

How To Train Your AI: Demystifying ChatGPT With Machine Learning, Neural Networks, and Deep Learning Basics

Workshop Overview

- **Duration:** ~1 hour (approx. 30 min presentation, interactive activity, & discussion, 30 min coding activity)
- **Required Materials:** Pen, paper, Microsoft accounts, and laptops with Internet access, preferably 1 per 4-5 students.
- **Audience:** Students ranging from high school to college, with skill levels spanning from little to no technical background to those with significant experience.
- **Goal:** By the end of this workshop, students should have a fundamental understanding of what AI and ChatGPT are doing under the hood, how Machine Learning (ML), Neural Networks (NNs), and Deep Learning (DL) fit together, and experience a hands-on coding activity to reinforce these concepts in a simplified form.

Introduction: What Is Artificial Intelligence?

Modern Artificial Intelligence (AI), including **Generative AI** like ChatGPT, might seem magical, but it's really about patterns, math, and practice—lots and lots of practice. It's not the sentient, futuristic robots of Hollywood (*iRobot*, *Avengers: Age of Ultron*, *Terminator*, *WALL-E*, etc.).

AI doesn't "understand" things the way humans do. When shown an image or given a prompt, it doesn't truly know what it's looking at or what you're asking. Instead, AI functions as an advanced computer system that analyzes data to detect patterns, makes educated guesses, verifies them against known outcomes, and improves over time through feedback. Eventually, through a process of trial and error, it gets so good at predicting and generating outputs that it seems intelligent.

Key Concept: At its core, AI is about training computers to learn from data. This is achieved through a process called Machine Learning, which can be thought of as a highly sophisticated, scaled-up version of guess-and-check. The AI (a complex computer program) makes a guess, compares it to the correct answer provided by humans—whether it's identifying an image of a cat or predicting the next word in a sentence—and adjusts itself based on the result. By repeating this millions or billions of times with real, human-created data, the AI refines its guesses until it produces results that seem like true understanding but are actually just well-tuned patterns.

The Foundations of AI: Machine Learning, Neural Networks, and Deep Learning

To understand how AI systems like ChatGPT work, it's essential to grasp three interconnected concepts:

Machine Learning, **Neural Networks**, and **Deep Learning**. These terms describe the methods and structures that allow AI to learn and make predictions. At its core, that's all AI is doing—taking input, analyzing it, and

guessing the output based on mathematical calculations and probabilities (those high school and college calculus and statistics classes really do matter!).

Here's the process broken down: AI simplifies the input it receives—whether it's an image or text—into numbers that computers can understand. For images, this means breaking them down into **RGB values** (the red, green, and blue components that combine to form every pixel of the image). For text, the process is similar but more abstract: the words are converted into numerical representations through techniques like word embeddings, which map words or phrases to arrays of numbers, capturing their meaning and relationships.

By processing these numerical inputs, the AI relies on math and probability to identify patterns and make predictions, continuously refining its approach through training and feedback. This combination of simplification, pattern recognition, and iterative learning is the foundation of how systems like ChatGPT work.

In Short:

- **Machine Learning** is the overall field of teaching computers to learn from examples.
- **Neural Networks:** A specialized Machine Learning approach that organizes many small computing units ("artificial neurons") into interconnected layers. Each layer refines the data based on what previous layers found, allowing the network to learn richer, more abstract patterns from the raw input.
- **Deep Learning:** A subset of Neural Networks that uses many layers of these artificial neurons. With more layers, the model can capture extremely complex relationships, resulting in more accurate and sophisticated predictions.

Different Types of AI Training

Before diving deeper into Machine Learning, Neural Networks, and Deep Learning, let's explore how AI models are trained. The way an AI learns depends on the type of data it's given and the task it needs to perform. This helps clarify how a system like ChatGPT is trained versus an AI designed to recognize cats in images.

1. Supervised Learning (Labeled Data):

Imagine you want an AI to recognize cats in images. In supervised learning, humans label data ahead of time—e.g., "this is a cat" or "this is not a cat." The AI looks at these examples, guesses, checks its guess against the label, and improves by adjusting how it processes similar data in the future. After seeing enough labeled examples, it becomes skilled at identifying cats in new, unlabeled images.

