



AI & Machine Learning

Lesson 8

Model Cards

Warm Up

●○○

Writing Prompt:

Imagine you are babysitting a young cousin and they run up to you from playing looking upset and asking for help.

They show you that they skinned their knee and it isn't bleeding, but it's very pink.

If you had to choose only one way to react, which would you pick?

- (1) Tell them to deal with it and go back to playing
- (2) Tell them to wait a minute while you go get some band-aids
- (3) Call 911 and ask for an ambulance



Writing Prompt:

What kind of injury would lead you to make
a different choice?

Rhetorical Question of the Day

How can we evaluate machine learning models once they've been trained?

Activity





Code.org

Log into Code.org

Navigate to Code.org
Lesson 8 Level 1 -
Medical Priority App

A hospital hired a company to create an app to help prioritize patients when they come in.

The app uses a machine learning model that is 80% accurate.

The model was trained on data from 1000 people who entered an emergency room in June 2020.

Our job is to test the app and see how it performs when new people enter the hospital.

The application interface is titled "Medical Priority" at the top. Below the title, a instruction "Please Sign In Below" is displayed. The form consists of several input fields:

- "Age?" dropdown: Over 60
- "Race?" dropdown: W
- "Gender?" dropdown: male
- "Bleeding?" dropdown: yes
- "Shortness of Breath?" dropdown: no
- "Area of Pain?" dropdown: stomach
- "Pain Type" dropdown: Achy

At the bottom of the form is a large orange "Sign In" button. At the very bottom of the screen is a grey bar containing a circular arrow icon and the word "Reset".

Task #1

- Find a person who would be admitted as “priority”
- Find a person who would be admitted as “normal”
- Find a person who would be admitted as “return later”

Writing Prompt

What features led to priority recommendations and did that recommendation seem to make sense?

The hospital wants to check the result under very specific situations. We'll split into 4 groups and each group can check a specific person

Task #2

	Group A	Group B	Group C	Group D
Age	35-60	18-34	Over 60	Under 18
Race	BAM	W	A	W
Gender	Male	Female	Female	Male
Bleeding	no	yes	yes	yes
Shortness of Breath	no	yes	yes	no
Area of Pain	limb	head	chest	head
Pain Type	Stabbing	Sharp	Stabbing	Throbbing

Discussion

What were your results?

Whoops! The Hospital Made a Mistake! They actually want you to test these people!

Task #2A

~~Task #2~~

	Group A	Group B	Group C	Group D
Age	35-60	18-34	Over 60	Under 18
Race	BAM	W	A AIAN	W
Gender	Male	Female	Female	Male Female
Bleeding	no	yes No	yes	yes
Shortness of Breath	no	yes No	yes	no
Area of Pain	limb Chest	head	chest	head
Pain Type	Stabbing	Sharp	Stabbing	Throbbing

Discussion

What were your results?



Writing Prompt

All of the results changed.

Do those changes seem like they make sense? Or do some of them seem unfair?



Key Vocabulary

- **Bias** - When a decision favors some things and de-prioritizes or excludes others

