Natural Language Engineering (2019), **25**, pp. 323–329
doi:10.1017/S1351324919000020

CAMBRIDGE UNIVERSITY PRESS

EMERGING TRENDS

Emerging trends: GANs vs. good enough

Kenneth Ward Church*

Baidu USA, Sunnyvale, CA, USA
*Corresponding author. Email: kenneth.ward.church@gmail.com

(Received 19 January 2019; revised 11 January 2019; accepted 11 January 2019)

Abstract
Generative Adversarial Networks (GANs) are hot. Murphy's Law, it is prudent to be paranoid. Best not to design for the average case. There is a long tradition of designing for the hundred-year flood (and five 9s reliability). What is good enough? Historically, the market hasn't been willing to pay for five 9s. Hard to justify upfront costs for future benefits that will only payoff under unlikely scenarios, and might not work when needed. If the market isn't willing to pay for five 9s, can we afford to design for the worst case?

Keywords: GANs; Generative Adversarial Networks; Five 9s; Defects per Million; Good enough; POTS

1. When pigs fly

Kolter and Madry gave a nice tutorial[a] at NeurIPS-2018 (formally NIPS) on General Adversarial Networks (GANs).[b] The tutorial starts by praising Imagenet[c] as a "home run." Deep nets are obviously doing extremely well. But "extremely well" may not be "good enough." The problem is that it is too easy to "attack" deep nets. Kolter and Madry credit (Szegedy 2013; Biggio 2013) and many others for identifying various vulnerabilities.

Their example of the vulnerabilities starts with a simple picture of a pig. Their net has no trouble classifying the picture as a pig. But then they attack the net by adding a tiny amount of "noise" to change the prediction from "pig" to "airliner." Their concern is: "the predictions are extremely good on average, but also, they are extremely brittle." At 16:38 in the video,[d] the audience has a good laugh after the punch line: "Machine Learning technology is truly magical; it can make pigs fly."

More seriously, these attacks demonstrate that it is too easy for an adversary to change the classification from something sensible to something completely different, perhaps with disastrous consequences, as suggested in (Papernot 2013), where an adversary causes a stop sign to be misclassified as a yield sign. A malicious adversary could cause car accidents by distorting stop signs in a way that would be hard for any of us to notice, but could cause lots of self-driving cars to misbehave at the worst possible time.

These attacks make for great theater, but there are additional unexpected benefits, as discussed at 1:40:30 in the video (and slides 45-46).[e] As shown in (Tsipras 2018), the losses are more intuitive and semantically meaningful with adversarial training.

[a] https://adversarial-ml-tutorial.org/
[b] https://en.wikipedia.org/wiki/Generative_adversarial_network
[c] http://www.image-net.org/
[d] https://www.facebook.com/nipsfoundation/videos/281865415800485/
[e] https://adversarial-ml-tutorial.org/adversarial_ml_slides_parts_1_4.pdf

Engineers have been talking (joking?) about Murphy's Law for a long time, well before the relatively recent interest in GANs. There is a long tradition of building bridges for the hundred-year flood. It is widely accepted that engineers should pay considerable attention to downside risks.

2. How bad is Murphy?

On the other hand, engineers do what engineers do. They like to build things (and sometimes overbuild things). Obviously, we shouldn't build a bridge for the average year, but should we build for the hundred-year flood, or should we build for a flood of biblical proportions, or something worse than that such as the worst case?

The lock on the front door of my house serves a useful purpose, but I understand that while it might deter an amateur thief, it won't stop a professional. It certainly won't stop a military invasion, or even the police. I could build a big beautiful border wall around my house, but why would I want to? It can be hard to justify upfront costs for future benefits, especially when the payoffs are uncertain and unlikely.

Similar concerns apply to robustness methods: GANs, five 9s, etc. When designing for robustness, we want to make sure that the benefits outweigh the costs.

Is Murphy's Law *always* out to get us? Or perhaps, Murphy's Law is only a rule of thumb. Perhaps, it often applies, but not always.

The larger question is the threat model. Training for the average case is like disrespecting Murphy. Like the Jim Croce song, you don't mess around with Murphy.[f]

You don't tug on superman's cape
You don't spit into the wind
You don't pull the mask off that old lone ranger
And you don't mess around with Jim

But how much respect should we have for Murphy? Is he like Leroy Brown, merely "the baddest man in the whole damned town, badder than old King Kong, and meaner than a junkyard dog,"[g] or does Murphy deserve more respect than that? Is Murphy merely

1. the baddest man in the whole damn town, or
2. the baddest man there ever was (or ever will be)

Is Murphy an all-knowing deity? Does he know what cards you have, and exactly which fibers to cut to do maximum damage to your network? Does Murphy play dice,[h] or can he predict the future?

