels. The goal is to learn a classifier with low discrimination score with respect to *S*, with minimal impact on the classification accuracy.

As an example assume that we are designing a recruitment system for a company to predict if a new candidate is suitable for a job or not. If the historical data contains more males than females, the prediction model may tend to favour the attribute gender. A prediction rule using gender or sensitive attribute like marital status, may achieve high accuracy, but it is not acceptable as it is discriminating, which is both unethical and against the law. Sensitive at- tributes such as gender, race and religion should be taken as an information carrier of a dataset, instead of distinguishing factors. Females may be less suitable for a given job as on average they might have less work experience or lower educational level. It is acceptable to use work experience and educational level in the prediction model.

5. Proposed Method

The measurement is designed by the degree of discrimination using probability of the class rule (R)

$$R = \frac{|P(C = low|gender = female) - P(C = low|gender = male)|}{P\ (C = low|gender = female) + P\ (C = low\ gender = male)}$$

If the sensitive attribute does not appear in that rule at all,
We define discrimination measurement to be 0. Therefore, the range of discrimination measurement is [0, 1). When DCI equals to 0, which means the probability of the class value to be y is the same given different sensitive attribute values, the rule is considered to be free of discrimination. Otherwise, discrimination measurement is monotonically in- creasing with the discriminatory severity of a rule, which means that the larger discrimination measurement is, the more discriminatory the rule is with regard to the sensitive attribute S.
If the sensitive attribute S is binary with values S_{male} and S_{female}.
The best case is when D is zero, which means that the probabilities of the class value to be C_{target}, for all different values of the sensitive attribute, are the same,
i.e. their value to be is no discrimination. Otherwise, higher D corresponds to higher discrimination severity. As the testing dataset has been labeled by the classifier, higher discrimination in the dataset indicates the classifier is biased, which should be prevented.

6. Result Analysis

Using probability class rule of discrimination aware classifier, the applicant statistics is presented in Figure 3 it shows the predicted ratio of male and female candidate.

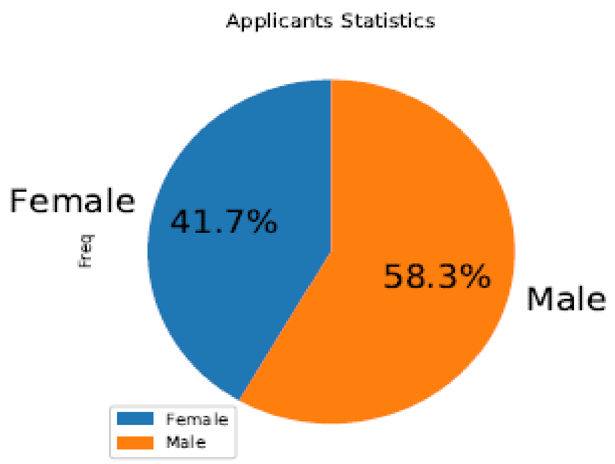

Figure 3: Predictive Male vs Female Ratio

The accepted and rejected ration for female candidate is presented in Figure 4 and the ratio of male to female candidate selection by various departments is presented in figure 5. The result with the help of association rule classification model shows that the prediction ratio of female to male is 41.7:58.3 along with predicted "accepted vs rejected" ratio for female is 39.3:60.7.

ISSN: 2233-7857 IJFGCN

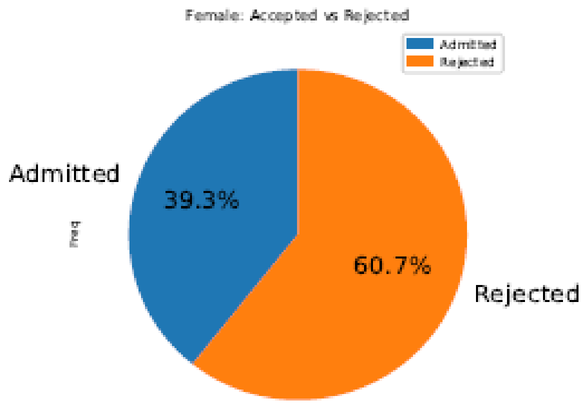

Figure 4: Predicted Accepted vs Rejected Ratio for Female

Predicted male Vs female ratio by various departments has also be compared in bar chart which is useful in discrimination measurement predicted model Four out of six departments have higher female admission rate. So, there is no evidence for gender discrimination. But overall female admission rate is much lower than male.