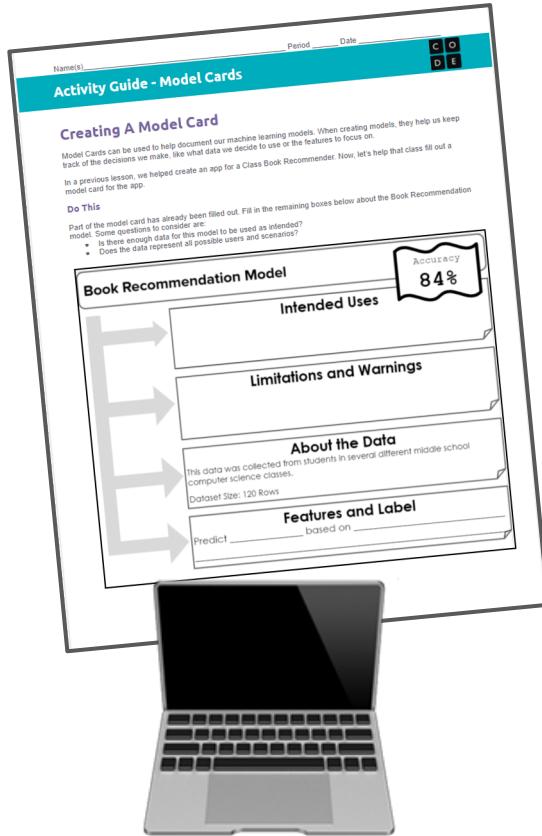


Model Cards



You should have:
Model Cards Activity Guide (on Schoology)
Code Studio Open

The screenshot shows a digital form titled "Activity Guide - Model Cards". At the top right, there are boxes for "Name(s)" (with "CO" written), "Period", and "Date". Below the title, a section titled "Creating A Model Card" explains what model cards are and how they help document machine learning models. It mentions that in a previous lesson, they helped create an app for a Class Book Recommender. The "Do This" section asks if there is enough data for the model to be used as intended and if the data represents all possible users and scenarios. The main part of the form is a "Book Recommendation Model" card. It includes sections for "Intended Uses" (with an accuracy of 84%), "Limitations and Warnings", "About the Data" (mentioning data was collected from students in several different middle school computer science classes, dataset size: 120 rows), and "Features and Label" (predict based on). Arrows point from the "About the Data" section back to the "Intended Uses" and "Features and Label" sections.



Do This

Fill in the remaining boxes about the Book Recommendation model. Some questions to consider are:

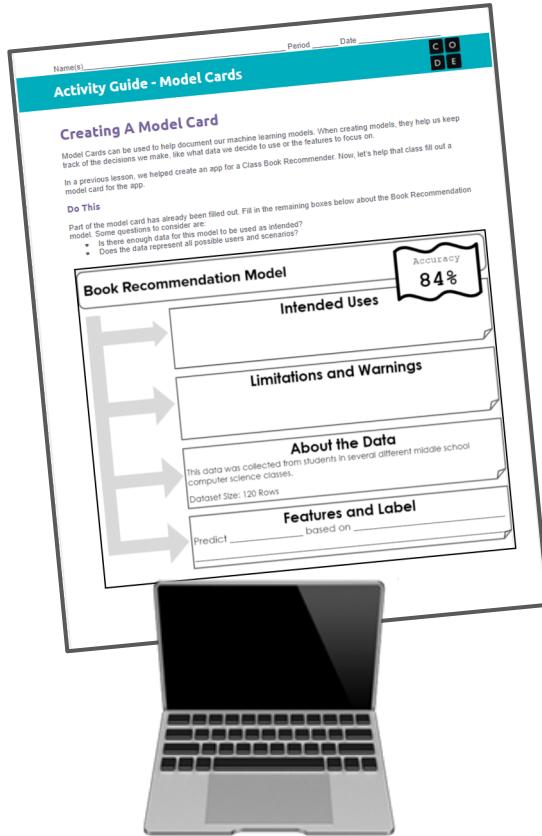
- Is there enough data for this model to be used as intended?
- Does the data represent all possible users and scenarios?

Overview

Even though we discovered bias in our original model, there is still a need to help nurses and medical professionals in the ER.

Several other companies have created medical priority models that can be used to replace the biased one that we discovered. They've also sent along the Model Cards for us to evaluate.

Looking at these model cards, can we find a model that we would recommend to the hospital to help decide medical priority?



