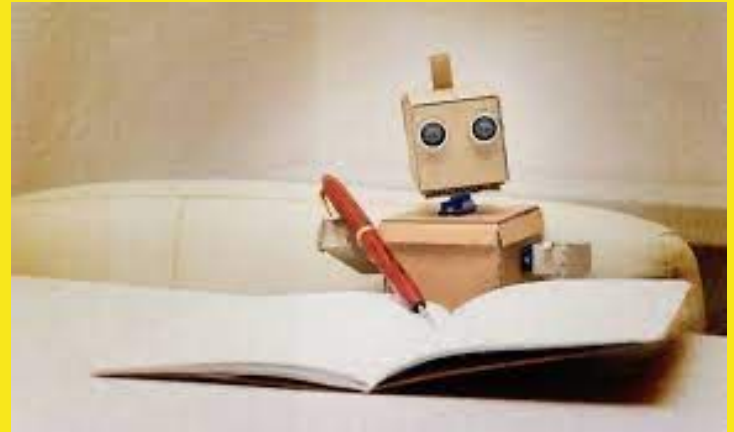


The Ethics of AI and Grading

Ian Scheffler + Michele Persaud

The Issue

Should AI be used in grading?



**Which piece of feedback was
written by a human?**

Piece of Feedback

#1

Problem: Get user input using python.

Close. There is a minor error with your logic to get input from user.

This could be something like forgetting to convert user input to a float.

Piece of Feedback

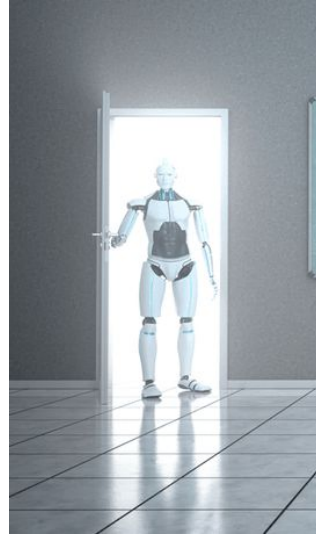
#2

Problem: Draw waves of tiny blue diamonds across a black-and-white grid.

Seems like you have a small mistake. Perhaps you are running into a wall after drawing the third wave.

Cons -

- Unfair Outcomes
 - The IBO used AI to predict final student grades (~175,000 students)
 - Used historical results
 - 15% protested (suspiciously low scores; revoked college acceptances)
 - UK used AI to predict A-levels (college entrance exams)
 - ~40% downgraded; 2% increased;
 - students with high grades from less-advantaged schools more likely to be downgraded; students from richer schools were more likely to have scores raised
 - Cheating or Inaccurate grades
 - Students (or parents) figure out the algorithm (scanning keywords)
 - Wide grading range
 - Edgenuity (0% no keywords; 100% one keyword)
- May replace teachers
- Can increase inequity (high income-AI and human teachers; others-device)



Cons con't

- AI only works with very specific tasks
 - Unable to evaluate things it has never seen before
- AI field is not diverse
 - Information AI is fed may be biased or skewed
 - Creates one view of the world
- Does not address all learners
 - Nuances in answers eg: ELL/SwD (cognates)

TENURE-TRACK FACULTY at CS DEPARTMENTS of
TOP UNIVERSITIES around the WORLD by GENDER,
AY 2019-20