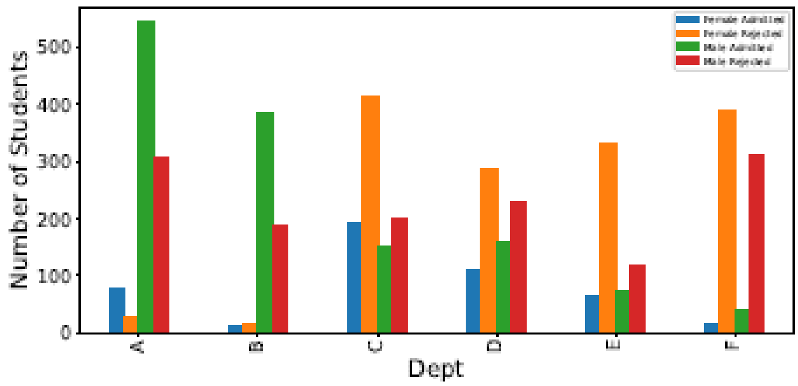

Figure 5: Predictive male vs Female ratio by Various Departments

This happens as various departments have high acceptance rate, but very few female applicants.

7. Conclusion

This exploratory result and analysis presents discrimination prediction of various departments. The result provides a comprehensive measure of equality between women and men relevant to the policy framework. The results have shown that the society is halfway towards gender equality, although there are large differences between male-female ratios in how close they are to the equality point. This research explains the ratio of gender differences for higher education subject-wise course such as Mathematics, Biology, Arts, Commerce and Agriculture. For higher education in rural social development of India, the proposed method investigated a gender discrimination-aware model. The proposed association rule classification model found to be useful for analysis of discriminatory patterns prediction. Finally, the result presented in this research explores psychological image and persuasion about women weakness especially in the department of mechanical engineering and heavy work jobs.

REFERENCES

[1] T. Dehdarirad, A. Villarroya, and M. Barrios, "Research Trends in Gender Differences in Higher Education and Science: a Co-word Analysis", Scientometrics, Vol. 101, No. 1, pp. 273–290, 2014.

[2] S. Stack, "Gender, Children and Research Productivity", Research in Higher Education, Vol. 45, No. 8, pp. 891–920, Dec 2004. doi: https://doi.org/10.1007/s11162-004-5953-z

[3] E. Leahey, "Gender Differences in Productivity: Research Specialization as a Missing Link," Gender & Society, Vol.20, No. 6, pp.754–780, 2006. doi: https://doi.org/10.1177/ 0891243206293030

[4] L. A. Hunter and E. Leahey, "Parenting and Research Productivity: New Evidence and Methods," Social Studies of Science, Vol. 40, No. 3, pp. 433–451, 2010. doi: https://doi.org/10.1177/0306312709358472

[5] D. K. Ginther and S. Kahn, "Does Science Promote Women? Evidence from Academia 1973-2001," National Bureau of Economic Research, Working Paper 12691, November 2006. doi: http://dx.doi.org/10.3386/w12691

[6] S. Schmitz, Race and Gender Discrimination across Urban Labor Markets: Susanne Schmitz, ser. Garland Studies in the History of American Labor & Urban studies Vol. 21. Routledge, 2018

[7] A. V. Dwivedi, Gender Discrimination, ser. The SAGE Encyclopaedia of World Poverty. Sage, 2015

[8] K. Soni and S. Gangele, "An Exploratory Survey on Improvement of Higher Education System and Learning Methods Using Big Data Analysis," in Journal Current Science, Vol. 20, No. 1, pp. 1–4, 2019

[9] S. Klasen, "From 'metoo' to boko haram: A Survey of Levels and Trends of Gender Inequality in The World," World Development, Vol. 128, p. 104862, 2020. doi: https://doi.org/10.1016/j.worlddev.2019.104862

[10] K. Soni and S. Gangele, "Predictive Analysis for Higher Studies in Rural Region Using Linear Regression Model," International Journal of Advanced Science and Technology, vol. 29, no. 8s, pp. 3829–3840, 2020.