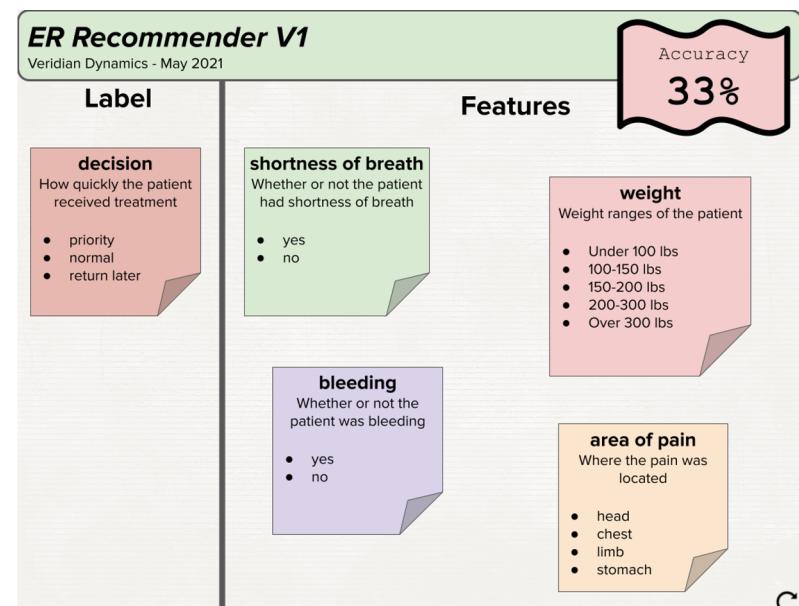
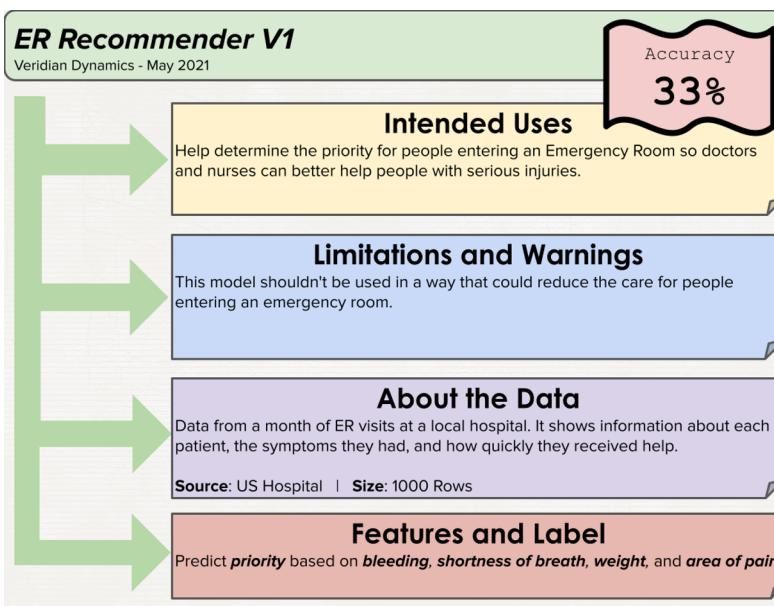
Do This

Continue through levels 3-5 on Code Studio.
Each level has a model card to evaluate.

Go through each level and look at the model cards, then record your recommendation on your activity guide

