MANDY HATHAWAY

Probabilities and Confidence

Instructor Guide · AI and Machine Learning Curriculum

Objectives

• Understand and explain the concept of probability
• Understand the concept of confidence
• Connect both concepts to how AI models work
• Build a simple machine learning model in Teachable Machine

Resources

• Teachable Machine: teachablemachine.withgoogle.com
• Companion slideshow (see link in course materials)
• No additional software required

Description

In this lesson, students get a real look inside how AI works. They learn that every AI model, regardless of its purpose, relies on the same two foundational concepts: probability and confidence. The lesson builds understanding through discussion, worked examples, and hands-on exploration with Google's Teachable Machine, where students build a simple image classifier and observe its predictions and confidence scores in real time.

Overview

- Phase 1: Motivation
- Phase 2: Probability, what it is and how to calculate it
- Phase 3: Confidence, how AI expresses certainty
- Phase 4: Explore with Teachable Machine

Instructor Note: *This lesson works well as an introduction before students begin building their own models. The concepts here recur throughout the AI curriculum and students who understand probability and confidence will have a much easier time interpreting model output later on.*

Phase 1: Motivation

Open with a discussion to get students thinking about AI before introducing any new concepts. The goal is to surface what they already know and create curiosity about what is happening under the hood.

Discussion Questions

- What kinds of things can AI do? (Prompt students to generate a list.)
- Have you used any AI tools recently? What did they do?
- Do you think all AI systems work differently, or do they share something in common?

Key Point to Land

Students often list very different AI applications: speech recognition, image generation, movie recommendations, self-driving cars. The surprising insight is that all of these systems, no matter how different they look, rely on the same underlying mathematics. Two of the most important concepts are ones students have likely already encountered in math class: probability. Confidence, as we will explore, is a related idea that AI systems use to express how certain they are about their predictions.

Phase 2: Probability

Goals and Learning Objectives

- Understand probability as a measure of how likely something is to happen
- Calculate simple probabilities and express them as fractions, decimals, and percentages
- Connect the concept of probability to how AI models make predictions

Discussion

- What is probability? Does anyone know?
- Where have you seen probability used before?
- What does probability have to do with AI?

Explanation

Probability is a measure of how likely something is to happen. It can be expressed as a fraction, a decimal, or a percentage. To calculate a simple probability, divide the number of outcomes you are interested in by the total number of possible outcomes.

Worked Example 1: The Spinning Wheel

Present students with a spinning wheel that has 8 equal sections: 4 labeled "Win" and 4 labeled "Lose." Walk through the calculation together:

- Count the winning outcomes: 4 sections show "Win," so the numerator is 4.
- Count all possible outcomes: there are 8 sections total, so the denominator is 8.
- Write the fraction: 4/8, which reduces to 1/2.
- Express in other forms: 1/2 = 0.5 = 50%.

- Conclusion: there is a 50% chance of winning on this wheel.

Worked Example 2: The Six-Sided Die

Ask students: what is the probability of rolling an even number on a standard 6-sided die? Walk through the same process:

- Even numbers on a die: 2, 4, and 6. That gives us 3 favorable outcomes.
- Total possible outcomes: 6 sides.
- Fraction: 3/6, which reduces to 1/2.
- Other forms: 0.5 or 50%.

Bonus Challenge

> *You have 3 pens in your backpack: 2 red and 1 blue. If you reach in without looking and pull one out at random, what is the probability that it is red? (Answer: 2/3, approximately 0.67 or 67%.)*

Connecting Probability to AI

Every output that AI produces is a prediction. When a streaming platform recommends a show, when a chatbot responds to a question, when an image classifier identifies an object in a photo: all of these are the model calculating which outcome has the highest probability of being correct, given the data it was trained on. AI is not guessing randomly. It is doing very sophisticated probability calculations at enormous scale.

- What kinds of things does AI predict? (Shopping preferences, weather, medical diagnoses.)
- Can AI probability calculations be wrong? What might cause that?
- What would happen if a model was trained on bad or biased data?

Training Data and Garbage In, Garbage Out

The probabilities a model learns are only as good as the data it was trained on. If a model is trained mostly on images of red apples, it may struggle with green ones. This is not a flaw in the math. The model simply never saw a green apple during training. This is sometimes called the garbage in, garbage out problem: a model trained on poor, incomplete, or biased data will produce poor, incomplete, or biased predictions, often with high confidence. This is one of the central challenges in AI development and a key reason why data collection and curation matter as much as the model itself.