Source: AI Index, 2020 | Chart: 2021 AI Index Report

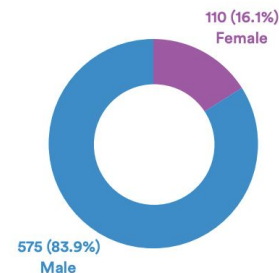


Figure 6.1.2

NEW U.S. RESIDENT AI PHDS (% of TOTAL) by RACE/ETHNICITY, 2019

Source: CRA Taulbee Survey, 2020 | Chart: 2021 AI Index Report

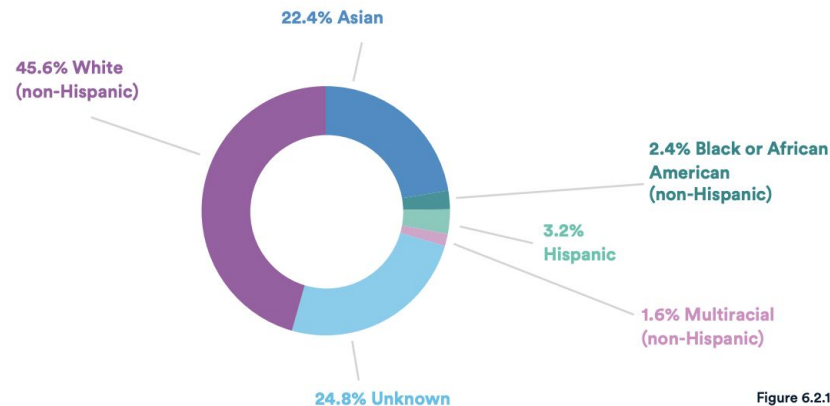
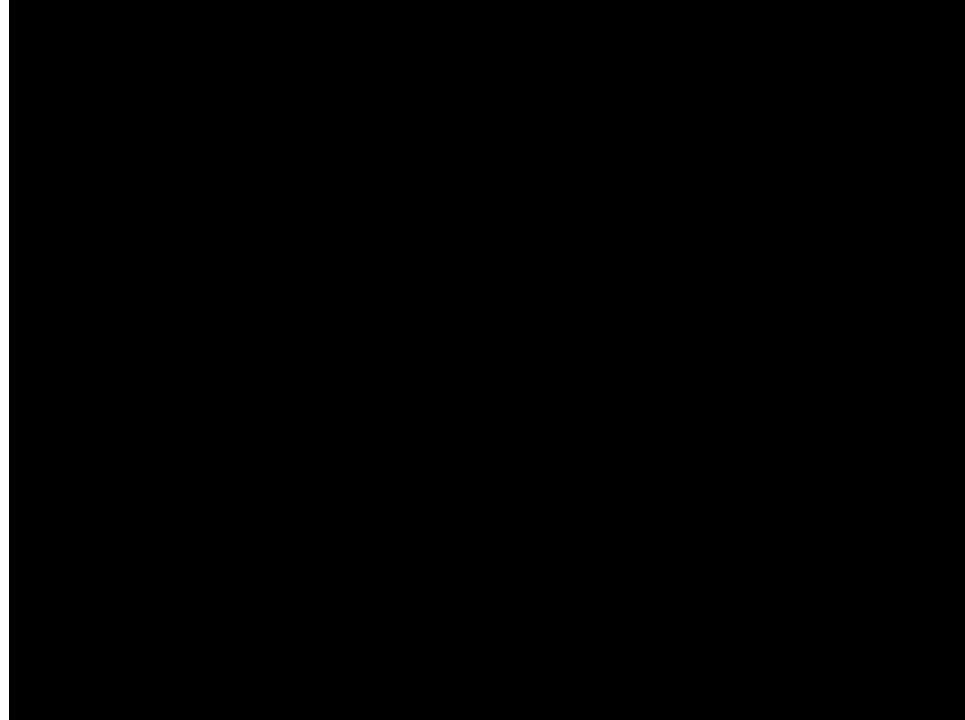


Figure 6.2.1

Pro #1 - Saves Time!

- In 2020, the [University of Massachusetts at Lowell](#), rolled out [Gradeoscope](#)
- Instructors scan student work into PDFs
- [Gradeoscope's AI](#) groups similar answers, allowing professors to grade more quickly
- The AI applies a rubric--created by instructors--to each scanned answer



Pro #2 - (Can) Expand Access to High-Quality Feedback

- In 2021, a lab at Stanford created ProtoTransformer, an AI that can learn how to give feedback
- Used in a Stanford MOOC to give feedback to 16k+ students
- The AI was trained on past midterms graded by human TAs
- Students agreed with the AI's feedback 97.6% of the time
- (Students in the course agreed with human feedback 96.7% of the time)

The screenshot shows a web browser window titled "Code in Place Feedback" with the URL "codeinplace.stanford.edu/diagnostic/feedback". The interface has tabs for "Overview", "Question 1", "Question 2", "Question 3", "Question 4", "Question 5", and "Wrap-Up". The "Question 1" tab is active, showing a "Feedback" section with "Back" and "Next" buttons. The feedback text reads: "GETTING INPUT FROM USER. This question requires you to get input from the user, convert it to a number, and save it as a variable. Did you correctly do all of these steps? Close. There is a minor error with your logic to get input from user. This could be something like forgetting to convert user input to a float. Do you agree with the feedback in the purple box? Please explain (optional):". Below the feedback text are thumbs up and thumbs down icons. To the right, the "Your Solution" section shows Python code:

```
def main():  
    # TODO write your solution here  
    height=input("Enter your height in meters: ")  
    if height < 1.6:  
        print("Below minimum astronaut height")  
    if height > 1.9:  
        print("Above maximum astronaut height")  
    if height >= 1.6 and height <= 1.9:  
        print("Correct height to be an astronaut")  
  
if __name__ == "__main__":  
    main()
```

 Blue arrows point from the feedback text to the code, highlighting the error. Another blue arrow points from the "Please explain (optional):" text to the feedback text. A third blue arrow points from the "Syntax error (missing \" here would prevent auto graders from being useful." text to the code.