Supervised learning works well for tasks like image recognition or spam email detection, where we can easily provide labeled data.

2. Self-Supervised Learning (Patterns in Raw Data, No Labels):

ChatGPT and other large language models use **self-supervised learning**, which doesn't rely on human-created labels for every piece of data. Instead, the AI trains itself by learning patterns in massive amounts of raw text.

Here's how it works for ChatGPT:

- The AI sees a sentence like "The cat sat on the ____" and tries to guess the missing word (e.g., "mat").

- If the guess is wrong, the system fine-tunes its process to handle similar patterns better next time.
- By repeating this process billions of times, the model becomes very good at predicting the next word based on context, which allows it to generate coherent sentences, understand grammar, and mimic human-like responses.

Unlike the cat example, ChatGPT isn't learning fixed answers like "this is a cat." Instead, it's learning how language flows, how words and ideas are related, and how to predict the best continuation for a given prompt.

Key Concept: Supervised learning is like giving the AI a multiple-choice quiz with the answers provided for grading. Self-supervised learning, used by ChatGPT, is like giving the AI a fill-in-the-blank test with no explicit answers—just patterns from the text itself. Regardless of the method, the core process remains the same: guess, check, adjust.

Machine Learning: Teaching Computers to Learn

Machine Learning is the broad field of teaching computers to learn from examples rather than following step-by-step instructions. Instead of explicitly programming a computer to recognize a cat or predict the next word in a sentence, you give it data and let it figure out patterns on its own.

How It Works

1. **Input Data:** The computer is given data (e.g., images, text).
2. **Prediction:** It makes a prediction (e.g., "Is this an image of a cat?", "Is 'roll' the next word in the phrase 'rock and ___'?").
3. **Feedback:** It checks its prediction against the correct answer provided in the data (e.g., "Yes, this is a cat" or "No, the correct word is 'roll'").
4. **Adjustment (Weights):** The computer adjusts **weights**, which are numbers that represent how strongly it should consider different features of the input. For example:
 - In an image, certain pixel patterns might become "important" for recognizing a cat.
 - For text, certain word combinations might weigh more heavily in predicting the next word.
5. **Repetition:** This process is repeated millions of times until the model gets highly accurate.

Key Concept: The computer doesn't "know" what a cat or a word is—it just analyzes the examples we provide and learns the patterns that humans have created in real-world data. These examples form the foundation for its guesses and improvements.

What Are Weights?

Imagine weights as dials that the computer adjusts to make better guesses. These dials control how much attention the system pays to different parts of the input data.

- **Example 1 (Image):** If the computer is learning to recognize a cat, it might initially pay equal attention to all parts of an image. Over time, it adjusts its weights to focus more on patterns like whiskers or triangular ears.
- **Example 2 (Text):** When predicting the next word in "The cat sits on the ___," the model adjusts its weights to give more importance to words like "sits" and "on," which help predict "mat".

Weights are the computer's way of learning which features in the data matter most for the task it's trying to solve. At first, these weights are random. As the model gets feedback, it tweaks the weights to improve.

Adjusting weights is like refining your guesses over time. Imagine if you guessed answers on a multiple-choice test without studying. After seeing which answers were correct, you'd (hopefully) start noticing patterns and improve your guesses on the next test. Weights are how the computer "learns" those patterns and becomes more accurate.

Pseudo Code Example

Visit the link below to view a simplified Python-based pseudo-code to connect the concept of Machine Learning with code. This example uses a basic "guess-and-check" loop for predicting whether an image is a cat.

[Pseudo Code | Machine Learning](#)

Key Concept: We used "brightness" as a stand-in for any feature the AI might consider. In reality, a model might look at many different features, not just brightness. Over time, by adjusting the weight, the model learns whether brightness is a good clue to determine if something is a cat. This is, of course, a simple example—real Machine Learning uses more complex algorithms and data structures.

Neural Networks: Mimicking How Brains Work

Neural Networks are a way to teach computers to recognize patterns, inspired by how the human brain works.