3. Threat models

What are we worried about? Acts of nature? Hurricanes? Earthquakes? Cosmic rays? Determined adversaries? All-knowing deities?

Traditionally, we tended to address threats with special purpose solutions designed specifically for each particular threat. Back in the 1970s, when I was a student at MIT, bored teenagers were the threat. Few people had access to computers back then. And even fewer had modems at home.

[f] https://www.youtube.com/watch?v=odkIEDi2x0g
[g] https://www.youtube.com/watch?v=QvwDohEEQ1E
[h] http://www.hawking.org.uk/does-god-play-dice.html

https://doi.org/10.1017/S1351324919000020 Published online by Cambridge University Press

These were the early days of networking.[i] Most of us used fancy so-called glass terminals[j] in a terminal room in the lab, but a few people (mostly in the military) could log into the system from home, connected over dial-up telephone lines via slow (300 baud) modems on "portable" (luggable) terminals that used expensive silver paper.[k]

In those days, we had no security (and no privacy). You could switch your glass terminal to any port you wanted, so it was as easy to see your neighbor's screen buffer as your own. There were no passwords. User ids and accounts didn't mean much. You could log in as yourself (or anyone else). Similar comments apply to logging out. You could log yourself out, or you could log out anyone you felt like logging out.

The security model was like a small town. Don't do anything you wouldn't want everyone to know about. You should assume that everyone was watching you (because they probably were).

The small-town security model had its advantages. Minsky once wrote a long flame to the director of the AI lab. This was a clever move to get the last word on some long-forgotten debate. Minsky knew that sending the flame to the director wouldn't bother him because the director didn't read his email, but Minsky also knew that the director was the only one in the lab that didn't read the director's email, so he could have his say with everyone else. None of us, of course, could respond to the flame because none of us would admit to reading email that we shouldn't have read.

Eventually, the small-town model became unworkable as the field became successful, and the community grew from a small town into a big city. Even in its heyday, the small-town model worked better among academics working in a shared space than teenage kids (in remote locations). Some teenagers logged into the system remotely via dial-up modems and behaved badly (deleting PhD theses, among other things). When we complained to their parents (largely in the military), the parents were unsympathetic. If you can't secure your system against our kids, how are you going to defend against a real threat? It took them a few years/decades to figure out that hacker kids were a real threat to be taken extremely seriously.

Our first attempt to deal with their kids felt good, but did little to solve the problem. We'd send their slow connection a block of line feeds, and they could watch their expensive paper advance painfully slowly (at 300 baud). That would force them to log out, but it didn't fix them problem since they often logged back in and caused even more childish mayhem.

The next attempt, PW, was more effective. The solution was to think like a kid. After logging out the offending target, PW printed "password please" on what remained of the network connection. The target would be given three tries to come up with the right answer, but there was no right answer (because we didn't have passwords). These kids rarely came back after they were "caught" by PW.

Later on, the threat model evolved from bored teenagers to organized crime. The solution in that case was to think like a mobster, and "follow the money." The mob does what they do to make money. Some of the mob businesses were scamming my employer (the phone company). If we could find them quickly, we could shut them down before they could do too much damage to us (and before they became profitable to the mob). Although the mob did what they could do to make it hard for us to find their businesses, the customers of these businesses knew where they were. The solution was to follow the frequent flyers of such scams.

The field became pretty good at dealing with old threats, using special case solutions such as the ones mentioned above. Over time, new threats emerged (spammers, fake news, digital warfare, etc.), and eventually, the good guys came up with new solutions to new threats. It would be nice if there was a general purpose solution to all possible threats, but that's probably too much to

[i] https://en.wikipedia.org/wiki/ARPANET
[j] https://en.wikipedia.org/wiki/Computer_terminal
[k] https://en.wikipedia.org/wiki/Silent_700

https://doi.org/10.1017/S1351324919000020 Published online by Cambridge University Press

ask for. We would all love to have a magic wand that can cure all ills, including ills that haven't been discovered/invented yet. Perhaps GANs are such a magic wand, but probably not.

4. Defects per million

When I started my first job at AT&T Bell Labs in 1983, there was a huge emphasis on reliability. The standard at the time was ambitious, perhaps too ambitious, but far less than worst case. The telephone standard was referred to as defects per million or five 9s. That is, the telephone system was designed to work 99.999% of the time. In fact, it probably only worked 99.99% of the time (four 9s), but even so, that was probably more reliability than the market was willing to pay for. If the telephone company hadn't been a monopoly, it is unlikely that the company could have found the money to pay for so much reliability.

