

Generative Adversarial Networks: Learning to Sample by Playing a Game

Lecture Notes Consolidated with Claude AI — CS6140 Machine Learning

Instructor's lecture plan for the note of the same title (`GAN_claude_0929.pdf`). Section numbers (§) refer to that note.

This note supports the second Week 10 lecture and **HW5 Problem 4** (and the optional Problem 7, DCGAN). It assumes Week 9's networks and backprop, logistic regression and its binary cross-entropy loss (Weeks 3–4), HW4's Gaussian discriminant analysis and GMM, and the first Week 10 lecture (**VAE** note: KL divergence, the reparameterization $x = \mu + \sigma\epsilon$). The demo notebook `gan_lecture_demo.ipynb` checks every number here and produces every figure.

Lecture plan (100 minutes, mostly on the board)

Time	On the board	Worked example / demo
0–10	Two generative models so far, both with a density: GMM (exact likelihood), VAE (a lower bound). Do we need a density to <i>generate</i> ? No, only a sampler. Counterfeiter vs. detective. Draw the big picture: two networks, the 3-step loop; leave it up all lecture. (§1.1, §1)	The overview diagram (§1).
10–22	A generator is a sampler: $x = G(z)$. Linear G = the VAE's $\mu + \sigma\epsilon$; nonlinear G bends noise into any shape, but has no density formula. (§2)	$G(z) = z + 2 \tanh(3z)$ makes two bumps (figure).
22–40	The detective is a logistic regression: real $\rightarrow 1$, fake $\rightarrow 0$, BCE. The best possible detective: $D^* = p/(p+q)$. (§3)	3-point example; $\mathcal{N}(0,1)$ vs. $\mathcal{N}(2,1)$ gives $D^* = \sigma(2 - 2x)$, HW4's LDA.
40–55	What the counterfeiter minimizes when the detective is perfect: $-\log 4 + 2 \text{JSD}$. Equilibrium: $D \equiv \frac{1}{2}$, D's loss = $\log 4$. (§4)	JSD = 0.216 on the 3-point example.
55–65	Training: alternate a D step and a G step; backprop from D's output into G; why fakes are detached in the D step. (§5)	Draw the two-network computation graph.
65–78	Saturation: the original G loss has almost no gradient exactly when G is bad; the non-saturating fix. (§6)	Gradient table ($D = 0.01$: 99×); demo figure.
78–90	Common failure modes: mode collapse, vanishing gradients, losses that don't go down. (§7)	Two-point collapse example; 3-mode demo figure.
90–97	Demo (5 min): a full 1-D GAN, snapshots, D's loss $\rightarrow \log 4$.	Notebook §6.
97–100	Three ways to build a generative model; HW5 Problem 4 map. (§9, §10)	