Imperfect Match

By Ellie Haberl

AI Story Series Editors: Bridget Dalton and Tom Yeh

University of Colorado - Boulder

June 2019

Word count 1404

Flesch-Kincaid reading level 6.5

Lucy pulled the blue folder out of the drawer filled with profiles of potential dogs for adoption. She looked across the desk at the couple and their two young children.

“Ah, here it is! Your match!”

She pasted on her big reveal grin, clapped her hands together and said, “It looks like you will be taking home Buttercup, a blond-haired golden retriever! She’s your family’s perfect match.”

But, inside, Lucy was just angry. Lucy had been working at the pet adoption shelter all summer. And she was starting to see a trend. As she handed the blue file to the children, she wondered to herself “Another golden retriever?”

Lucy loved animals. As a kid, she taped pictures of hamsters, cats, and birds around the house each time she wanted to convince her family she needed another pet. Her parents drew the line when they found a picture of tiny hedgehogs taped to their toothbrushes. “We are not getting a hedgehog.” Her mom declared mid-brush, making it sound a little more like, “Re war mot betting a hejwog.”

“Betting a hejwog?! Mom, I would NEVER bet a hejwog. They are much too precious to lose on one bad hand in poker!”

Lucy still remembers the day she found out that most shelters have to kill the animals that are not adopted within a certain timeframe. She threw herself on her mom’s bed, wailing: “No! We have to rescue them!”

That afternoon, her mom took her to the shelter to choose one dog that they would adopt. As soon as she saw Mac, she felt as though an enormous fleece blanket had been wrapped around her shoulders. Mac snuggled into her lap and nuzzled his face against her leg. It was simple. Mac was love.

Now, as a sophomore in high school, Lucy was the president of her high school’s pre-veterinarian club. In the summer, she worked at the pet shelter, helping animals get matched with their new owners. She loved the work, but it was hard sometimes to work at a shelter that is not a no-kill shelter.

Last December, the owners of the shelter, Rebecca and Mario Thornton, came into the office with a new laptop for Lucy. “Ta-da!” Rebecca placed the laptop on the counter in front of Lucy and proclaimed,

“This is it! This computer program is The Official-It-Will-Blow-Your-Mind-Game-Changer-For-Pet- adoption across the land. She bowed in front of Lucy like Lucy was some kind of queen of the pet shelter. Rebecca’s enthusiasm was annoying.

“Wow, really!” Lucy did her best Rebecca impersonation. “Across the entire land? The Kingdom, even? Hath it been tested with the lowly subjects yet? I mean, have we tried using it to match real pets with real people?” Lucy, a long-time member of the robotics club, knew how complex the program would have to be, and felt a little suspicious of Rebecca’s confidence.

Mario knelt, searching for the best place to plug in the charger. “The team that built the program used machine learning and artificial intelligence (AI) to devise a model that can predict the best fit between an animal and a family. They built the model by taking data from hundreds of pet owners and their reported satisfaction with their pets. This “training data” produced a model that can essentially use this information to predict the pet that will fit best with a family. We actually think it has the potential to reduce the need for shelters. I mean, people won’t return pets or abandon pets if they’re a good fit.”

“Oh, Lucy responded, “So it’s like a dating site for pets, only people never get to see the options. The computer makes the match.”

“Exactly. It even sounds like a dating site. It’s called PerfectMatch.” Mario turned the computer on to start the program for a quick demo.

Lucy was thinking about the first time she held her dog, Mac. She couldn’t imagine the model would be able to predict the loving connection she felt.

Weeks passed and Lucy enjoyed using the program. It wasn’t until the third or fourth week that she noticed a pattern. Diego, a fellow programmer and her best friend in the robotics club, visited the shelter to bring her ice cream one night, and Lucy shared her observation. “Here’s the thing,” Lucy said, “I just have this feeling. Some breeds are getting matched over and over again. And some breeds are never getting chosen. I’m telling you, it’s weird. And I’ve noticed pets with disabilities never, ever get matched. And I feel like I have given away more blond golden retrievers than there are blond golden retrievers on the planet, or at least a billion of them!”

