Name: Ayodeji Abdulfatai Oyinloye
Course: Machine Learning with Python
Exercise 1.2: Ethics and Direction of Machine Learning Programmes
Task 1.2: Issue of Ethics in Weather Machine Learning Studies
Date: 13.04.2025

Part 1

1. Read or watch at least one of the examples of machine learning ethics below. Each example looks at how machine learning was used to address different technical or societal issues.
    - What are Ethics and Why Do They Matter? Machine Learning Edition;
    - Machine Learning Moderating Social Media: The Growing Role Of AI In Content Moderation (pro) and The Darker Side of Machine Learning (against);
    - Machine Learning and Human Bias.

Answers

I. Ethics and why they Matter

This video above pictured one Dr. Rachel Thomas of the University of San Francisco, and director of the university's Centre for Applied Data Ethics and fast.ai

The 12 minutes video teaches a wide range of tech-based ethics issues from olden days (allegation of IBM collaborating and providing tech support for the Nazi Government of Hitler and the killings of the early 1930s) to modern ones involving Facebook, YouTube, Goggle, other faulty machine learning applications that somehow did not enrich themselves with feedback mechanism to collect and attend to possible post-roll out errors, etc. An example of lack of avenue for collection and prompt attendance to post implementation error is seen in the reported case of an algorithm-based deployment cutting the health benefits of deserving members of the society in the US. This is even more pathetic when it involves individuals who have physical, mental, emotional or other challenges.

Another case depicting possible glitches that can happen during or after deployment of machine learning is the reported issue of Amazon's face recognition model "matching members of congress to criminal mugshots"

I will leave below a trail of key content views from her brief video lecture in images.

