

## PS 643 Notes (3/09/24)

- No AI can tell you your internal state, like your suitability. It can only detect observable characteristics.
- There are a lot of these companies which claim to predict these using AI which have also gotten a lot of funding.
- Why do they get funding? This is to induce terror in the minds of the other people.
- **Chilling effect:** Enacting any policy not to solve a problem but to induce fear.
- In South Korea, there are companies that teach people how to smile so that the AI doesn't think that it is suspicious
- Prof Narayan: Most of the AI discussion happens on Twitter.
- AI softwares claiming to increase productivity: How does it work?  
You get scared and start working and become productive.
- Whether a technology is productive or not? What is productivity-more output for the same input. But what these AI softwares do is increase the input and hence the output
- Massive effort to influence public opinion
- Though AI has made genuine, rapid progress: What is common between all these softwares that work?  
That you don't ask it for any type of interpretation but only identification, like speech to text, deepfake.
- Softwares that are dicey but have improved:  
Content recommendations: there are ethical concerns, can be used in ways that are unethical as well
- Softwares that don't work: predicting societies, eg: predicting job performance
- Princeton study, 457 scientists, 12942 variables; but then someone found out that even if you use 4 variables, it gives the same results
- COMPAS tool: to study the characteristics of prisoners and predict who needs help with mental health. But they tried to also predict which prisoner is the most dangerous.
- But in case of logistic regression using just 2 features; age and number of pending cases against him/her; that has better accuracy
- Science: Natural Philosophy
- Everything is philosophy
- Basic pillars of philosophy:
  1. Ontology: What is all this?
  2. Epistemology: How do I know if this is right or wrong?
  3. Ethics/Moral Philosophy: What do I need to know?
- Mind and Ethics: Ethics and morality is the same for philosophers.
- How do you decide what is ethical? (Normative ethics)  
How people decide what is right or wrong? This has changed over time.  
**Virtue Ethics:** Does a person have good qualities? Quality of a person. Doesn't look at the actions at all. This is a very old way of doing ethics.  
**Deontological Ethics:** what was the intention behind doing something? Doesn't care about the circumstances of doing something

**Consequentialism:** What the actions impacts society as? First thought of by British (Bentham). But who decides the utility? They said state decides. But people asked who is the state.

Criticized by Germans, saying the state is not representing all the people.

**Pragmatic:** American philosophy of ethics. Get everyone from different classes on the table and decide what is right. Our constitution is heavily influenced by pragmatism.

- What is ethics? (Meta ethics)
  - How do I translate all of this into my actions? (Applied ethics)
- How societies should work should be decided by all the people (democracy) and not by some bunch of rich people: Why effective altruism is a faulty philosophy?
- Whatever is economically just, should that be considered the right decision? Only on the basis of economical equality.