Some of the more expensive preventive measures were:

1. Avoid depending on the power grid (by overbuilding the power grid with a separate power system for the entire network including the customer's handset).
2. Avoid depending on long-distance cables using redundancy such as SONET self-healing rings (SHR)[l] (Wu 1992).

5. Backup power

They felt it was necessary to overbuild the power grid because the standard grid was too far from the five 9s standard. A single blackout lasted 13 hours.[m] To make five 9s, we can afford only 5.26 minutes of down time per year.[n] Thirteen hours of downtime would blow the reliability budget for about 150 years. And there have been many blackouts since then.

But overbuilding the power grid is extremely expensive. Approximately 20% of the cost of a data center goes to batteries and generators. It isn't clear that people would be willing to pay that cost, if they had a choice. How important is it to you that you can access your backup photos during a power failure? Most people don't look at their old photos much. Are you willing to pay a 20% premium for the privilege of looking at these photos during a power failure? Most people pay this premium because they don't have a choice. But even though the clouds will probably stay up during a power failure, most of the edge will not. Few people install backup batteries on their Wi-Fi routers in their homes...

If you don't have backup batteries on your Wi-Fi router, are you concerned? Do you expect your home network to work during a power failure? Are you willing to pay what it costs for 99.999% reliability? Would you be willing to pay for more than that? Do we need your home network to work during a military attack? Do we need it to work in all cases (including the worst case)?

6. Avoiding single points of failure

There are a number of methods like SONET rings and error correcting (ECC) memory[o] that protect against a single fault (fiber cut/bit flip), but not against a double fault. Such measures do not protect against a determined adversary, especially an adversary that knows exactly where to cut your fibers.

When estimating the benefit of such measures, it is common to assume independence (and Poisson processes) such as this:[p]

[l] https://classes.engineering.wustl.edu/2012/fall/ese571/SONETpaper1(WU).pdf
[m] https://en.wikipedia.org/wiki/Northeast_blackout_of_1965
[n] https://en.wikipedia.org/wiki/High_availability
[o] https://en.wikipedia.org/wiki/ECC_memory
[p] https://en.wikipedia.org/wiki/Cosmic_ray#Effect_on_electronics

https://doi.org/10.1017/S1351324919000020 Published online by Cambridge University Press

> Studies by IBM in the 1990s suggest that computers typically experience about one cosmic-ray-induced error per 256 megabytes of RAM per month.

Discussions of the benefits of error correcting memory can be hard to follow (because of unstated assumptions), but in addition to Poisson assumptions, there are also suggestions that cosmic rays can cause (non-Poisson) bursts of bit flips.[q]

The best laid plans often don't work out, but other things sometimes work out better than expected. Although SONET rings failed in New York on September 11, 2001, the mobile network took a licking, but it kept on ticking.[r] Unlike SONET rings, the mobile network was never designed for reliability (it was designed for mobility), but the mobile network worked when SONET rings failed.

What happened? SONET rings failed when building 7 collapsed, causing multiple fiber cuts between the Verizon switch and Lower Manhattan.[s] Much of the load transitioned to the mobile network which continued to work despite the loss of several cell towers (on both the North and South Towers of the World Trade Center), as well as damage to the nearby Verizon switch. The main challenge for the mobile network was not connectivity (thanks to cell towers in nearby New Jersey and elsewhere), but a surge in demand as news of the disaster spread around the world.[t]

7. POTS

In general, people have limited willingness to pay for disaster prevention, especially as memories fade from the last disaster. New homes these days are no longer designed with inside wire for POTS, and consequently, telephone service is no longer as reliable as it used to be.

What is POTS? POTS is plain old telephone service.[u] It doesn't do much, but what it does, it does reliably (designed for 99.999%). When the telephone company was a monopoly, the company was willing to invest huge sums to make sure that people could reach 911 in an emergency, no matter what the cost. Features were less of a priority, whether people wanted them or not, leading to an image problem and jokes such as: "We don't care. We don't have to. We're the Phone Company."[v]

These days, there are different trade-offs. People are probably not willing to pay for five 9s, let alone worst case. I remember shopping for a telephone handset with my parents (who still owned a landline phone), and discovering that it is hard to find a handset that would work during a power failure. The store had a whole aisle full of all sorts of fancy phones, but only one of them would work during a power failure. The store did its best not to sell that phone, sticking it in a corner on the bottom shelf, because that phone was the cheapest one they had, and the store probably didn't make money on that phone. Most phones offer all sorts of fancy features: voice mail, address books, cordless, etc. The only thing this phone offered was reliability (POTS). Clearly, the market prioritizes features over reliability.

