**Read-me-once training script (6–8 minutes)**

Read in your **natural teaching voice**. Pause at punctuation. If a section feels unlike you, skip it—authenticity beats quantity.

**1) Warm-up & vowels (30–40s)**

“Hello, I’m Anirban. This is a short reference recording for my presentation voice.  
A, E, I, O, U: *ah, eh, ee, aw, oo*. fAther, bAd, sEEk, sAW, bOOk  
‘Data’ as in *day-tuh*, ‘process’ as in *pro-cess*, ‘project’ as in *proj-ect*.  
Today, I’ll speak the way I teach: calm, clear, and a little curious.”

**2) Numbers, dates, times, currencies (40–60s)**

“Zero, one, two, three, four, five, six, seven, eight, nine.  
Ten, twenty, thirty, forty, fifty, sixty, seventy, eighty, ninety, one hundred.  
Two thousand twenty-five; 17 September 2025.  
Times: 9:00 a.m., 12:30 p.m., 5:45 in the evening.  
Values: 0.001, 3.14, 1e-3 1 times 10 to the minus 3, minus 7.5.  
Money: A$199, $1.2 million.  
Percentages: 0.5%, 7%, 99.9%.”

**3) Acronyms & code-adjacent (40–60s)**

“AI, ML, DL, NLP, TTS, GPU, TPU, API, SDK, IDE.  
CSV, JSON, HTML, SQL, PyTorch, TensorFlow, Keras, NumPy, pandas, Matplotlib.  
ETA, KPI, MVP, ROC-AUC.  
‘pip install’, ‘conda create’, ‘git commit’, ‘version control’, ‘vectorized operations’.”

**4) Core DS/AI lecture tone (90–120s)**

“In supervised learning, we map inputs to outputs using labeled data.  
For regression, we predict a continuous value like house price; for classification, a discrete label like fraud or not.  
We split our dataset into training, validation, and test sets. We standardize features, handle missing values, and watch for data leakage.  
Gradient descent iteratively updates parameters to minimize a loss function.  
With mini-batches, Adam adapts learning rates per parameter and typically converges faster on noisy objectives.  
We evaluate using mean squared error for regression and F1-score for imbalanced classification.  
In deep learning, we stack layers with nonlinear activations.  
We monitor validation loss, use early stopping to prevent overfitting, and plot learning curves to diagnose under- or over-fitting.  
For interpretability, we examine feature importances, SHAP values, and partial dependence plots.  
For deployment, we containerize models, log predictions, and track data drift.”

**5) Technical detail & math voice (60–90s)**

“The cost function J of linear regression is one over two-m times the sum of squared errors.  
The gradient is X-transpose times the residuals, divided by m.  
In logistic regression, we use the sigmoid function, one over one plus e to the minus z.  
Regularization adds a penalty term; L-two shrinks weights towards zero to reduce variance.  
For time series, we check stationarity, seasonality, and autocorrelation.  
For neural nets, we initialize with Xavier or He; we set learning rates carefully and sometimes apply cosine decay.”

**6) Domain examples you’ll actually narrate (60–90s)**

“This chart shows Australian emissions trending downward since 2005, with sharper declines after 2020.  
The forecast suggests we’re on track for the 2030 target if current policies hold.  
On the right, the confusion matrix shows strong recall but precision needs improvement—false positives are still costly.  
Here, the ablation study demonstrates that removing the text normalization step drops accuracy by three percentage points.  
In our abalone age classification, adding batch normalization stabilized training and improved ROC-AUC.”

**7) Classroom/explainer tone shifts (60–90s)**

“Quick recap: start with a clean dataset, split carefully, choose a baseline, and iterate.  
Common pitfall: leaking target information into features.  
If your validation score is suspiciously high, double-check the split.  
Question I often get: should we use Adam or SGD with momentum?  
Short answer: Adam for speed and noisy gradients, SGD-momentum for ultimate generalization when you can tune it.”

**8) Emotion & emphasis samples (30–45s)**

**Neutral:** “Let’s review the results.”  
**Encouraging:** “Great progress—your loss is trending down!”  
**Concerned:** “Notice the variance: the model may be overfitting.”  
**Excited:** “This improvement halves inference time without sacrificing accuracy.”

**9) Names, places, and proper nouns (20–30s)**

“Sydney, Melbourne, Canberra, Brisbane, Adelaide, Perth.  
UNSW, Anodiam, Gypsee, Syan Gypsee.”

**10) Phoneme-rich sentences (30–40s)**

Read a handful of phonetically balanced lines to help consonants/vowels:

* “The quick brown fox jumps over the lazy dog.”
* “She sells sea shells by the sea shore.”
* “A noisy voice annoys an oyster.”  
  (Phonetically balanced lists like the *Harvard Sentences* are commonly used for coverage.) [Grateful Labs](https://gratefulabs.com/free-resources/phrases-to-speak-for-accurate-voice-cloning-harvard-sentences?utm_source=chatgpt.com)

**11) Closing (10–15s)**

“This concludes my reference recording. I’ll speak naturally, as I do in lectures, with clear articulation and steady pace.”