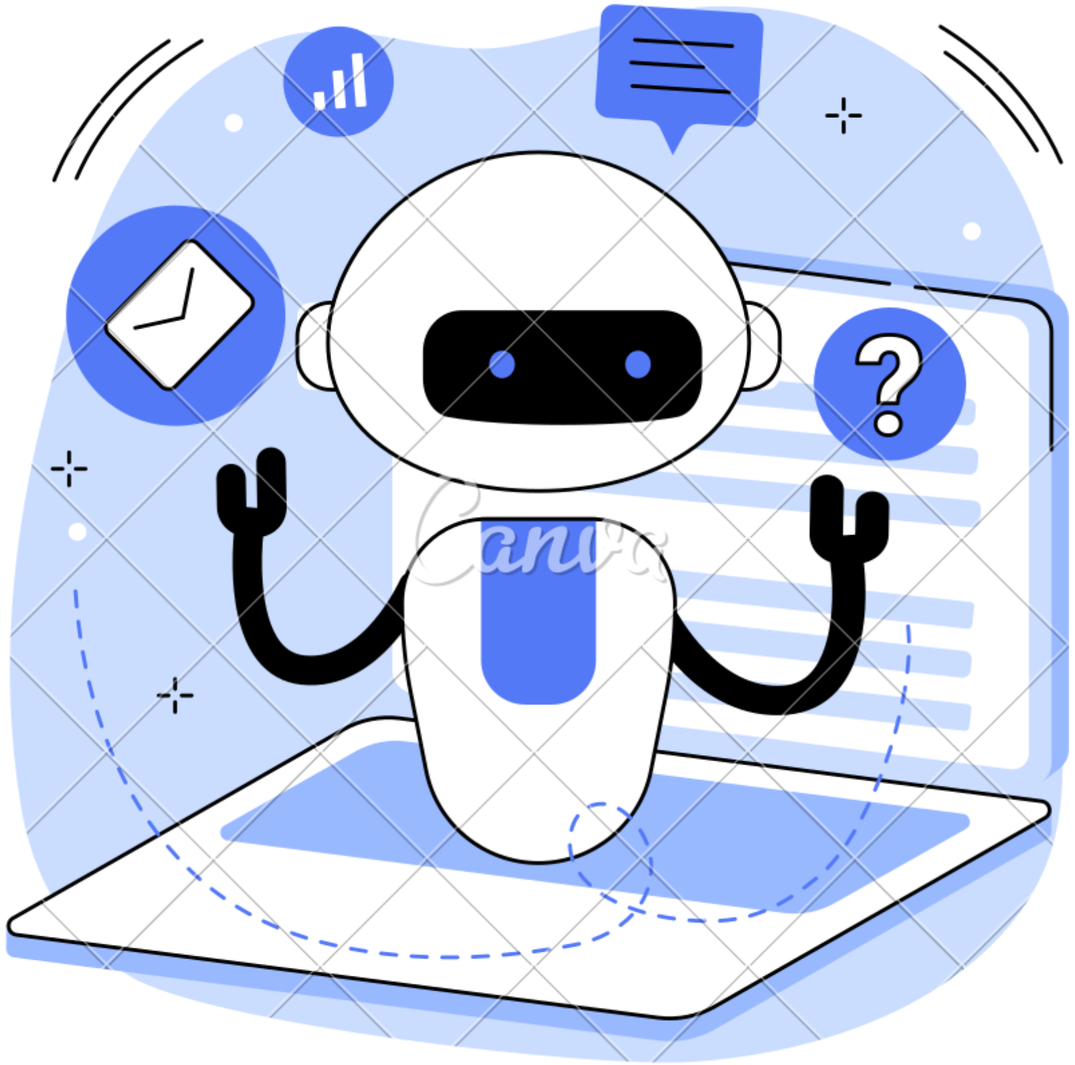
5 UNEXPECTED SOURCES OF BIAS IN A.I.