Neural Networks are a subgenre of Machine Learning. They're a specific approach within the larger field. Your brain has tiny units called neurons, each responsible for processing specific details, like detecting colors, edges, or shapes in a picture. Neural Networks mimic this process in computers, but their "neurons" are virtual—small mathematical functions or programs that analyze bits of information to recognize patterns.

These neurons are organized in **layers**, like an assembly line. Each layer has a specific job:

1. The first layer picks out very basic details.
2. The second layer builds on those details to find bigger patterns.
3. The third layer uses those patterns to figure out what the input might be.

By stacking layers, the network can go from seeing small details to recognizing the bigger picture.

How It Works

When Neural Networks process data, they pass it through layers, each one doing a little more work than the last. Each layer has a job that builds on the previous one, going from simple details to more complex patterns. This layered approach allows the network to recognize almost anything—from images to text—by breaking the problem into smaller, manageable steps. It's like solving a puzzle: each layer contributes one piece until the big picture emerges.

Let's break it down with **two examples**:

Example 1: Image-Based Input

Suppose you're training a Neural Network to recognize pictures of cats:

1. First Layer (Basic Details):

The first layer of neurons looks at the raw data (the image, which is really just a grid of numbers representing pixel colors).

- Example: It might notice lines, edges, or areas with similar colors.

2. Second Layer (Patterns):

The second layer takes those lines and edges and looks for shapes.

- Example: It might spot circles (eyes) or triangles (ears).

3. Third Layer (Recognition):

The third layer combines those shapes into objects.

- Example: It might say, "Two circles and a triangle arranged like this? Probably a cat."

Example 2: Text-Based Input

What about recognizing text or generating words, like **ChatGPT**?

1. First Layer (Basic Word Features):

The first layer takes a sentence as input. Each word is converted into numbers (called vectors) that represent its meaning or context.

- Example: The word "cat" might be represented by numbers showing its relationship to words like "pet" or "animal."

2. Second Layer (Relationships Between Words):

The second layer analyzes how the words are related.

- Example: It might see that "cat" and "sits" often appear together.

3. Third Layer (Understanding Context):

The third layer uses this relationship to understand the overall meaning or predict what comes next.

- Example: If the input is "The cat sits on the ___," the network might predict "mat" because it has seen that pattern many times during its training on real text that humans wrote.

Key Concept: Each layer passes its output to the next layer. The first layer might pick out very simple patterns, and the next layer builds on that to find more meaningful patterns. Finally, we make a guess at the end. If we're wrong, we adjust the weights throughout the network so we do better next time.

Pseudo Code Example

Visit the link below to view a Python-based pseudo code example showing how a Neural Network processes data through layers. We'll keep this as simple as possible!

[Pseudo Code | Neural Networks](#)

That's it! It's the same guess-check-adjust process as before, but now we have layers:

- Layer 1 transforms the input

- Layer 2 uses Layer 1's output to make a final guess
- If we're wrong, we tweak weights in BOTH layers so the entire network improves next time.

Deep Learning: Neural Networks, Supercharged

Deep Learning takes the idea of Neural Networks a step further by adding many more layers—making the network "deeper." More layers mean the network can recognize much more complex patterns. It's still the same guess-check-adjust cycle you've seen in Machine Learning and Neural Networks, but on a much larger scale and often using more powerful computers to handle all the extra calculations.

Key Concept: Deep Learning enables systems like ChatGPT to understand the context of sentences, recognize objects in complex images, or even generate realistic-sounding text. Each additional layer refines the data a bit more, allowing the model to identify very subtle relationships. It's the backbone of the most advanced AI applications today.

How It Works

1. **Many Layers:** Instead of just one or two layers, Deep Learning models might have dozens, hundreds, or even thousands. Each layer focuses on a different aspect, starting from basic patterns and moving toward very complex concepts.
2. **Complex Patterns:** More layers let the model detect extremely intricate patterns. For an image, early layers might find edges and shapes; middle layers might recognize parts of objects (like eyes or ears), and later layers might piece these parts together into entire objects (like a cat sitting on a mat). For text, early layers understand letters or word pieces, middle layers understand word meanings and sentence structure, and later layers grasp entire paragraphs or even the style of writing.
3. **Massive Data and Computation:** Deep Learning usually involves huge amounts of data and often requires specialized hardware (like GPUs) to process everything quickly. This makes sense because with many layers, there are many more weights to adjust and a lot more learning to do.