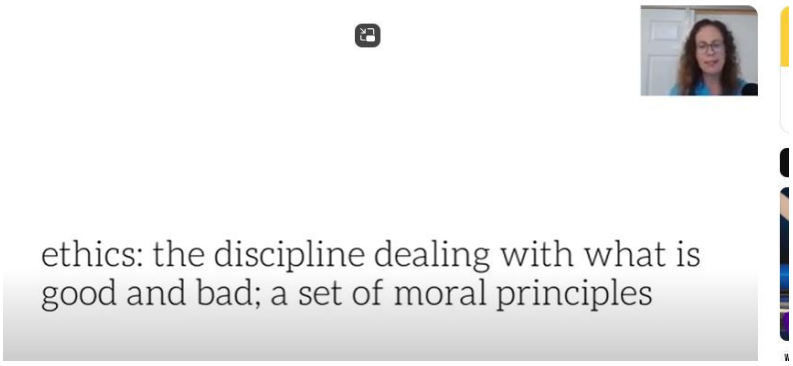
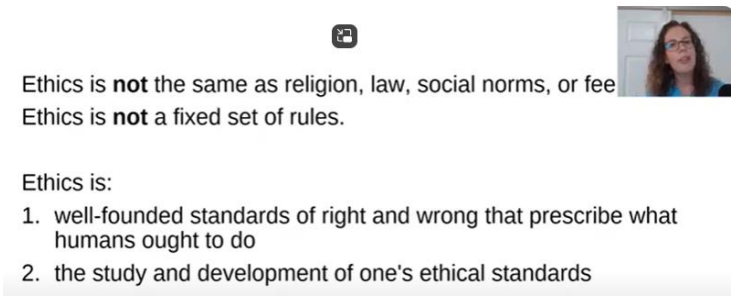
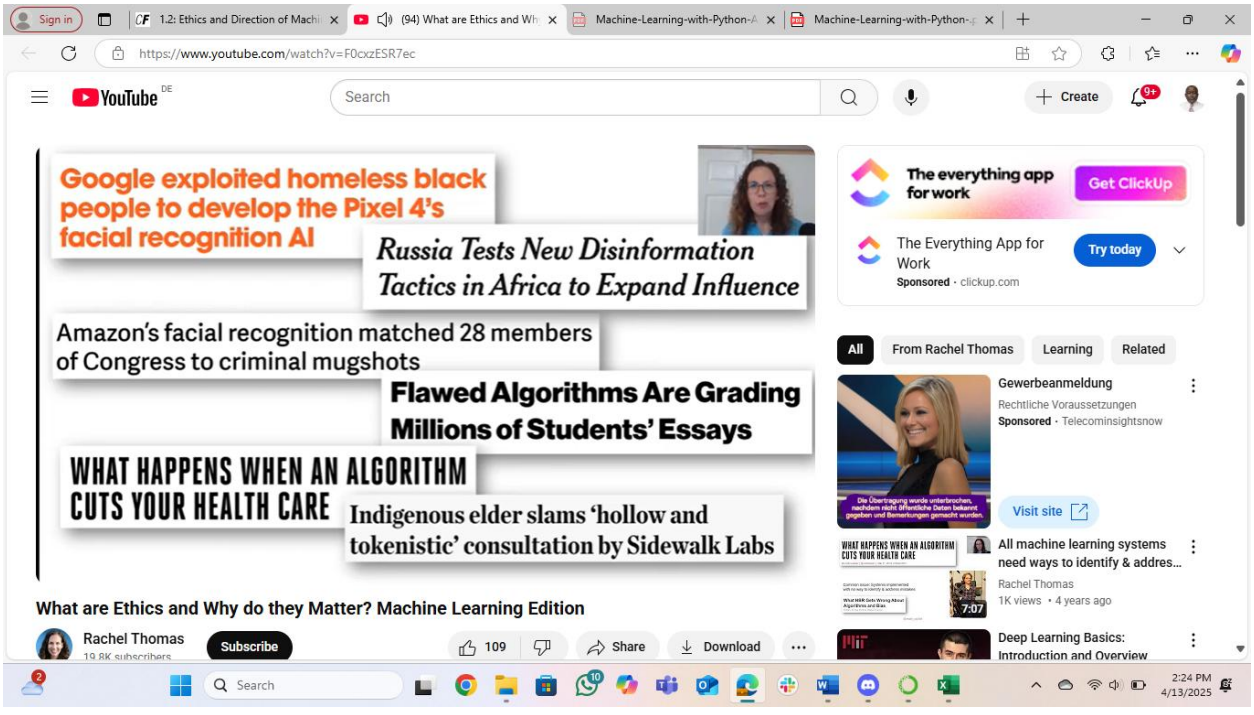
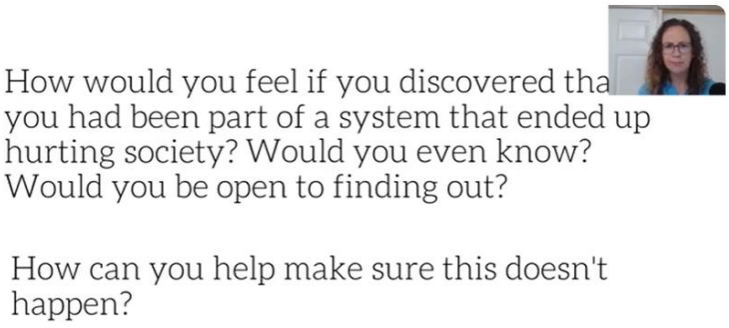
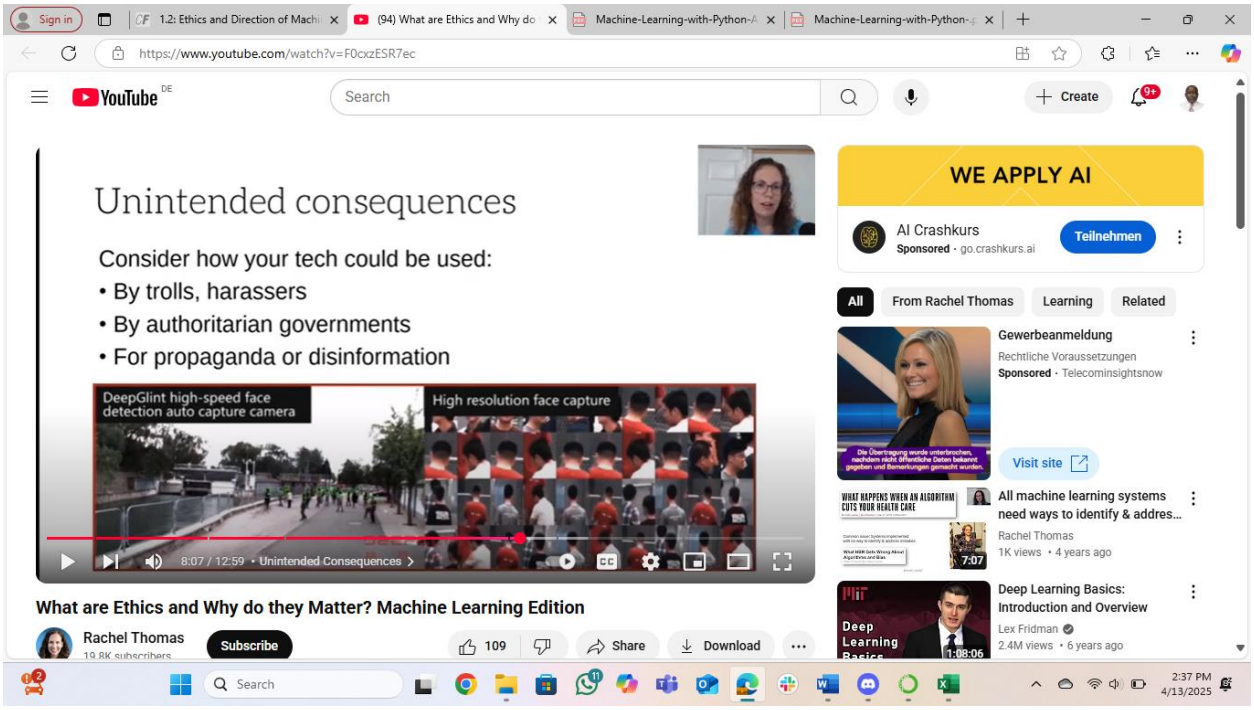


