AI ETHICS

Fill In the blanks:-

1. The guiding principles to decide what is good or bad is known as _____. (Ethics)

2. When building AI solutions, we need to ensure that they follow _____(AI Principles).

3. Praneet has taken extra packets of mouth freshener after dinner from a restaurant. Is it considered as theft?" Is it -Moral or Ethical concern? (no, it is not considered theft, it is an ethical concern)

4. Rakshit and Aman are talking about purchasing a new mobile. They discuss various features which

they want in their mobile. Aman finds that, he started getting notification of various models of Mobiles that meets his requirement? Write which ethical concern the above example depicts. (Data Privacy)

5. "Preference for one over the other" is known as _____. (Bias)

6. Artificial Intelligence and machine learning systems can display unfair behaviour if not trained

properly. (True/False) (True)

7. Search for images of personal secretary on Google, displaying predominantly the images of Women is an example of _____. (Bias)

8. An Ethical AI framework makes sure that transparency, fairness and accountability is develop into the systems to provide unbiased results. (True/False) (True)

Answer the following:

1. Differentiate between Ethics and Moral with suitable examples.

Ans:-Morals are nothing but the rules dictated by society, these are not fixed it can be different for different societies

It is a guiding principle to decide which is good or bad, these values chooses by the person for his life.

2. Define principles of AI.

Ans:- The Principles of Artificial Intelligence (AI) Ethics for the Intelligence Community (IC) are intended to guide personnel on whether and how to develop and use AI, to include machine learning, in furtherance of the IC's mission. These Principles supplement the Principles of Professional Ethics for the IC and do not modify or supersede applicable laws, executive orders, or policies. Instead, they articulate the general norms that IC elements should follow in applying those authorities and requirements.

3. Explain Data privacy.

Ans:- Data privacy referred to as information privacy is concerned with the proper handling of sensitive data including personal data and other confidential data, such as certain financial data and intellectual property data, to meet regulatory requirements as well as protecting the confidentiality and immutability of the data.

● Understanding what data, you have collected, how it is handled, and where it is stored.

● Necessary data required for a project should only be collected.

● User consent while data collection must be of utmost importance.

4. Craft a description of how considerations for inclusivity are addressed during the development of AI models.

Ans:-

AI MUST NOT discriminate against a particular group of population, causing them any kind of disadvantage.

Here are a few things you should take care of

▪ Does your AI leave out any person or a group?

▪ Is a rich person and a poor person benefitted equally from your AI?

▪ How easy is it to use your AI?

▪ Who does your AI help?

▪ How can we make AI more inclusive?

5. Write Major Issues around AI Ethics.

Since AI gained so much power that it can change the lives of people, it also has to look into ethical issues around it as AI can have enormous impact on societies and even nations. Some of Major Issues of AI are

- Bias and Fairness
- Accountability
- Transparency
- Safety
- Human AI-interaction
- Trust, Privacy and control
- Cyber security and malicious use
- Automation and impact over jobs
- Democracy, civil right, robot right etc.

6. A company had been working on a secret AI recruiting tool. The machine-learning specialists uncovered a big problem: their new recruiting engine did not like women chefs. The system taught itself that male candidates are preferable. It penalised resumes that included the word "women chef". This led to the failure of the tool.

a. What aspect of AI ethics is illustrated in the given scenario?

b. What could be the possible reasons for the ethical concern identified?

Ans:- The company developed AI recruiting tool preferred men chefs instead of equal preference has to be given both, which leads to AI Bias. AI Bias means inclination or prejudice for or against one person or group, especially in a way considered to be unfair.

a. In the above scenario it is against a group of people who are not preferred for the chef job which leads to AI Bias.

b. The possible reasons for this ethical concern include:

Biased Training Data: If the training data used to develop the AI system contained biases, such as a historical preference for male chefs, the AI would learn and replicate these biases1.

Lack of Diversity in Data: Insufficient representation of women chefs in the training data could lead the AI to undervalue their resumes

Inadequate Oversight: Without proper monitoring and intervention, the AI system might perpetuate and even amplify existing biases

Flawed Algorithm Design: The algorithms might not have been designed to detect and mitigate biases, leading to discriminatory outcomes

7. As Artificially Intelligent machines become more and more powerful, their ability to accomplish

tedious tasks is becoming better. Hence, it is now that AI machines have started replacing humans

in factories. While people see it in a negative way and say AI has the power to bring mass

unemployment and one day, machines would enslave humans, on the other hand, other people

say that machines are meant to ease our lives. If machines over take monotonous and tedious

tasks, humans should upgrade their skills to remain their masters always.

What according to you is a better approach towards this ethical concern? Justify your answer

Ans:- The ethical approach to AI replacing humans in factories should focus on upskilling and reskilling the workforce. While AI can handle monotonous and tedious tasks, humans should be encouraged to develop new skills that complement these technologies. This approach not only mitigates the risk of mass unemployment but also ensures that humans remain central to the workplace, leveraging AI to enhance productivity and innovation.

By investing in education and training programs, we can create a future where humans and AI work together harmoniously, driving economic growth and improving quality of life. This balanced approach addresses ethical concerns while maximizing the benefits of technological advancements.