Diego loved the way Lucy exaggerated everything, always using phrases like “one million billion.”

Lucy added, “During last week’s robotics club meeting, I started doing a little research on the way the PerfectMatch model works. I noticed that the model weights certain variables more than others, like hair color. Don’t you think that’s a terrible metric for pet satisfaction? No one in the trillions of people ever says, I love my pet so much because she has blond hair. Plus, the model never matches big dogs with families who live in apartments! Maybe the family lives near a park where the dog can run around and maybe they really want a big dog! The model should be more nuanced and include that information.”

Suddenly, Lucy heard Mac barking outside of the door. He often came to work with Lucy and she let him run around outside for a while before coming in. “I think about the way Mac would be described in the data, before I found him. He barked all the time when I first adopted him, but he actually just needed a better trainer!”

Diego added, “I see what you mean. When I look at the model, I can see it is giving more weight to breed than personality. So, retrievers might be chosen more frequently than dogs that are also easygoing, but maybe belong to a different breed.”

Now Mac was curled up in Lucy’s lap. It made her so angry to imagine that she wouldn’t have been matched with him. She continued, “And here’s the other thing, Diego. I know fewer people adopt dogs with disabilities in the first place! So maybe there is a smaller amount of data on those dogs in the training data that built the original model. Or maybe the training data relied more heavily on owners who had returned dogs with disabilities. This is anecdotal evidence, of course, but Mac is blind in one eye and I much preferred him to some dogs I know with fancy perfect vision.” Lucy laughed.

Diego joined in, “So there you have it. Mac is the match that makes you happy, but the AI model as we’re looking at it now probably would have recommended a blonde lab with 20/20 vision because the hundreds of pet owners who reported their experience with pets was already a biased sample.

Lucy wondered, “But Mario said the model is trained from hundreds of families. It should be pretty good, right? Why is this strange thing happening?"

Diego responded, “I guess the sample isn’t large or diverse enough. As just one example, I can see that the training data was primarily taken from suburban neighborhoods where larger houses and larger dogs are more common, so that is probably skewing the recommendations for other families.”

Lucy lifted her spoon “Ok. Here’s my plan. Let’s eat the rest of this ice cream carton, and then write up a case for ditching PerfectMatch that we can present to Mario and Rebecca.

“And maybe offer to build a fairer model by using fairer training data?” Diego asked.

“Yes! Exactly! It’s the bias in the training data that is producing a biased model. We can help improve the model by uncovering the way this bias was perpetuated in this data in the first place and using a fairer sample for a new model. Maybe PerfectMatch will buy our model to improve their business or maybe we can make the program open source, so more pets will be matched with more families. Just think of all those animals and people who are waiting to be a family! I bet there are millions.”

This time, they sensed that her exaggeration was pretty close to accurate. And they both smiled.

Discussion Guide

A few questions to support discussion and reflection (and/or prompts to support a freewrite or turn and talk to prep for larger conversation):

1. Does your family have any pets?  If so, what factors went into your family choosing the pet in the first place?
2. If your family does not have a pet...do you want one?  What factors would go into you choosing your own pet?
3. Let's pretend like the data set this AI program is using as the training data is based on this classroom.  What is your ethnicity, family size, and pet of choice?
4. With this information in mind, will we see a pattern in the data within our own classroom?  Just because there are patterns does this mean all of us want what the data suggests?
5. Why did Mario want to use AI for his company? Why is he so excited?
6. Why is Lucy upset?
7. Is everyone really going to be happy if the model is made more fair? Who might not be happy?
8. Does this story remind you of other uses of artificial intelligence that you know about from the news or your own experience? Do you think those other uses of AI might also have bias training data creating the program? How might it be bias?
9. Does a larger sample size always solve the problem of bias in training data?
10. How can we improve this AI program? Lucy and Diego think about a few approaches they would take like getting a fairer sample for the training data, but can you think of other approaches that might also improve the AI program?