Resume of AI and bias

Artificial intelligence is a set of theories and methods used to allow computers to simulate human intelligence. Machine learning is a branch of AI that focuses on statistical or mathematical approaches in order to give computers the ability to learn without being explicitly programmed. Many industries and enterprises use machine learning algorithms in their processes, this can be the attribution of credits or hiring process etc. To be able to realize those tasks, machine learning algorithms will use datasets as inputs, but the encountered problem is that those datasets can be explicitly or not biased. Explicitly biased when informations of candidate are explicitly described in the datasets (the gender, the race etc…) or implicitly with what we call “latent variables”, an example is given in the first article (we know that the person is a teacher in primary school or she is a secretary so the person is probably a women) the algorithm by learning those implicit or explicit informations can be biased and the predictions that it will make also subject to bias. This happens because the algorithm tries to learn the distribution of the data, but in some cases the data have been manually or automatically generated by humans and human decisions in everyday life are subject to various biases.

In these articles, one of the recurrent biases is gender bias, the algorithm in this case will discriminate women against men because it will learn this implicit or explicit bias during the learning procedure.
The good aspect of using AI is that it will simplify a lot the process and the enterprise will reduce some costs (example : with an algorithm that automatically classify cv, the cost will be reduce because we don’t need to have many humans whose will be manually read this)
The main problem is the fact that those algorithms can implicitly learn the biases inside the data and so many kinds of discriminations will occur.
One solution to this problem which has been proposed is algorithmic transparency, so we should tell a customer why a decision has been made, and which elements of their data were the most significant.
for my viewpoint, I think that upwind processes should be done to try to debias the dataset or to debias the model trained