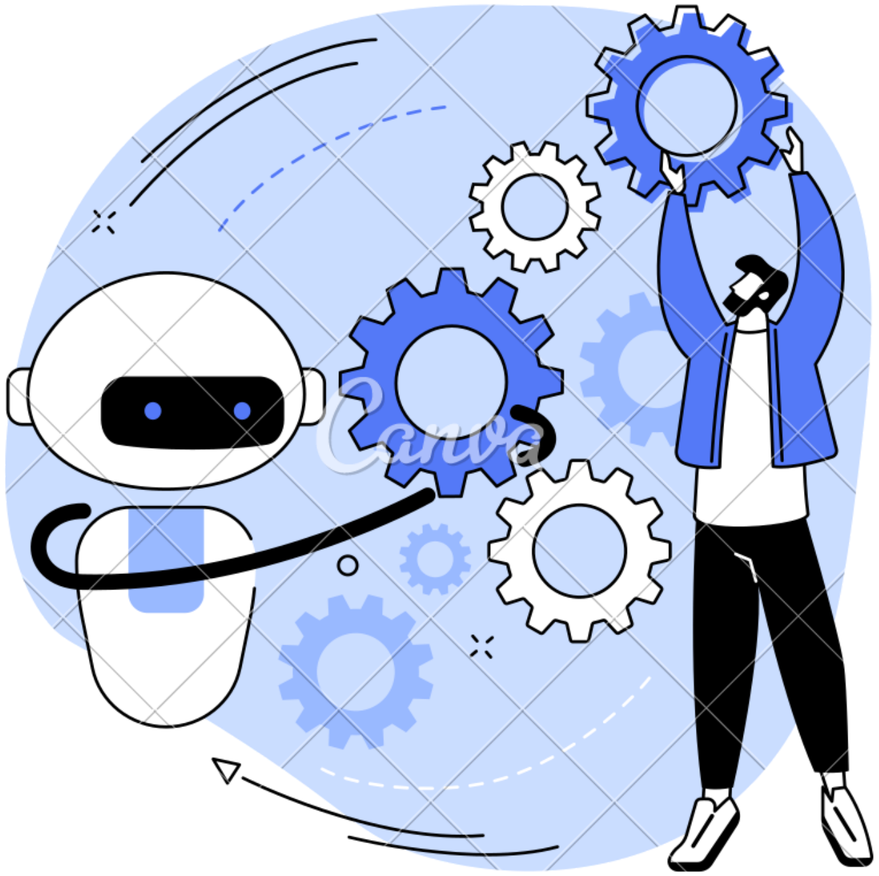
WHAT IS ARTIFICIAL INTELLIGENCE?

Artificial Intelligence (AI) is the use of computer systems to simulate human mental processes, such as interpreting and generating language.



EXAMPLES OF ARTIFICIAL INTELLIGENCE

- Chatbots
- Smart assistants
- E-Payments
- Search algorithms
- Media streaming



- Smart cars
- Navigation apps
- Facial recognition
- Text editors
- Social media feeds

BIAS IN A.I.

It is a phenomenon that occurs when an algorithm produces results that are systemically prejudiced due to erroneous assumptions in the machine learning (ML) process.

BIAS IN ARTIFICIAL INTELLIGENCE - HISTORICAL BIAS

Training data often reflects historical biases present in society. If historical data is biased, the AI system could perpetuate those biases. For instance, if past hiring practices favored a certain gender or ethnicity, a model trained on that data might make similar biased recommendations.

BIAS IN ARTIFICIAL INTELLIGENCE - PROXY VARIABLES

AI systems may inadvertently learn bias from proxy variables that are correlated with the target variable but not causally related. For example, if a hiring model learns that certain educational institutions are associated with certain demographics, it could perpetuate bias based on school names rather than qualifications.