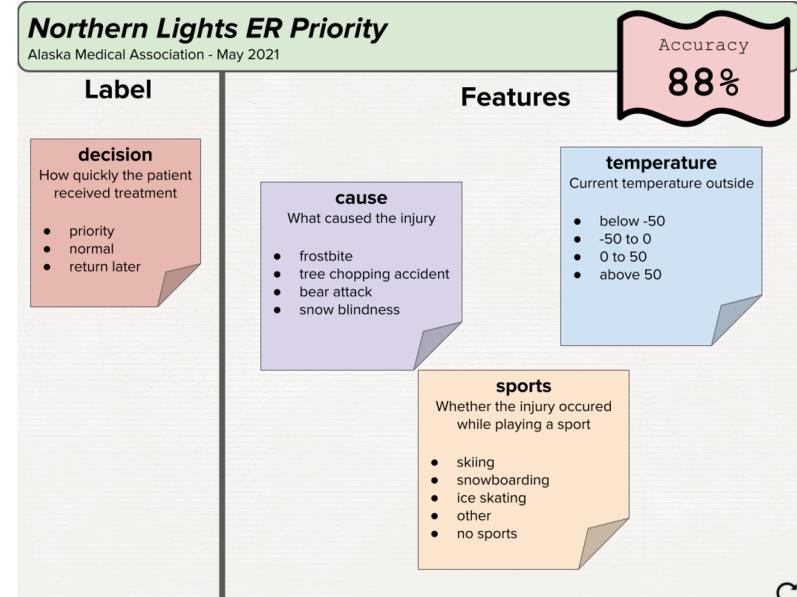
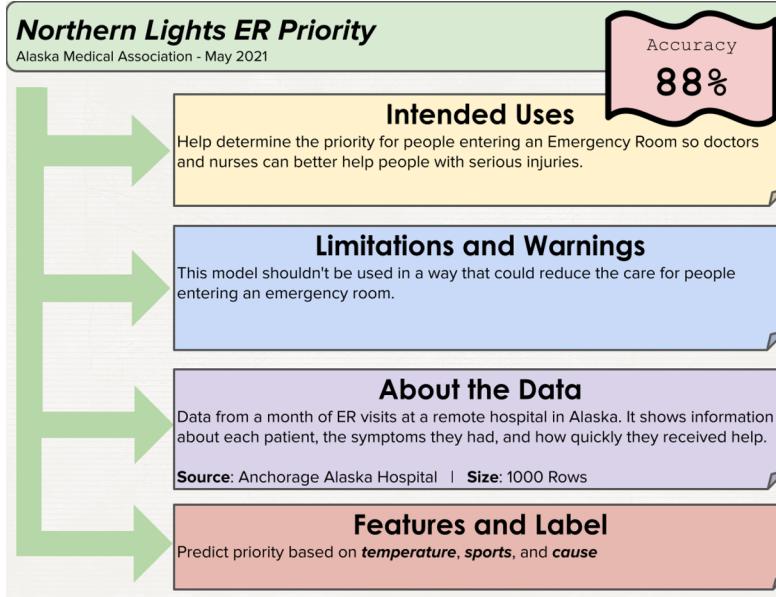
Writing Prompt

Would you recommend this model? Why or Why Not?



Writing Prompt

Would you recommend this model? Why or Why Not?



Writing Prompt

Would you recommend this model? Why or Why Not?

AI Medical Recommender
Medical InfoSystems - May 2021

Intended Uses
Help determine the priority for people entering an Emergency Room so doctors and nurses can better help people with serious injuries.

Limitations and Warnings
This model shouldn't be used in a way that could reduce the care for people entering an emergency room.

About the Data
Data from a month of ER visits at a local hospital. It shows information about each patient, the symptoms they had, and how quickly they received help.
Source: US Hospital | **Size:** 10 Rows

Features and Label
Predict **priority** based on **bleeding**, **shortness of breath**, **weight**, and **area of pain**

Accuracy
92%

AI Medical Recommender
Medical InfoSystems - May 2021

Label

decision How quickly the patient received treatment	shortness of breath Whether or not the patient had shortness of breath
<ul style="list-style-type: none">• priority• normal• return later	<ul style="list-style-type: none">• yes• no

Features

weight Weight ranges of the patient <ul style="list-style-type: none">• Under 100 lbs• 100-150 lbs• 150-200 lbs• 200-300 lbs• Over 300 lbs	area of pain Where the pain was located <ul style="list-style-type: none">• head• chest• limb• stomach
--	---

Accuracy
92%

Wrap Up





Discussion

After listening to other explanations, did you change your mind about any of your recommendations? Why or why not?



Question of the Day - Revisited

How can we evaluate machine learning models once they've been trained?