From: Mariia Luzan, luzan@umich.edu

Christopher Brooks, brooksch@umich.edu

15/02/2024

To whom it may concern,

We are writing to address the comments provided by the reviewers on the paper "Evaluation of the instance weighting strategy for transfer learning of educational predictive models", which we submitted for the AI4Ed Workshop at AAAI, 2024. First and foremost, we would like to express our gratitude to the reviewers for their thorough evaluation and valuable feedback.

The following table contains our responses to the reviewers' analysis and comments:

	Comment / Cons	Response
Reviewer v2nX	The results will be more convincing if it provides some statistical significance of the experiments	Statistical significance has been included in p.5: "The null hypothesis was tested to determine if the two AUC values (direct and weighted transfer) were equal: z = 2.378, p-value = 0.017; the standard deviation of the AUC difference for the testing purposes was estimated using a paired stratified bootstrap procedure."
	Detailed visual representations of the methodology could enhance comprehension for readers, making the approach more accessible	Visual representations have been incorporated into p. 4
	The paper could also benefit from a deeper discussion on potential real-world challenges when using transfer learning across institutions	Additional elaboration on real-world challenges has been integrated into p. 8, in the last two paragraphs.
Reviewer SdQt	The differences between the compared methods aren't very significant.	The difference between the methods under comparison wasn't expected to be substantial due to the similar characteristics of the two groups of students from a single institution. Despite the slight difference in the AUC values, this difference is statistically significant
	The dataset is quite limited, and results might not be generalizable to realistic settings.	This limitation is identified in the Discussion section, indicating the necessity for further analysis using an alternative dataset with more different student groups
	There aren't novel advancements in the methods.	While the applied transfer method itself lacks new advancements, its novel application lies in predicting dropout rates among undergraduate students.
	Question: When choosing a threshold of 10% students to be intervened, what would the evaluation metrics (TPR, precision, FPR, etc) be for Direct transfer and IWS?	Additional metrics have been incorporated into p. 6: "The recall values corresponding to this 10% threshold for direct transfer and weighted transfer are 55.1% and 56.7%, respectively. The precision values are 7.1% and 7.3%, respectively. These precision values are expected to be low due to a dropout rate of only 1.2%, with

		university support directed towards the most struggling 10% of students.”
	fewer resources for for this -> fewer resources for this	Incorporated, p. 1
	β are unknown on practice -> β are unknown in practice	Incorporated, p. 3

We appreciate your consideration of our revised paper.

Sincerely,

Mariia Luzan

Christopher Brooks