**AI and society**

* AI and hype
* Limitations of AI
  + Bias
  + Adversarial attacks
* AI, developing economies and jobs
* Conclusion

**Goldilocks rule for AI**

* Too optimistic: Sentient/super-intelligent AI killer robots coming soon
* Too pessimistic: AI cannot do everything, so an AI winter is coming
* Just right: AI can’t do everything, but will transform industries

**Limitations of AI**

* Performance Limitations
* Explainability is hard(but sometimes doable)
* Biased AI through biased data
* Adversarial attack on AI

**AI learning unhealthy stereotypes**

* **Why bias matters**
  + Hiring tool that discriminated against women
  + Facial recognition working better for light-skinned than dark-skinned individuals
  + Bank loan approvals
  + Toxic effect of reinforcing unhealthy stereotypes
* **Combating bias**
  + Technical solutions
    - E.g., “zero out” the bias in words
    - Useless biased and/or more inclusive data
  + Transparency and/or auditing processes
  + Diverse workforce
    - Create less biased applications

**Adversarial attacks on AI:**

* Minor perturbation.
* Physical attack.

**Adversarial defences:**

* Defences do exist but incur some cost.
* Similar to spam vs anti-spam, we may be in an arms race for some applications.

**Adverse uses of AI:**

* DeepFakes
  + Synthesize video of people doing things they never did
* Undermining of democracy and privacy
  + Oppressive surveillance
* Generating fake comments
* Spam vs anti-spam and fraud vs anti-fraud

**Developing Economies:**

* Leapfrog

**How developing economies can build AI:**

* US and China are leading, but all AI communities are still immature.
* Focus on AI to strengthen a country’s vertical industries.
* Public-private partnerships to accelerate development.
* Invest in education.

**AI’s impact on jobs worldwide:**

* 400-800 million jobs are displaced by 2030.
* 555-890 million jobs are created by 2030.

**Some solutions**

* Conditional basic income provide a safety net but incentive learning
* Lifelong learning
* Political solutions