In Short: Deep Learning is just neural networks with lots and lots of layers, allowing the model to learn incredibly detailed and sophisticated patterns.

Pseudo Code Example

We've seen how Machine Learning and Neural Networks work in simplified code. For Deep Learning, the idea is exactly the same—guess, check, adjust—but with more layers. Let's show a very high-level, extremely simplified pseudo code example. Visit the link below to check it out.

[Pseudo Code | Deep Learning](#)

What's Important Here:

- We're doing the same basic steps—multiplying inputs by weights, adding them up, making a guess, checking the guess against the label, and adjusting weights.
- The difference is just that we do this many times in a row, through many layers, and adjust many weights.
- Over time, this lets the model learn extremely complicated relationships that a single-layer model couldn't handle.

Putting It All Together

- **Machine Learning:** One layer, guess-check-adjust cycle, simple features.
- **Neural Networks:** Multiple layers, each building on the last, letting the model understand more complex patterns.
- **Deep Learning:** Many layers, allowing the model to become even more powerful and capable of understanding very intricate patterns and relationships—just what we need for advanced AI systems like ChatGPT.

Each concept builds on the previous one, and the core idea never changes: we guess, we see how far off we are, and we adjust our weights. The more layers we have, the more complex and nuanced patterns the model can recognize, and the more impressive the predictions and outputs can be.

Activity 1: From Rules to Patterns: How AI Learns

Goal: Help students grasp how AI moves beyond human-written rules to learn patterns from data. They'll first try to identify cats in images and write rules by hand (simulating traditional programming logic), then discover why giving examples and letting the AI figure it out (supervised learning) is more flexible. After that, they'll explore text-based fill-in-the-blank guessing to understand self-supervised learning—the technique ChatGPT uses.

Materials Needed: The PowerPoint presentation accompanying this workshop found at [How To Train Your AI: Demystifying ChatGPT With Machine Learning, Neural Networks, and Deep Learning Basics](#). Pen and paper for each student.

Estimated Time: 25–30 minutes total (about 15 minutes for images/cats activity, 10 minutes for text/fill-in-the-blank activity).

Step-by-Step Instructions

Part 1: Image-Based (Supervised Learning)

1. Set the Stage (2 minutes):

Show a slide with a very simple "Is this a cat?" question and an easy cat image. Say:

"Right now, I'm the computer and you're the programmers. I need rules to know if an image contains a cat. Let's see how you might do this."

2. Identify Cats Without AI (5 minutes):

Show a series of images: some obviously cats (domestic cats), some clearly not cats (dogs, cars, etc.), and then some tricky ones (lions, lynxes, ocelots, panthers, caracals, cartoon cats like Garfield, plush cat toys). After each image, ask:

"Is this a cat or not a cat?"

Have students hold up a piece of paper with "Cat" or "Not Cat" written, or just say it out loud. Encourage quick answers.

3. Write Human-Made Rules (5 minutes):

Ask each student:

"Try to write a set of rules that would let a computer decide if an image is a cat. Be specific: maybe 'has whiskers', 'has pointy ears', 'four legs', etc."

Give them a couple of minutes to write down their rules.

4. Test the Human-Made Rules (3 minutes):

Show some trick images:

- A lion: Does it follow their rules? Is it still a "cat" by those rules?
- A cartoon cat (like Garfield): It's a cat but doesn't look realistic. Does it break their rules?
- A cat that only has 3 legs as a result of an accident, or a cat with some other form of disability. Does it break their rules?
- A fox or a dog that looks somewhat cat-like: Does it fool their rules?

Ask volunteers to share their rules. Point out how their rules fail in certain edge cases.