AI generated feedback

Students evaluate the feedback

Algorithm uses attention to highlight where in the code the error comes from

Syntax error (missing " here would prevent auto graders from being useful.

**Would you use AI to grade?
Why or why not?**

Sources

- Pros

- <https://www.nytimes.com/2021/07/20/technology/ai-education-neural-networks.html>
- <https://arxiv.org/pdf/2107.14035.pdf>
- <https://ai.stanford.edu/blog/prototransformer/>
- <https://www.uml.edu/news/stories/2020/gradescope-software.aspx>
- <https://www.gradescope.com>

- Cons:

- <https://www.axios.com/england-exams-algorithm-grading-4f728465-a3bf-476b-9127-9dfo36525c22.html>
- <https://hbr.org/2020/08/what-happens-when-ai-is-used-to-set-grades>
- <https://www.edweek.org/technology/teachers-the-robots-are-coming-but-thats-not-a-bad-thing/2020/01>
- <https://www.wired.com/story/algorithm-set-students-grades-altered-futures/>

Further Reading

- <https://www.vox.com/the-goods/2019/4/15/18311367/essaybot-ai-homework-passing>
- <https://towardsdatascience.com/getting-a-machine-to-do-my-english-homework-for-me-5d339470fe42>
- <https://www.economist.com/open-future/2019/10/01/how-to-respond-to-climate-change-if-you-are-an-algorithm>
- https://www.sciencedirect.com/science/article/pii/S0748575109000074?casa_token=foyZ6Rxn06IAAAA:IDb2l4FpoqdsBDeTA3tYegdFGO77EbdfoxXZi2hoq7ecS4hiYPhXr8ZRDt0ww69_2cnxIDvh_mZwTw
- <https://www.proquest.com/openview/4fa03dae1df0915508fcf849b696f97/1?pq-origsite=gscholar&cbl=2031153>
- <https://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS13/paper/view/5938/6134>
- <https://news.google.com/articles/CAliEDY-N-Nh-2aBVkboiaIMrWgqGOgEKhAIACoHCAowsvvyCjDh-9kCMK3gqwM?hl=en-US&gl=US&ceid=US%3Aen>
- <https://www.americanprogress.org/issues/education-k-12/reports/2021/09/16/502663/future-testing-education-artificial-intelligence/>
- <https://www.govtech.com/education/k-12/nualang-introduces-ai-to-help-students-learn-new-languages>
- <https://www.theverge.com/2020/9/2/21419012/edgenuity-online-class-ai-grading-keyword-mashing-students-school-cheating-algorithm-glitch>
- <https://hechingerreport.org/from-admissions-to-teaching-to-grading-ai-is-infiltrating-higher-education/>
- <https://www.edsurge.com/news/2021-01-11-how-professors-can-use-ai-to-improve-their-teaching-in-real-time>
- <https://www.nytimes.com/2019/12/18/education/artificial-intelligence-tutors-teachers.html>
- <https://venturebeat.com/2018/05/28/chinese-schools-are-testing-ai-that-grades-papers-almost-as-well-as-teachers/>
- <https://www.analyticsinsight.net/will-artificial-intelligence-replace-teachers-in-the-future/>
- <https://hechingerreport.org/researchers-say-ai-will-greatly-impact-the-future-of-education/>
- <https://www.wsj.com/articles/when-the-bot-that-grades-your-kids-schoolwork-earns-an-f-11605618000>
- <https://www.iol.co.za/business-report/opinion/tech-news-artificial-intelligence-teachers-making-their-way-into-the-classroom-7a2b81f2-7fd8-4169-9aef-7721246be992>
- <https://www.zdnet.com/article/ai-can-write-a-passing-college-paper-in-20-minutes/>
- <https://www.wired.com/story/algorithm-set-students-grades-altered-futures/>
- <https://www.brookings.edu/blog/the-avenue/2019/09/26/ai-is-coming-to-schools-and-if-were-not-careful-so-will-its-biases/>
- <https://www.nytimes.com/2021/03/17/technology/learning-apps-students.html>
- <https://www.iotforall.com/the-place-for-artificial-intelligence-in-education>
- <https://readwrite.com/2020/11/16/the-future-of-education-can-ai-make-us-smarter/>
- <https://thenextweb.com/news/gpt3-ai-college-essay-grades-compared-students>
- <https://hechingerreport.org/ai-can-disrupt-racial-inequity-in-schools-or-make-it-much-worse/>
- <https://futurism.com/grad-student-neural-network-write-papers>
- <https://news.mit.edu/2013/automatically-grading-programming-homework-0603>
- <https://dl.acm.org/doi/10.1145/3234825.3234839>