→ Chilling Effect: Enacting policies to hide existence of problem via fear, instead of solving the problem

Relevance: AI companies claim sentiment understanding via visual media ⇒ Bogus, smiles faked in South Korea by employees

Snake Oil ⇒ Exploiting fear to sell goods ⇒ AI snake oil

→ Massive effort to influence public opinion ⇒ Successful

→ AI (Crude differentiation)
- Genuine Progress: Does & should (can be harmful)
- Useful, but not much progress: Works but bad performance
- Fraud: Doesn't, Won't, Harmful

→ Genuine progress: Facial recognition, speech to text, Medical diagnosis from scans, Deepfake (moral dilemma) — Perception

→ Improving: Spam detection, Copyright detection, Automated essay grading, Hate speech detection, Content recommendation — Automated judgement
↓
Can be done but far from perfect rn.

→ Fraud: Predicting. policing, terror risk, risky kids, job performance, criminal recidivism, anything related to human behavior — Predicting social outcome

~~Isolated corpus - Gutenberg~~

→ ~~Members of UNs need to acknowledge Human rights~~

→ 13000 parameters based modeled ≃ 4 parameter based linear regression (in social prediction)

→ Accuracy of recidivism prediction:
137 features: COMPAS tool :- 65% ± 1% (slightly better than random)
~~Logistic regression~~: ≥ features: Logistic Regression :- 67% ± 2%
↑ Age & No. of priors

☆ Narayana claims AI is shit at predicting social outcomes

→ Harms of AI in social prediction :- Hunger for personal data from domain experts & workers (privacy harm)
- Massive power transfer ^ to unaccountable tech company
- Lack of explainability - Veneer of accuracy - Distracts from interventions

~~Takeaways~~ Takeaways:

→ AI shit at social prediction
→ Funding is a cheap way of ~~half assed~~ solving problems
→ Sometimes, Manual > Automation testing

Philosophy → Mind
→ Ethics
↳ what is this

→ Ontology: What is reality composed of?
→ Epistemology: What is truth? Is it real or my mind's playing tricks?
What is knowledge?
How do I know what it is
→ Ethics / Moral Philosophy

Ethics

→ Moral philosophy: What is ethics, is it different from morality?

→ Normative Ethics: (Good/Bad quality)
- Virtue ethics: Quality of a person (not concerned with actions)
- Deontological ethics (German): "Majority of human civil." (Good/Bad action)
  Focuses on the action's nature, action's intention
  - Immanuel Kant (POI) — certain things are bad (no ifs, ands & buts) → Categorical Imperative; Critique of Reasoning
  - John Rawls → Veil of Ignorance → Remove everything, become completely impartial
  - No regard for consequence or cause.
- Consequentialism (British): Focus on the impact
  - Jeremy Bentham → Utilitarianism
    - reaction (kind of) → State consequentialism
    - Contextual (By Karl Marx) (critised ⇒ Good for one, bad for another, state not ~~defined~~ clear here
  - Welfarism

- Pragmatism: American way, representative of everyone, mutual decision (also Indian)
  - John Dewey
    - student → B.R. Ambedkar (Dr.)

Applied Ethics

→ ~~AI control problem~~ → Machine ethic: Instructor believes machines shouldn't even be in discussion with ethics

→ AI control: AI taking harmful decisions for task completion (murder for winning a game)

→ Economic justice cannot come at the cost of social justice

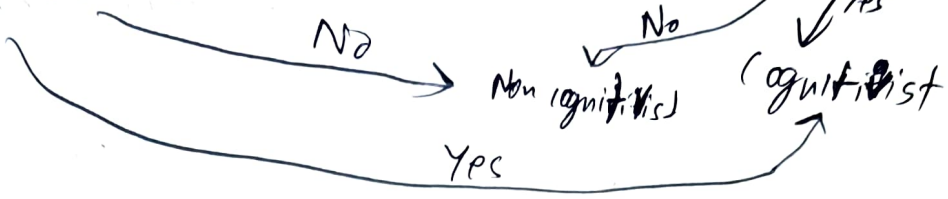
Metaethics

Policy brief: List of demands from policymakers

→ Moral statements: Is it similar to a descriptive statement? → Ambiguous answers
(describing reality)

↓

Is it a truth statement?