[11] K. Soni and S. Gangele, "Predictive Analysis for Higher Studies in Rural Region Using Linear Regression Model," International Journal of Advanced Science and Technology, Vol. 29, No. 8s, pp. 3829–3840, 2020.

[12] S. Gangele, K. Soni, and S. Patil, "Data Mining Approach towards Students Behavior Assessment Methods for Higher Studies," International Journal of Computer Applications, Vol. 181, No. 30, pp. 11–14, Nov 2018. doi: http://www.ijcaonline.org/archives/volume181/ number30/30171-2018918099

[13] E.-M. Sent and I. van Staveren, "A Feminist Review of Behavioural Economic Research on Gender Differences," Feminist Economics, Vol. 25, No. 2, pp. 1–35, 2019. doi: https://doi.org/10.1080/13545701.2018.1532595

[14] N. Le Feuvre, "Contextualizing Women's Academic Careers in Cross-National Perspective," GARCIA Working Paper Series, Tento, Tech. Rep., 01 2016.

[15] M. Dufwenberg and A. Muren, "Discrimination by Gender and Social Distance," Stockholm University, Department of Economics, Research Papers in Economics 2002:2, Jan. 2002. doi: https://ideas.repec.org/p/hhs/sunrpe/ 2002 0002.html

[16] M. Castillo, R. Petrie, M. Torero, and L. Vesterlund, "Gender Differences in Bargaining Outcomes: A Field Experiment on Discrimination," Journal of Public Economics, vol. 99, no. C, pp. 35–48, 2013. doi: https: //ideas.repec.org/a/eee/pubeco/v99y2013icp35-48.html

[17] D. Sharma and C.Venkateswaran,"Impact of Gender Discrimination on Professional Life of Working Women in Education Sector of Haryana Universities", International Journal of Engineering and Advanced Technology (IJEAT), Vol. 9, No. 3, pp. 2008–2013, Feb 2020. doi: https://dx.doi.org/10.35940/ijeat.B3452.029320

[18] F. Kamiran and T. Calders, "Classifying without discriminating," in 2009 2nd International Conference on Computer, Control and Communication, 2009, pp. 1–6. doi:

ISSN: 2233-7857 IJFGCN

https://doi.org/10.1109/IC4.2009.4909197

[19] O. J. Sankay, A. K. J. Yi, Z. Othman, and S. Jusoh, "Gender Inequality and Economic Growth in the Nigerian Civil Service", International Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol. 9, No. 2, pp. 1375–1382, Dec 2019. doi: https://dx.doi.org/10.35940/ ijitee.B6238.129219

[20] R. Krishnan, H. Beegum, and P. Sherimon, "Cademic Performance Based on Gender Using Filter Ranker Algorithms-an experimental Analysis in Sultanate of Oman", International Journal of Engineering and Advanced Technology (IJEAT), vol. 8, no. 6S, pp. 2008–2013, Aug 2019. doi: https://dx.doi.org/10.35940/ijrte.F1211.0886S19

[21] L. Luo, W. Liu, I. Koprinska, and F. Chen, "Discrimination- Aware association Rule Mining for Unbiased Data Analytics," in Big Data Analytics and Knowledge Discovery, S. Madria and T. Hara, Eds. Cham: Springer International Publishing, 2015, pp. 108–120. doi: https://doi.org/10. 1007/978-3-319-22729-0 9