5. Transition to Supervised Learning (2 minutes):

Explain:

"As you saw, writing rules by hand is tough. What if we just give the computer many images of cats and not-cats, all labeled, and let it figure out which patterns matter? This is supervised learning. Instead of you writing rules, you give lots of examples and let the computer guess-check-adjust until it can recognize cats on its own."

Key Message: With supervised learning, we don't have to write explicit rules. The computer develops them by seeing many labeled examples and adjusting its internal 'knobs'. This is how we have AI systems that can recognize images, voice, spam emails, and generate similar content.

Part 2: Text-Based (Self-Supervised Learning)

1. Fill-in-the-Blank Sentences (3 minutes):

Show popular phrases on slides with one word missing. For example:

- "Rock and ___" (students will say "roll")
- "Let's get ready to ___" (Students: "rumble")
- "Very mindful, very ___" (Students: "demure")
- "Started from the ___ now we here" (Students: "bottom")
- "May the ___ be with you" (Students: "force")
- "That's what she ___" (Students: "said")
- "To infinity and ___" (Students: "beyond")
- "With great power comes great ___" (Students: "responsibility")
- "It's about drive, it's about ___" (Students: "power")
- "One does not simply walk into ___" (Students: "Mordor")
- "Bond, James ___" (Students: "Bond")
- "A Lannister always pays his ___" (Students: "debts")
- "Actions speak louder than ___" (Students: "words")
- "The pen is mightier than the ___" (Students: "sword")
- "You reap what you ___" (Students: "sow")
- "Out of the frying pan, into the ___" (Students: "fire")
- "What we do in life, echoes in ___" (Students: "eternity")
- "May thy knife ___ and shatter" (Students: "chip")

Each time they fill in the blank easily, say:

"See how you guessed the right word based on what you've heard before?"

2. Harder Phrases Without Context (3–4 minutes):

Now show sentences where the next word is not obvious without context. For example: a random sentence from a blog or a tricky phrase they haven't heard. Ask them to guess the missing word. They'll struggle.

- "The rain fell softly against the window as she picked up her favorite ____" (Answers: "book, pen, mug")
- "Without hesitation, she stepped onto the platform and grabbed the ____" (Answers: "microphone, handrail, package")
- "The study revealed an unexpected correlation between sleep and ____" (Answers: "creativity, memory, productivity")
- "Philosophers often debate whether free will is an illusion or a ____" (Answers: "necessity, myth, paradox")
- "To all reasonable people, the best Capri Sun flavor is, without question, ____" (Answers: "Strawberry Kiwi, Pacific Cooler, Fruit Punch, Tropical, Orange")
- "Among pastries, ____ are the best" (Answers: "Cinnamon Rolls, Croissants, Donuts, Eclairs, Danishes")

Explain:

"If we tried to write a computer program by hand to guess the next word, it'd be really hard! There's no neat set of rules."

3. Introduce Self-Supervised Learning (3 minutes):

Say:

"ChatGPT learns by seeing billions of sentences from human made examples like books, articles, social media posts, etc., and guessing the next word over and over. It doesn't have labels telling it what the words mean. Instead, it teaches itself patterns of language by filling in blanks. Every time it gets it wrong, it adjusts how it picks the next word next time. Eventually, it becomes so good at predicting words that it can hold complex conversations, write stories, and explain concepts!"

Key Message: Self-supervised learning (what ChatGPT uses) doesn't need explicit labels for every example. It just learns patterns from raw data (text) by trying to predict missing parts, adjusting itself, and getting better over many, many attempts.

Activity 2: Guess, Check, Adjust: Training AI One Step at a Time

Goal: Transition from understanding concepts to hands-on coding experience. Students will explore how a basic text prediction model works by writing and running simple Python code. They'll start with a basic model that produces nonsensical predictions, then make small adjustments (like tweaking parameters for word selection) and rerun the code to observe improvements. By the end, they'll have an interactive AI-like model they've built and improved themselves.

Materials Needed:

- Laptops with Internet access for students to visit the workshop link: [Guess, Check, Adjust: Training AI One Step at a Time](#). Visual Studio Code for Education provides a browser based coding environment

that the students will work in.

