

A watercolor illustration featuring several stylized robots and a woman. The robots are depicted in various colors and designs, including blue, red, and grey. One robot in the center has a blue body and a red head. Another robot on the right has a red body and a blue head. A woman with blonde hair, wearing a red top and an orange skirt, stands on the right side. The background is a light, textured grey.

Human AI Interaction

Lecture 13: Power
aidesignclass.org

Today

- A continuation of data - generating synthetic data, and why
- Assignment 3
- AI and power - and what it means for data collection

A few terms to know

- Supervised data: <input> -> <output>
 - E.g. Instruction -> Results of instruction
 - “Write a poem” -> “<poem>”
- Unsupervised data: only <input>
 - E.g. a collection of images, webpages etc (but see semisupervised)
- Semi-supervised data: <input> -> <output mechanically found from input>
 - E.g. fill in the blank (covering consecutive words)
 - Eg. breaking a sentence “X because Y” to “Why X?” -> “Because Y” (input -> output)
- Synthetic data: artificially created data, to serve a particular purpose

Why synthetic data

- A lot cheaper, faster to create.
- Can precisely target some kinds of data that you are missing
- Can explore new areas that people are unlikely to create in supervised data

Not an exclusive approach - can be combined with other methods of data generation.

Synthetic data allows for exploring new areas

MOVE BY MOVE

Google's AI won the game Go by defying millennia of basic human instinct



Man versus machine.

Image: AP/Lee Jin-man

<https://qz.com/639952/googles-ai-won-the-game-go-by-defying-millennia-of-basic-human-instinct>

Creating synthetic data

- For text, why... it's just predicting the next words!

The following is a conversation between a human and an intelligent AI Assistant:

USER: I'd like a simple recipe for pancakes.

ASSISTANT:

Creating synthetic data: set the right context

- For text, why... it's just predicting the next words!

Below are a series of dialogues between various people and an AI assistant. The AI tries to be helpful, polite, honest, sophisticated, emotionally aware, and humble-but-knowledgeable. The assistant is happy to help with almost anything, and will do its best to understand exactly what is needed. It also tries to avoid giving false or misleading information, and it caveats when it isn't entirely sure about the right answer. That said, the assistant is practical and really does its best, and doesn't let caution get too much in the way of being useful.

USER: I'd like a simple recipe for a healthy salad.

ASSISTANT:

Creating synthetic data: add an example!

- For text, why... it's just predicting the next words!

Below are a series of dialogues between various people and an AI assistant. The AI tries to be helpful, polite, honest, sophisticated, emotionally aware, and humble-but-knowledgeable. The assistant is happy to help with almost anything, and will do its best to understand exactly what is needed. It also tries to avoid giving false or misleading information, and it caveats when it isn't entirely sure about the right answer. That said, the assistant is practical and really does its best, and doesn't let caution get too much in the way of being useful.

USER: If the top 10% of Americans own 80% of the wealth, does that mean 1 in 10 people I see on the street have significantly more money than me?

ASSISTANT: If you took every single American, put them in a big mixer bin, and then used a crane to fish out 10 of them at random, you would expect to find one of them to have a significant amount of money compared to the others. You may or may not actually get that result due to luck of the draw, but if you repeated this over and over, you'd average that amount.

Just walking down any street, though, it depends a lot on who actually visits that street. If it's a back alley in a small town in the Midwest, you probably won't meet any people who make a lot. But if it's Wall Street in New York City, probably everyone there makes quite a bit.

USER: I'd like a simple recipe for a healthy salad.

ASSISTANT:

Creating synthetic data: Separate out instructions and responses

- For text, why... it's just predicting the next words!

Come up with a diverse list of tasks:

Task 1: Explain: If the top 10% of Americans own 80% of the wealth, does that mean 1 in 10 people I see on the street have significantly more money than me?

Task 2: I'd like a simple recipe for a healthy salad.

Task 3: How do I do a handstand?

Some other things that work well

- Filtering out quality of responses (Ask an LLM!)
- Asking to create a diverse set
- Suggesting kinds of tasks (make a list, explain...)
- Suggesting ways to evolve the list of instructions (take on different roles, etc.)

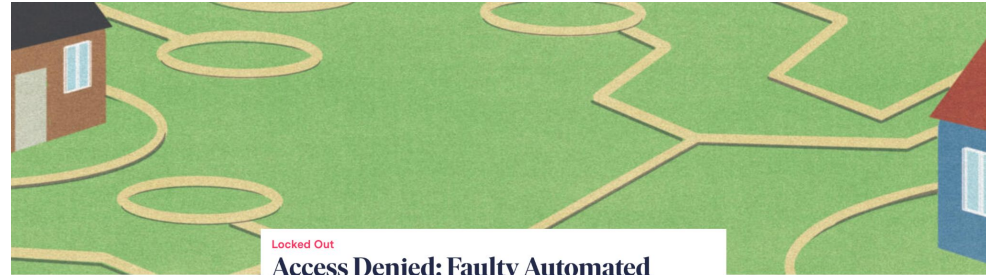
Assignment 3: create your own synthetic data!

See canvas for details

- 1000 instructions, using the model we saw today
- You'll create high-quality data
- You're rewarded for creativity

Read the article: what went wrong?

<https://themarkup.org/locked-out/2020/05/28/access-denied-faulty-automated-background-checks-freeze-out-renters>



May 28, 2020 05:00 ET

Andrea Ucini

Locked Out

Access Denied: Faulty Automated Background Checks Freeze Out Renters

Computer algorithms that scan everything from terror watch lists to eviction records spit out flawed tenant screening reports. And almost nobody is watching

By [Lauren Kirchner](#) and Matthew Goldstein, of The New York Times

Share This Article

Copy Link

This article is copublished with
The New York Times

Burglary and domestic assault in Minnesota. Selling meth and jumping bail in Kentucky. Driving without insurance in Arkansas. Disorderly conduct. Theft. Lying to a police officer. Unspecified “crimes.” Too many narcotics charges to count.

That’s what the landlord for an apartment in St. Helens, Ore., saw when he ran a background check for Samantha Johnson, a prospective tenant, in 2018.

But none of the charges were hers.

The growing data economy and the [rise of American rentership](#) since the 2008 financial crisis have fueled a rapid expansion of the tenant screening industry, [now valued at \\$1 billion](#). The companies produce cheap and

