# 11-695/17-645 Midterm, Spring 2021

Christian Kaestner and Eunsuk Kang

**Name:**

**Andrew ID:**

Instructions:

* Fill in answers in this document in between questions (as you would do on paper) or write all answers in a separate document. Ideally, start each question on a new page. Upload the solution as a PDF to Gradescope, mapping questions to pages as in homework assignments.
* All questions in this midterm refer to the scenario on the first 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 **56** points. The point value of each problem is indicated. We allocated approximately one point per minute.
* 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 and remove those lines. However, we strongly recommend writing concise, careful answers; short and specific is much better than long, vague, or rambling.
* This is an open book exam. You may use notes, books, bots, GPT-3, and the internet, but do not interact with other humans.

## 

## Scenario

Delivery robots are expected to be a massive market. Robots navigating sidewalks are considered by many a better alternative to flying drones for safety and capacity reasons. The Pennsylvania legislature has recently classified such robots as pedestrians for the purpose of the law (SB 1199), imposing almost no restrictions on their deployment and allowing weights up to 550 pounds (empty) and speeds of up to 12 miles per hour on sidewalks and 25 miles per hour on roadways. For example, within Pittsburgh, a company called Starship Robots has deployed robots (shown in the picture) on the University of Pittsburgh campus for over a year now that delivers orders from restaurants on campus.

With the increasing demand of online grocery orders that are expected to persist beyond the pandemic, you are part of a competing startup working with grocery chains like Trader Joes and Aldi to produce similar delivery robots. However, in contrast to the small Starship robots for delivering meals from restaurants, your robots have a much bigger carrying capacity, about the size of a full shopping cart, making use of the 550 pound limit. The robot is battery powered (for up to 3h of operation before needing to recharge), has a detailed map of the sidewalks in the delivery range of the supermarket, has cameras in all directions and GPS sensors, but no lidar sensor. The robots have a touch-screen display to interact with humans. Hardware costs are about $3000 each and you plan to lease the entire system to grocery chains for a monthly fee. The robots have a mobile data connection to call for assistance and can be remotely operated by a human operator when needed. A typical supermarket would have one human operator overseeing 10 to 20 robots. One support technician of your startup is responsible for about 20-30 supermarkets.

Your startup works with the grocery chains with a leasing system, where your startup provides the hardware system, making repairs when needed, while the supermarkets do day to day operations with software you startup provides. Your startup receives monthly leasing payments from the supermarket independent of how much the robots are used.

You are overseeing the software team that handles navigation on sidewalks. Two data scientists on your team are currently heavily focused on developing an obstacle detection component that uses camera images to detect obstacles on the sidewalk. In particular, the robot should be able to detect various physical obstacles (e.g., pedestrians, pets, trash cans) and safely navigate around them or slow down if its pathway is blocked.Since obstacle detection is a fairly well studied problem, they plan to train deep neural network models with architectures found in academic papers. They are mostly using public data available from research competitions, but have also been collecting their own data from manually operating the robots in some test areas.

## Question 1: Goals [10 points]

For the project of the scenario, your startup’s organizational goal is to make money from leasing contracts, providing these robots and a corresponding software system to grocery chains. Identify a corresponding leading indicator and a user outcome measure that are reasonable in the context of the scenario. For each describe what data you would collect and what measure you would operationalize from that data (described with enough detail to be independently measured):

(a) **Leading indicator**

Goal:

Data to collect:

Measure and operationalization:

(b) **User outcome (with user = grocery chain that is leasing the system)**

Goal:

Data to collect:

Measure and operationalization:

## 

## Question 2: Telemetry design [10 points]

Collecting data in production is valuable to learn how well the system is doing and to train future versions of the obstacle detection model, but the multiple cameras of the robot produce a massive amount of data, about 600MB per minute, or 1TB per day. Constantly streaming all that data over a mobile data connection seems prohibitively expensive.

(a) [6p] Discuss what telemetry data you would collect and how. Describe the data sources (e.g., video, predictions, human interactions, …). If you decide to sample and store only part of the data explain how you would sample. If computational costs are involved (e.g., for compression or sampling) discuss when and where you perform those activities.

(b) [4p] Briefly discuss the tradeoff between cost (storage, bandwidth, computation, battery) and quantity/quality of data collected and justify why your design is suitable for the scenario.

## Question 3: Trade Offs [12 points]

For obstacle detection, you probably will need to use some sort of deep neural network, but it is still worth considering constraints on possible solutions or identifying important qualities for the entire system and its components, in case local decisions influence other parts of the system design (e.g., hardware requirements). Due to bandwidth costs and latency requirements, your team decides it is necessary to deploy the obstacle detection model directly on the robot.

For each of the three following qualities, briefly discuss how important the quality is for decisions in the system and how it may affect the design of the obstacle detection model and the rest of the system (software or hardware). Make sure your answer is grounded in the realism of the scenario.

(a) Required time and effort to deploy a new version of the obstacle detection model on all robots

(b) Inference latency and energy consumption of the obstacle detection model when the robot is driving

(c) Explainability of the obstacle detection model (to what degree humans can understand and debug why the model predicts that it has detected an obstacle).

## 

## Question 4: Model and Data Quality [12 points]

(a) [4 points] Accuracy results, when you evaluate your obstacle detection model offline, seem too good to be true. You suspect that you ignored the *independence assumption* when splitting training and test data. Briefly given an example of this kind of problem that could have happened in the scenario.

(b) [4 points] Even though accuracy numbers seem high for the obstacle detection model, you are concerned that your model does not perform equally well in all conditions. You consider to curate multiple test sets to cover important scenarios. Characterize at least three slices of test data you would want to consider and monitor separately and briefly describe why you think these may deserve special attention.

* Test sets:
* Brief justification:

(d) [4 points] In the context of the obstacle detection scenario, give an example of data drift (not concept drift) that may occur and degrade the performance of the obstacle detection model over time, and briefly describe how you would address it.

## Question 5: Requirements Decomposition [6 points]

When the delivery robot encounters an intersection during its operation, it must cross it safely by waiting for the traffic light to turn green. The robot must also consider vehicles on the road and other pedestrians in the intersection. You may assume that the robot software contains additional ML models for traffic light detection and tracking pedestrians and cars.

State (i) one safety requirement related to the traffic crossing task and (ii) one environmental assumption and one software specification *that are needed to establish the requirement*. Make sure you decompose the three parts correctly according to the world vs machine distinctions.

Safety requirement:

Environmental assumption (needed to establish requirement with spec.):

Software specification (needed to establish requirement with assumption):

## Question 6: Mitigating Mistakes [6 points]

Consider the following safety requirement for the delivery robot: *If its pathway is blocked by an obstacle, the robot must come to a complete stop in time to avoid a collision.*

A fault tree below shows possible ways in which this requirement may be violated. Describe a mitigation strategy for (i) a fault in software (e.g., ML model) and (ii) a fault in hardware or mechanical components (e.g., sensor or motor). Your answer must describe how each strategy reduces the likelihood of the robot failing to come to a stop in time. You do not need to update the diagram.

Failure to come to a stop in time

OR

Failure to recognize an obstacle in time

OR

Model fails to detect an obstacle in time

Inaccurate data received through camera

OR

OR

(i) Mitigation for a fault in software:

(ii) Mitigation for a fault in hardware/mechanical components: