Machine Learning in Production
Practice Midterm 2 Questions, Fall 2024

Christian Kaestner and Sherry Wu

Name: _____

Andrew ID: _____

Instructions:

- Including this cover sheet and the scenario, your exam should have _ pages. Make sure you are not missing any pages. *You may detach the last page and recycle it after the exam.*
- All questions in this midterm refer to the scenario on Page __. Answers are graded in the context of the scenario; **generic answers that do not relate to the scenario will not receive full credit.**
- The exam has a maximum score of _ points. The point value of each problem is indicated. We designed the exam anticipating approximately one minute per point.
- **Please write legibly**. We are unlikely to be able to grade your solution if we can't read it.
- We give an amount of space commensurate with what we expect you to need for each question. We use horizontal lines to suggest where to not use the full page. You may exceed those limits if it is clear where to find the rest of your answer. However, we strongly recommend writing concise, careful answers; short and specific is much better than long, vague, or rambling. However, **do NOT write anything you want us to grade on the back of pages.** We will scan the exam and will not look at the back sides.
- This is a **closed book exam**; no books or electronics allowed. You may refer to 6 sheets of notes (handwritten or typed, both sides).

Question 1: Scaling and Operations

All questions in this exam relate to the scenario on the last page. You may detach the last page if you like.

- In this scenario, the training data comfortably fits on a single hard drive. What are potential benefits to (a) partitioning and (b) replicating the data regardless?
- The model is currently deployed as a service that responds with analysis results immediately when a ___ sends a request. You are considering whether to switch to batch processing, stream processing, or the lambda architecture. Within the realism of the scenario, please recommend one approach, and justify your choice (your justification must demonstrate an understanding of the tradeoffs involved):
- A colleague working on ___ is skeptical of the current MLOps hype and thinks we can just write custom scripts to deploy models from a notebook. What arguments would you make for adopting an MLOps mindset and what kind of MLOps tools would you recommend to adopt in the context of the scenario? Justify your answer. Your answer should demonstrate an understanding of MLOps concepts or principles.
- A software engineer in your team argues that Computational Notebooks are terrible tools because they encourage data scientists to write flat code without abstractions, with lots of global variables, and without tests; they argue you should all adopt proper IDEs instead and write standard and testable Python code. What plausible arguments can you make in support of data scientists using notebooks for their work on ___?
- You plan to measure the number of predictions of your machine both for the last 5 minutes and for the last 7 days. Describe briefly how you would do this with Prometheus. The answer must explain which and how many metrics (counters, gauges, summaries, or histograms) you would create for this.
- In the scenario, most of your data is ___. What versioning strategy would you recommend for this data and why?
- Your colleague skipped ___ because they didn't know about how useful it might be. In this scenario, would you consider this as prudent technical debt? Explain why in a way that illustrates an understanding of technical debt and prudent technical debt.

Question 2: Safety and Security

- For the scenario, explain how an adversary could try to attack the system with an evasion attack using *adversarial examples*:
    - Attacker goal (intended outcomes):
    - Which security property does the attack undermine:
      ☐ Confidentiality ☐ Integrity ☐ Availability
    - Attack method (what the attacker would do):
    - Mitigation strategy (how the attack could be prevented/made harder):
- In the scenario, you are worried about the confidentiality of ____ data. Describe (a) one possible attack against the non-ML parts of the system that could break a confidentiality requirement and (b) one attack against the ML model in the system that could break a confidentiality requirement
- Describe one *integrity-related* security requirement for the system in the scenario.
- Robustness alone is not sufficient to ensure safety. Explain with an example relevant to the scenario how a more robust model may lead to worse safety outcomes

Question 3: Fairness

- In the scenario, the project developer is invested in an equity-based approach to fairness with regard to gender. What fairness measure would be suitable in the context of the scenario and how could it be measured in practice?
- You have evaluated that the model is fair according to group fairness (almost perfect parity between groups). Yet customers still complain about unfair allocations. What is a plausible reason that the model may still be unfair?
- Fairness testing revealed that your model heavily relies on the ___ attribute, which is problematic because it is a legally protected category. You pursue anti-classification for the attribute. How can you assure anti-classification in the system *without* changing the model?
- Consider the source of the bias that could lead to ___ unfairness. Select two of the following sources of biases discussed in class – *tainted labels; historical bias; skewed sample; limited features; proxies* – that may be responsible in this scenario. For each, briefly explain how it might have led to the different model behaviors regarding ___.

Question 4: Explainability and Transparency

- In the scenario, suggest two plausible transparency mechanisms that can help end users better interact with the system beyond traditional technical explainability approaches for developers.
- In the scenario, what would be a plausible purpose of providing explanations to ___.
- A colleague argues that counterfactual explanations are intuitive because they align with how humans typically provide explanations to other humans?
    - Sketch a possible counterfactual explanation the system could provide (you can assume any model internals as needed).
    - Would you agree with your colleague? Justify your answer.
- Name a technique to explain which features a model generally relies on the most that could be used to provide high-level documentation for an external audit and explain briefly how the technique works (roughly).