BIAS IN ARTIFICIAL INTELLIGENCE - DATA PREPROCESSING TECHNIQUES

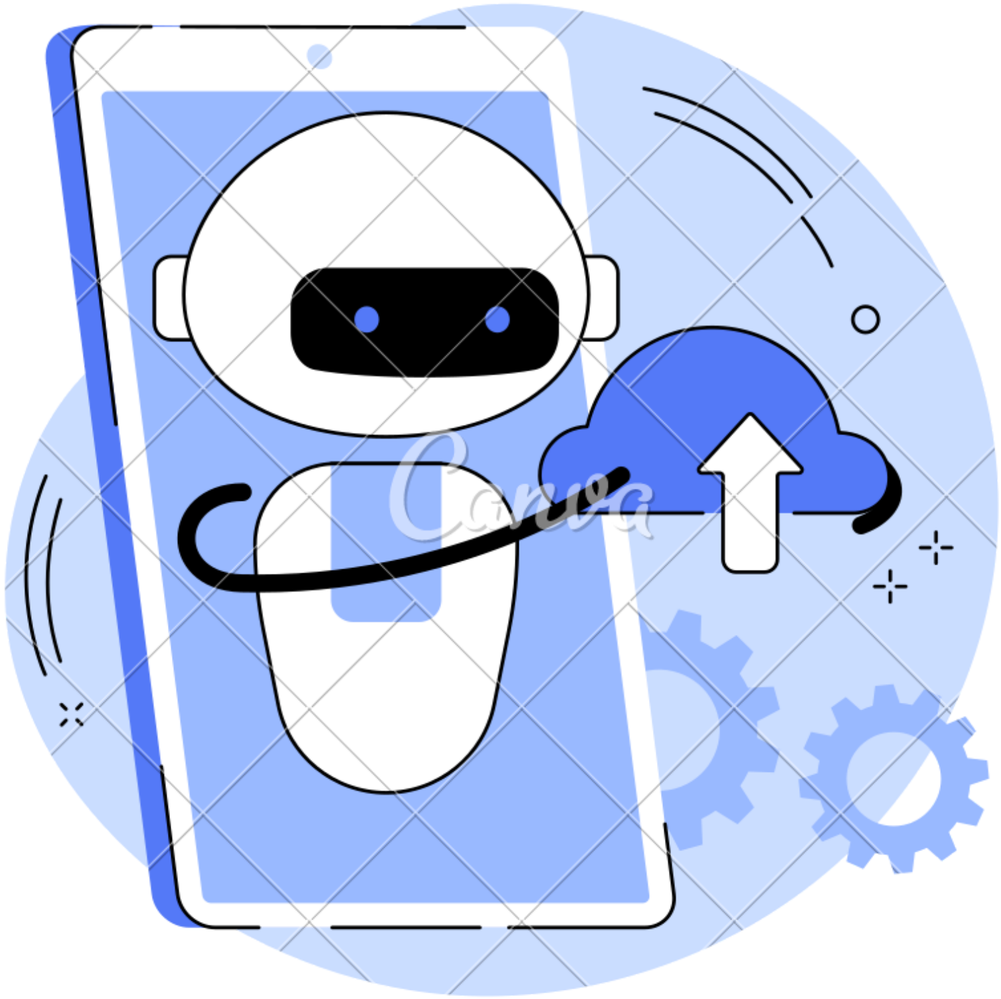
Preprocessing steps like data cleaning, normalization, and feature engineering can inadvertently introduce bias. Biases present in these steps can carry forward into the final AI model.

BIAS IN ARTIFICIAL INTELLIGENCE - CONTEXTUAL BIAS

The phenomenon where AI systems exhibit biased behavior or make incorrect decisions due to their inability to understand the nuanced context of a situation. This bias arises when AI models fail to grasp the subtle meanings, cultural nuances, or specific circumstances that human beings can easily comprehend. As a result, the AI system might generate inaccurate outputs, reinforce stereotypes, or produce inappropriate responses because it lacks a deeper understanding of the context in which it is operating.

BIAS IN ARTIFICIAL INTELLIGENCE - LABELING BIAS

Labeling bias in AI occurs when human annotators introduce their own biases while labeling training data. This bias can lead to skewed learning patterns in AI models, perpetuate stereotypes, underrepresent certain groups, and amplify societal biases. It influences how AI systems make predictions and can result in unfair or inaccurate outcomes, impacting fairness, ethics, and societal implications of AI technologies.



THANK YOU!!!!!!!!

😀

BY - NAVYA & PRESHA