Cognitivist: Moral statements are truth-apt

→ Error theory: (Theorist) All moral statements are false (are cognitivist)
(All statements have an aus., it is false)

→ Some are true/false

→ Objective i.e. exist in reality: Moral Realist / Objectivist

→ Only a social construct, subj. only: Moral subjectivist / Relativist

Non cognitivist: Moral statements are not truth-apt

→ Emotivist: Morality is an emotion guided.

→ Prescriptivist: Moral statements are not true/false, they're wills/orders.

→ Quasi-realist: Not real, humanity adopted it as pseudo-real stuff to live life

→ Is-ought gap: (Realized by David Hume) Most moral args start with logical statements (is) and then go to an ought statement (should, ought). where does this description change to prescription? is what is

- Normative statements come somewhere in all arguments, the theory doesn't question its existence, rather it is concerned with how, why & where does this transition occur

Ethics

→ Ethics initiatives
- By trans-national orgs, like UN, IEE

→ Ethics boards
- Companies future proofed themselves to avoid investment in "unethical" activities

→ Lack of actionability
- No one can do anything if "unethical" or non compliance with certain standards, due to ethics being entirely theoretical & hence no legal basis for the compliance of standards by AI companies

→ Project Maven
- Idea by Pentagon after 9/11, to track everything of every person on Earth by satellites (theoretical ambition)
- Pitched to Google, 1st time a strike occurred & employees left (due to ethical concerns), hence Google had to refuse.
- Palantir picked it up

→ Lack of Transparency

Issues with Tech Ethics

→ Individualist Solutions
- Solutions are not preventing the generation of ethical issues, they're targeting individual problems, which will always keep popping up.
  Ex: Project Maven stopped by Google, when it shouldn't have been thought of.

→ Corporate logics
- Company won't let anything impact its bottom line

→ Contradiction & Vagueness
- Ethics suffers from these

→ Ethics washing
- Describe earlier

→ Determinism & Solutionism
- Solutionism
  - Everything ought to be solved
  - Wrong, not only morally, but because it is reductionist/facile as shown by

What is a mind?

→ what changes ~~bet~~ b/w consciousness & unconsciousness? We don't know

→ Chinese Room: Analogy for A.I, ~~NN~~ neural network, computer (originally)

↓

An AI, a computer can't ~~ever~~ ever have a mind.

→ Dualism

- Mind is separate from matter

→ Monism

- No such distinction b/w mind & matter.

→ Idealism: Use mind to think about "matter", hence only ideas & its realm is what is thought about

→ ~~Mate~~ materialism: Reality can be measured and known, not just ideas in mind.

~~Philosophy~~ Philosophy of Mind

→ Can't know other ppl have minds

→ If matter is what makes up the mind & it's governed by physics, so is the mind, and if so, does free will exist?

→ ~~Thinki~~ Can think of things with which there is no material contact.

Materialism

→ Behaviourism: Behaviours has a 1 to 1 connect with the mind, hence behaviours determines mind.

↓

Wrong

→ ~~Physicalism / Identity theory~~ → Type Identity theory: ~~Mente~~ It : Black box

→ Token " "

→ ~~Qualia~~ Subj. conscious experiences → No way to know if everyone has the same or diff. qualia

EAI

→ Explainability ⟷ Transparency. Explainability ; in general & in to an engineer is not the same.

Procedural expl. & Technical expl.

Issues

→ Ignorance due to no outward looking & considering the society

→ Mcnamara's philosophy says can't be measured ⇒ not knowledge

Mcnamara fallacy ⟸ ignore

→ Epistemic contestation: which kind of knowledge should have privilege?

→ Political eco. contestat^{n}: Doesn't count for how tech affects politics & economics