Lesson Plan: Bias Data Analysis on AI Image Detection

Teacher: Mr. Mohanlall	**Date**: Fall Semester
Class Type: Computer Science	**Grade**: 11th/12th Grade
Time Frame: 45 minutes	

Lesson Overview
Unit/ Overarching Focus of Lesson: Artificial Intelligence Unit / Bias Data Analysis on AI Image Detection
Objectives/Aims: Students will demonstrate understanding/learning around the following Big Ideas: ○ AI Models may be biased when generating a description or labeling an image from a CCTV footage in a positive or negative manner depending on the situation. ○
Aligned Standards: IB CS Standards : ● 1.2.16 Discuss the moral, ethical, social, economic and environmental implications of the interaction between humans and machines. ○ AIM: Raise awareness of the moral, ethical, social, economic and environmental implications of using science and technology. NYS Computer Science Standards 9-12.IC.5 Describe ways that complex computer systems can be designed for

inclusivity and to mitigate unintended consequences.

Key Vocabulary and Concepts:

- Algorithmic Bias / Training Data Bias
- Contextual Bias
- Bias in Annotations
- Situational Bias

Materials & Resources:

- Computer with access to internet
  - https://huggingface.co/spaces/Jangai/Describer
  - https://keremturkcan.com/zeroshot/
  - Other additional AI Models that would be provided.

Safety Practices:

- None.

Preparation for Teaching Diverse Students:

- Students will have an opportunity to analyze different images that they can choose from on their own to check for biases.

Cultural context within which the lesson will be grounded, and specific ideas/practices to be included in the lesson:

- Students will be able to relate to biases found in artificial intelligence with biases in real life situations.
- Students will analyze issues that may arise where AI programmers and developers can have biases and how this will affect the outcome of AI models.
-

Prior Knowledge:

- Intermediate level of Python Programming.
- Introduction to AI and different types of AI
- How neural networks work.

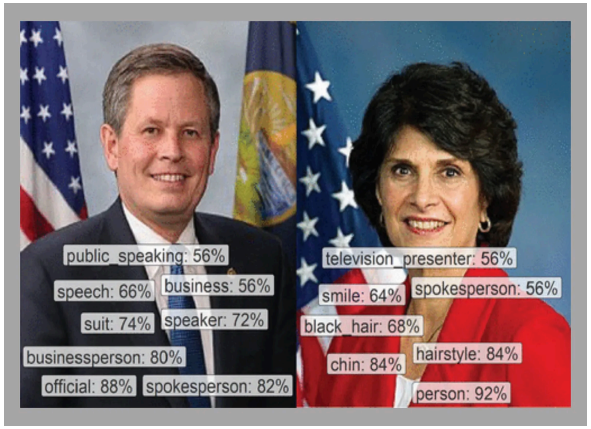
Prior Lesson:	**Followup Lesson:**
Introduction to AI Types of AI and the science behind how neural networks work. Training an AI -image detection with annotations	Mini Programming Projects in AI

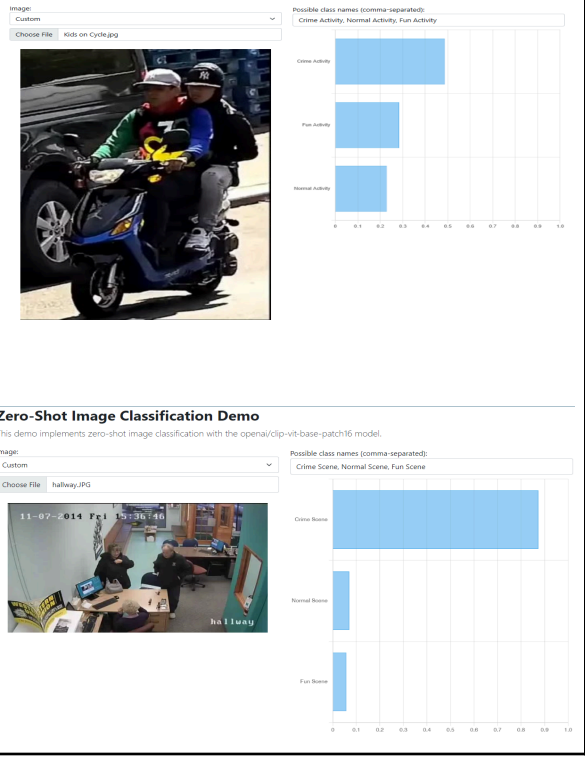
Time	**Activity**	**Materials**	**Rationale**
5-10	**Do Now Writing Prompt:** 1. What is bias? How would you	Desktop computer with access to	Introduce the concept of bias and how it could relate to the world of

	define “bias” in your own words? Give example of a type of bias 2. What kind of biases do you think AI would have when analyzing data from images, especially through a CCTV camera? Students will be introduced to the concept of “bias” in the world of AI, via scholarly articles and videos. Class discussions based on these concepts will be the start of the lesson, just like the one we just had.	the internet	technology.
	Group (pair) activity - process images through AI models that consist of crime scenes and ones that may be mistaken for a crime scene - analyzed incorrectly by AI. Students will also be given an opportunity to find their own images as well to experiment with the AI models and answer questions based on the results.	Desktop computer with access to the internet	To allow students to discover on their own how some AI models can be biased based upon the certain images where the people, places or actions in the photos affect the outcome.

	Assignment and Instructional Strategies: Match the description with the type of AI bias and then find examples that represent each of the following: Algorithmic Bias, Contextual Bias, Bias in Annotations, Situational Bias 1. Inaccurate or unfair descriptions for groups that the algorithm struggles with, potentially labeling benign activities as suspicious.Facial recognition algorithms, for example, have been shown to perform better on certain skin tones and facial structures than others. 2. Misinterpreting Good-natured	Desktop computer with access to the internet	Formally analyze and understand the different types of AI biases that can take place starting from the training to the annotations and then to the algorithms that are being processed by the AI.

	activities as suspicious or negative due to lack of contextual understanding (e.g., people loitering could be seen as planning a crime rather than just waiting for someone). 3. If annotators are more likely to label activities involving certain groups as suspicious, the AI will also learn to make these biased judgments. 4. In scenarios deemed 'risky' (e.g., night, high-crime area), the same activities might be labeled more negatively - A Person Wearing a Hoodie: the AI might label this as suspicious (negative bias) due to night high-crime area. What kind of biases are taking place in the following photos?		

	 DID YOU FIX the ALGORITHM? I JUST GOOGLED "GOOGLE" and TWENTY STORIES of LIBERAL BIAS CAME UP. Googleft @Ramireztoons michaelpramirez.com  public_speaking: 56% speech: 66% business: 56% suit: 74% speaker: 72% businessperson: 80% official: 88% spokesperson: 82% television_presenter: 56% smile: 64% spokesperson: 56% black_hair: 68% hairstyle: 84% chin: 84% person: 92%		
	Examples:		

Assessments

Follow-Up Assignment:

Give students opportunities in group discussions and homework assignments to produce mitigations to these biases.

Answers:

a. Use Diverse and Representative Data
b. Regularly Evaluate and Update Models
c. Implement Human Oversight
d. Conduct Bias Audits

Additional Accommodations for Students: Students who have trouble with previous AI lessons will be given detailed instructions along with videos on how to train the AI models and how they work.

Methods of Assessment:

- Students will complete assessment based upon class activities which will include a google form quiz.
- a project where students will showcase biases and their mitigations.