Figure 1: Dr. Rachel Thomas, defining ethics









II. The paired resource links for AI systems in social media

The first resource discussed how big social media has grown as at 2022 (4.62 bn people, 10% growth over the 2021 figure) and how huge the derivable market from it was growing as well (USD 3 bn in 2020, at compound annual growth rate, CAGR, of 27.1%, to reach USD 20 bn by 2028). The author described the huge volume of daily expected user-generated content by 2025 (463 exabytes or > 200 m DVDs). Even when we categorize this huge volume by providers and platforms, each will still be huge for human moderators to humanly handle. Thus, AI systems could be deployed to scan, filter and categorize various online contents with speed, scale and efficiency. The filtered contents can be flagged for human final overview before they are determined suitable or not for their cases. This is a positive deployment of machine learning.

The second link discussed the harmful aspects of machine learning deployment. A grave example is seen in threat to personal security and correct use of information – criminal identity theft possibility which may mean we can't keep our identities or that what we did not say or do may be attributable to us. This can happen if criminals deploy machine learning with historical data about a victim to generate untrue new misinformation about the person, thereby mimicking the victim successfully. In one example is the reported case of the unrestricted access a face recognition app, *FindFace*, has to over 200 m social media data of subscribers of *VK.com* (both are Russian tech companies). In another example, researchers at the University of Texas were said to have successfully used a regularly available and popular ML procedure to train algorithms to achieve 90% undermine of the privacy benefits of content blurring and obfuscation protocols – the researchers had the objective of warning tech providers of such protocols. These are instances of dangers of negative deployment of machine learning.

2. What potential is there for bias or ethical issues when dealing with climate change data? Where would ClimateWins need to be cautious about using machine learning to develop answers? Consider the following questions and write an answer of around 200 words showing what pitfalls ClimateWins should avoid:

- Is there personal information that may be exposed?
- Are there regional or cultural biases in climate change that might be made worse by using machine learning?
- Is there human bias in climate change that might be propagated while training machine learning?
- Could machine learning potentially make incorrect decisions about where weather conditions might worsen and cause harm?

Answers

I. **Exposure of Personal Information**

this can happen if and when location specific weather information is fed into learning models. Feeding people-specific information can deepen the risk here. Compounding this risk in machine learning may be averted by data anonymization and observance of due diligence to remove unnecessary geo-specificity.

II. Regional and Cultural Biases

The eighteen (18) weather stations across Europe may have, overtime, generated data that is too locality specific and thereby not necessarily representative of other possible geographic units within the continent. The weather data collection methods of the 1800s may also meant exclusion bias pattern such that localities with less sophistication were poorly reported. If such exist in data, we may need to use statistical methods and weighting to remove such bias risk which is an ethical issue.

III. Human Bias in Existing Data

The period of time between the late 1800s and present days represents a huge layers of measurement technologies and instrumentations history. This may mean earlier instruments had much less accuracy and precision compared to modern ones, which is a measurement bias and an ethical issue. We may restrict historical reference to eras of comparable accuracy or calibrate data for machine learning to avoid this ethical issue.

IV. Potential for Erroneous or Misleading Conclusions

A strong risk here is the fact that weather data dating long back in history are devoid of the reality of present-day impact of global warming (of course, data completeness is also key). Using such data without necessary compensation or factorization may lead to underestimated forecast of extreme weather events that may be great risks for people and property, meaning associated weather advisory would be unreliable and unethical. Machine learning by *ClimateWins* has to deliberately mitigate this.