- Each student will need a Microsoft account to log in.
- Remind students to click the play button in the upper right hand corner of the Visual Studio Code for Education environment in their browser to run each code file. The files prompt students for input, providing instructions on options they can use to tweak the training parameters.
- Ideally, one laptop per 3–4 students for collaborative group work, but more individual devices are even better.

Estimated Time: 30–40 minutes, depending on students' technical background and prior coding experience. Adjust timing for additional guidance or discussion as needed.

Overall Flow of Coding Activities

Each of these files is accessible from the VS Code for Education project found at [Guess, Check, Adjust: Training AI One Step at a Time](#). Students will progress through the workshop one file at a time, each building on the previous.

- **File 1: 01_basic_predictor.py**

This file picks words at random with no context or learning. Students see nonsense output and understand this as the baseline—pure randomness before any intelligence is added.

- **File 2: 02_markov_improved.py**

In this file, students learn about **Markov chains** and **temperature**:

- **Markov Chain (window_size):**

A Markov chain picks the next word based on the words that came right before it. By setting a `window_size`, the model looks at that many previous words to decide what comes next. For example, if `window_size=2`, it checks the last two words and picks a word that often followed those two words in the training text. Increasing `window_size` generally gives the model more context, leading to more coherent output.

- **Temperature:**

Temperature controls how "adventurous" the model is when choosing the next word. A lower temperature (e.g., 0.2) makes the model stick more closely to the most common or "safe" choices, resulting in more predictable and coherent text. A higher temperature (e.g., 1.0) allows the model to pick less common words more often, adding variety and creativity but sometimes producing less coherent sentences.

Students start with less ideal defaults (like `window_size=1` and `temperature=1.0`) and must experiment to achieve more coherent and satisfying results.

- **File 3: 03_interactive.py**

In the final file, students take full control. They adjust `window_size`, `temperature`, and `output_length` interactively and choose a starting word. They can try the suggested starting words (e.g., "On", "I", "The") or pick their own. By experimenting, they see how their choices influence coherence, style, and variety, directly building on all previous lessons. They should realize it's still limited because it only knows the small dataset they trained it on, unlike ChatGPT's massive training.

Each file in the Visual Studio Code for Education project includes detailed comments and instructions, telling students what to change and what to observe.

Before Starting

- Ensure students have access to the Visual Studio Code for Education environment and the `data.txt` file.
- Students should see: `README.md`, `data.txt`, `01_basic_predictor.py`, `02_markov_improved.py`, and `03_interactive.py`.
- Let them know the `data.txt` file has a small dataset. Remind them that small datasets limit the complexity and realism of the text but still demonstrate the core concepts.

Ideal Settings for Facilitators (Suggested Parameters)

As students progress through the files in this workshop, they will experiment with training a mock AI system by providing input settings such as `window_size` and `temperature`. The goal is for them to tweak these settings on their own and observe how changes impact the generated output. However, if students struggle or ask for guidance, here are the recommended settings based on the `data.txt` file provided with the Visual Studio Code for Education project. These settings are not "right" answers but are optimized for this dataset to produce the best learning outcomes.

File: `01_basic_predictor.py`

Purpose: Establishes a baseline of random output. No parameters to tweak here—just run it and see nonsense. Encourage students to look at `data.txt` after running this file to observe the training data the code is using.

What Students Will Observe: Complete randomness, demonstrating the need for learning and context. This file has no settings. It's meant to be run as a baseline to show how useless computers are until we, humans, introduce some logic teaching them how to make guesses, check them, adjust, and then eventually improve, so much so that we end up with something like ChatGPT in our world today.

File: `02_markov_improved.py`

Purpose: Introduces `window_size` (context) and `temperature` (predictability vs. creativity).

- **First Run (Defaults):**
 - `window_size = 1`
 - `temperature = 1.0`
 - `fixed_output_length = 30` **Observation:** Slight improvement over pure random but not very coherent.
- **Suggested Improvement:**
 - `window_size = 3`
 - `temperature = 0.5`
 - What Students Will Observe:** More coherent phrases but not quite perfect.
- **Further Exploration:**
 - Try optimal values of `window_size = 6` or even higher values within the acceptable range.

