1. What are Deepfakes?

- Application of data science that involves generating images and videos using machine learning and neural networks

- Data science is the field I want to work in, and I find the ‘Generative Models’ section to be especially interesting

2. Reconstructing Faces

- For this presentation, focusing on videos of human faces generated by artificial intelligence models

- Premise is taking one person’s speech and making another person say it

- Figure on left shows how this is done: the input faces are compressed by artificial intelligence models and then reconstructed with another person’s speech

- In GIF on right, woman’s speech and movements are overlayed onto man’s face

3. Misinformation in Deepfaked Videos

- GIF on left shows a subject creating a deepfake of Vladimir Putin

- The ability to make our world leaders say anything can have very negative consequences

- Could undermine the truth very easily on the internet

- News outlets are very scared of the threat deepfakes pose, as shown by headlines on right

4. Combating Deepfakes with Algorithms

- Microsoft, Facebook, and others hosted a challenge for creating the best deepfake detection algorithm

- Detection algorithms are also artificial intelligence models

- They separate the faces into small features and use those to figure out if the face is real or not

- Winning algorithm was only 60% accuracy

5. Spotting Deepfakes

- Luckily, humans have very good intuition when it comes to recognizing faces

- GIF is of Jordan Peele’s speech being reconstructed onto Barack Obama’s face

- There’s a lot of stuff wrong with Obama’s face that you can see somewhat easily

- Deepfakes don’t pose much of an issue as long as we are willing to stop and think about the validity of what we see online

6. Conclusion

- All depends on our digital literacy, especially critical thinking and evaluation

- Only as much of a threat as we allow them to be