Phase 3: Confidence

Goals and Learning Objectives

- Understand confidence as a measure of how certain a model is about its prediction
- Recognize confidence scores in real AI output
- Understand why confidence matters when evaluating AI output

Discussion

- What does the word "confidence" mean to you?
- If I asked whether you are confident that you are in class right now, what would you say?
- Can you think of a time you were confident about something and turned out to be wrong?

Explanation

Self-confidence is a feeling of certainty in your own abilities. In AI, confidence works the same way: it is a measure of how certain the model is that its prediction is correct. AI systems typically express confidence as a percentage on a scale from 0 to 100. A prediction with 95% confidence means the model is very sure of its answer. A prediction with 52% confidence means it is barely more certain than a coin flip.

Reading Confidence Scores

Show students examples of AI confidence output. A good pair of examples to contrast:

- An image classifier shows "Metal: 47%, Not Metal: 53%." The model is barely leaning one way and is not confident at all.
- The same classifier shows "Not Marshmallow: 100%." The model is completely certain. This tells us the training data for this category was strong and clear.

> ***Key distinction:*** *In many classification models, including Teachable Machine, the confidence score and the output probability are the same number. What is important for students to understand is that high confidence does not mean the model is actually correct. If a student holds up an object the model has never seen before, it may still report high confidence, because it can only compare the input to what it was trained on. Confidence reflects certainty within the model's experience, not certainty about the real world.*

Discussion

- Would you trust a model with 51% confidence as much as one with 99% confidence? Why not?
- What could cause a model to have low confidence even on a simple object?
- How might confidence scores be useful when AI is used in high-stakes situations like medicine?

Phase 4: Explore with Teachable Machine

Goals and Learning Objectives

- Build a simple image classification model using Teachable Machine

- Observe the model making live predictions and generating confidence scores
- Experiment to understand what affects confidence

> ***Instructor Note:*** *Effectiveness is not the goal here. Small datasets of 20 to 50 webcam images per class are fine. The focus is on watching the model predict and noticing the confidence scores, not on building an accurate classifier.*

What Happens During Training

When a model trains, it is looking at every example in the dataset and learning the probabilities associated with each class. For each image, it adjusts its internal settings to make it slightly better at distinguishing one class from another. By the time training is complete, the model has built a mathematical picture of what each class tends to look like, expressed as probabilities across thousands of features. When a new image is shown during evaluation, the model compares it to everything it learned and reports the class it finds most probable, along with how confident it is in that match.

Walkthrough

Step 1: Open Teachable Machine

Go to teachablemachine.withgoogle.com and select "Image Project," then "Standard image model."

Step 2: Name your classes

Rename each class to the object you plan to train on (for example, "Apple" and "Orange"). Clear class names make the output easier to interpret later.

Step 3: Collect training images

Click the webcam button for each class. Hold the object fully in frame, keep your face out of the shot if possible, and slowly rotate the object to show different angles. Press "Hold to Record" and capture 20 to 50 images per class. More angles and variety will generally produce better results.

Step 4: Train the model

Click the "Train Model" button and wait for training to complete. This typically takes under a minute for small datasets.

Step 5: Evaluate and observe

Once training is complete, turn on the webcam preview in the Output panel. Hold up each object and watch the model predict in real time. Direct students to watch both the prediction label and the confidence bars.

Final Discussion

After students have had time to experiment, bring the class together and work through these questions:

- Where are the model's predictions displayed? Where is the confidence score?
- Is the model predicting correctly? How confident is it?
- Can you make the confidence change? What did you do to cause it to change?

- What does it mean when the confidence drops? Why does that happen?
- Now that you understand probability and confidence, how would you explain AI to someone who has never heard of it?

Bringing It Together

Close the lesson by revisiting the key ideas. Every AI output is a prediction. Every prediction is based on probability. And every prediction comes with a confidence score that tells us how certain the model is. These three ideas connect everything students will encounter as they build more complex models throughout this curriculum.

- How does probability relate to AI?
- How does confidence relate to AI?
- What is the relationship between the two?
- Has this lesson changed how you think about or explain AI?

© 2025 Mandy Hathaway · mandyhathaway.com