- Experiment with optimal **temperature** of around **0.7** to **1.0**.

Observation: More context from a higher window size often leads to more coherent sequences. Lower temperature makes the output more predictable and structured, while higher temperature adds variety and sometimes nonsensical choices.

Encourage students to try several combinations. The point is for them to discover that some settings yield better, more coherent text than others.

File: **03_interactive.py**

Purpose: Students choose **output_length**, **window_size**, **temperature**, and a starting word **interactively**. The **default** settings are intentionally set to be **suboptimal** (small **window_size**, high **temperature**) so students see **chaotic** text first, motivating them to **improve** the parameters.

- **First Run (Defaults):**

- **output_length** = 30
- **window_size** = 1
- **temperature** = 1.8
- Suggested starting words: "On", "I", "The" (or **any** word in the data).

What Students Will Observe: Output can feel **random** or **chaotic**, demonstrating the **need** for better parameters. Explains the concept of "meh" defaults, reinforcing that small **window_size** means minimal context, while high **temperature** pushes the output toward unpredictability.

- **Suggested Tweaks for More Coherence:**

1. **window_size** = 5 or 6

- Greater context → more coherent, thematically consistent text.

2. **temperature** ≈ 0.7

- Balances creativity with predictability.
- Lower values (< 0.5) become more repetitive and predictable; higher values (> 1.2) get chaotic.

3. **output_length** = 30

- Already a good midpoint. Students can raise it (up to 50) for more extended text, but should keep in mind a **larger window_size** often helps maintain coherence for longer outputs.

- **Encourage Variety:**

- Try a higher temperature (like **2.0**) with a small window size to see more playful but less coherent text.
- Experiment with different starting words and notice how the initial prompt guides the generated text.

Observation: With the "better" settings, students can produce **fluent, story-like** text that still has some creative variety but feels **less random** than the defaults.

General Recommendations

- **Encourage Experimentation:** Remind students that there are no right or wrong answers. The point is to try different values and observe changes in coherence and style.
- **Facilitator Guidance:** If students struggle to find good results, suggest `window_size=6` and `temperature=0.7` in `02_markov_improved.py`, and in `03_interactive.py`, try `output_length=30`, `window_size=6`, `temperature=0.7` (the defaults), and let a default starting word be chosen.
- **Explain Trade-Offs:**
 - Larger `window_size` often means more coherent text but you need enough output length to see it.
 - Lower `temperature` makes text more stable but possibly repetitive; higher `temperature` adds creativity but can lead to nonsense.
- **Dataset Constraints:**

Remind students that the dataset is small and handmade. Real AI models like those behind ChatGPT use huge datasets. This tiny model will never produce perfect text, but it shows the core idea of context and probability shaping the output.

Step-by-Step Instructions

1. Introduce the Concept (2 minutes):

Facilitator Says:

"We've discussed how AI 'learns' by guessing and adjusting. Now you'll see a tiny example. We'll start from random nonsense, then use previous words as hints (Markov chains), and finally tweak how the model picks words (temperature) to produce more coherent output. Let's see how this evolves step-by-step."

2. Run `01_basic_predictor.py` (3 minutes):

Facilitator Says:

"Open `01_basic_predictor.py`. This code picks words at random with no context, just to show you the baseline. Click 'Run' and observe the output. Now open `data.txt` to view the training data for the code. Are there any phrases you recognize? Can you map the random output to what's seen in the training data? The system creates nothing new, it only draws from what has been provided to it!"

What Students See: Nonsensical text—pure randomness.

Facilitator Explains:

"This is our starting point: no learning, no pattern, no understanding. Just random words from the dataset."

3. Discuss Results (2 minutes):

Facilitator Says:

"See how random it is? There's no pattern because we're not considering what came before. This illustrates why context matters and why adjustments are needed to get something that feels more like 'real' text."

4. Move to `02_markov_improved.py` (5–10 minutes):

Facilitator Says:

"Now open `02_markov_improved.py`. Here, we introduce two key concepts: 'window_size' and 'temperature'."

- **Markov Chain (window_size):**

"A Markov chain picks the next word based on the previous words. `window_size` controls how many previous words we consider. Start with `window_size=1` (the default) and run it. Compare to the random output—notice any improvement?"

"Now try `window_size=3` by entering '3' when prompted. With more context, the model often picks a more suitable next word. Does it sound more coherent? Keep experimenting!"

- **Temperature:**

"Next, change `temperature`. Lower temperature (e.g., 0.5) makes the model more predictable—like it's choosing the most common word that fits. Higher temperature (e.g., 1.0 or more) introduces variety but can lead to stranger sentences."

Facilitator Encourages Experimentation:

"After seeing the default output, try `window_size=2` and `temperature=0.5`. Check if the text feels more natural. Adjust these values further to see how output changes. Explore `window_size=6` or try temperatures from `0.7-1.0` to see the differences."

5. Check Understanding (2 minutes):

Facilitator Asks:

"What happened when you increased the `window_size`? Did it become more coherent? What about lowering temperature—did it feel more predictable? This shows how adding context and adjusting probability can guide the model toward better results."

6. Move to `03_interactive.py` (5–10 minutes):

Facilitator Says:

"Open `03_interactive.py`. Now you're fully in control. You choose how many words to generate, the `window_size`, the temperature, and even the starting word. The model will continue from what you start with."

Facilitator Suggestions:

"Try using the settings you found effective in `02_markov_improved.py`. For example, `window_size=6`, `temperature=0.7-1.0`, and `output_length=30`. Start with a suggested word like 'The' or 'On' and see how it goes."

"If you pick a different starting word, the style might change. A higher temperature might make it more creative, while a lower temperature might keep it consistent."

7. Encourage Further Exploration (2–3 minutes):

Facilitator Says:

"Play with different parameters. Try higher `window_size` for even more context, or a higher temperature for wilder output. Notice how your starting word influences the topic of the generated text. Experiment until you find a combination you like."

8. Discussion and Wrap-Up (3–5 minutes):

Facilitator Summarizes:

"You started from random guesses and ended up guiding the model to produce more coherent text by adding context (Markov chain) and adjusting how it picks words (temperature). Real AI models do this at a

massive scale, using huge datasets and intricate training methods. Here, you've seen the core idea: guess, check, adjust."

"This simple demo helps you understand that what feels like 'intelligence' is really about patterns, context, and careful parameter tuning. If you've got the hang of this, you've taken your first steps into understanding how tools like ChatGPT work under the hood."

Key Takeaways

In real life, companies like OpenAI, Google, Meta, and Anthropic gather enormous amounts of data from the internet—text from books, articles, forums, and more—and use huge computer servers to train their models. They run through this guess-and-check process billions of times. That's why these models can cost millions of dollars to train. It's also why smaller labs often can't do this at home: you need lots of computers and tons of electricity!

- **AI as Patterns, Not Magic:** AI isn't alive or thinking. It's a pattern-recognition system that predicts what comes next based on statistical probabilities.
- **Math and Scale:** Underneath the hood, AI is math-driven. Big models run trillions of guess-check-adjust steps on huge datasets, using powerful computers to speed up the process.
- **No True Understanding:** AI doesn't "know" what words mean or what images actually look like. It identifies patterns and probabilities rather than understanding concepts.
- **Human Feedback is Crucial:** Adjustments guided by human input (fine-tuning) make AI more useful, coherent, and aligned with what we need it to do.
- **Core Concept—Guess, Check, Learn:** Whether it's labeling images or predicting words, AI starts clueless, guesses, checks against real data, and learns from mistakes. Over time, this iterative process transforms random guesses into outputs that feel remarkably intelligent.

Ultimately, what appears as intelligence is the result of careful pattern analysis, massive amounts of data, and continuous refinement. The "magic" is just math, practice, and scale—over and over until it feels like understanding.

Additional Learning Resources

Check out the resources below to keep learning about AI!

- [From Padawan to Jedi Master to Whatever Yoda Was: A Curated List of AI Learning Resources](#)
- [Google's "A Beginner's Guide to Machine Learning"](#)
- [Non-Technical AI Concepts Explainer Series](#)