8. Good enough

I first heard the term "Good Enough" when working at Microsoft. They understood that the market wasn't willing to pay for five 9s. Engineers do what they do. They like to build things that work.

[q] https://stackoverflow.com/questions/2580933/cosmic-rays-what-is-the-probability-they-will-affect-a-program
[r] https://www.youtube.com/watch?v=_NHq3Yze6s0
[s] https://www.nap.edu/read/10569/chapter/4#23
[t] https://www.nap.edu/read/10569/chapter/4#37
[u] https://en.wikipedia.org/wiki/Plain_old_telephone_service
[v] https://snltranscripts.jt.org/76/76aphonecompany.phtml

https://doi.org/10.1017/S1351324919000020 Published online by Cambridge University Press

But marketing has other priorities. They want things that are good enough to sell. Features sell. Reliability, not so much.

This became extremely clear when when a military (not the US military) came to the Microsoft executive briefing room. You might think the military would want mil spec[w] from a beltway bandit,[x] but no one wants that, not even the military. The military had figured out that consumer grade electronics were better value for money than their more obvious alternatives.

Some vendors talk a lot about security and reliability, but much of the talk is just talk. It isn't clear that more expensive solutions are more reliable than less expensive solutions. My grandmother used to say that fruits and vegetables are best when they are cheapest (because that's when they are in season). So too, cheaper consumer electronics are likely to be best (and most reliable) because they are cheapest.

That might seem counter-intuitive, but Moore's Law favors whoever has the larger market share. Consumers buy more than others (enterprise & government), and therefore, consumer electronics are better in every way (including reliability).

I had a friend that used to go to toy stores for game controllers for a military application. My friend configured used cars for target practice. The military would drive these cars around in the desert by remote control, while others would see if they could destroy them with drones. It was like a video game, except it wasn't exactly a game. I asked my friend why he went shopping in toy stores, and he gave me the same answer that I heard in the executive briefing room. Toy stores offer better value for money than mil spec.

9. Conclusion

The market isn't willing to pay for reliability, except on rare occasions. Most people (and even most enterprises) probably don't want mil spec, and stuff designed for five 9s. The best laid plans often don't work out as planned. Attempts to protect against double faults (like SONET rings, error correcting memory, RAID, etc.) introduce obvious costs. The benefits often depend on independence assumptions that may not be appropriate in practice. It can be hard to justify upfront costs for future benefits, especially when the payoffs are uncertain and unlikely.

IMHO, I have more confidence in postmortems based on empirical data than in theoretical models based on assumptions that may not be appropriate in practice. Here is an example of a sensible post mortem with reasonable conclusions (assuming the future will be like the past).[y]

> These results confirm the conventional wisdom that more interconnected networks are more reliable. However, it is interesting that the higher reliability is not a consequence of fewer faults, but of the smaller consequences of most faults. From the available data, it cannot be concluded why the number of faults is not lower in more interconnected networks, as it only contains key performance indicators and a basic description of the causes.

Of course, if Murphy should turn over a new leaf and change from merely the baddest man in the whole damn town, to become the baddest man there ever was (or ever will be), then maybe we should design for the worst case (as GANs do).

Conventional wisdom suggests that redundancy improves reliability, but of course, that doesn't follow from worst case analysis. If Murphy is truly out to get us, he can do so twice. Double faults may be less likely than single faults, but double faults are not impossible.

Adversarial training is very appealing, but we don't want to take it too far. The business case will justify a few 9s, but probably not five 9s, and certainly not worst case.

[w] https://en.wikipedia.org/wiki/United_States_Military_Standard
[x] https://en.wikipedia.org/wiki/Beltway_bandit
[y] https://ses.jrc.ec.europa.eu/sites/ses.jrc.ec.europa.eu/files/publications/1-s2.0-s0378779612001071-main.pdf

https://doi.org/10.1017/S1351324919000020 Published online by Cambridge University Press

References

Biggio B., Corona I., Maiorca D., Nelson B., Šrndić N., Laskov P., Giacinto G. and Roli F. (2013). Evasion attacks against machine learning at test time. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Prague, Czech Republic: Springer, pp. 387–402.

Papernot N., McDaniel P., Goodfellow I., Jha S., Celik Z.B. and Swami A. (2017). Practical black-box attacks against machine learning. *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*. Abu Dhabi, UAE: ACM, pp. 506–519.

Szegedy C., Zaremba W., Sutskever I., Bruna J., Erhan D., Goodfellow I. and Ferbus R. (2013). Intriguing properties of neural networks. arXiv:1312.6199.

Tsipras D., Santurkar S., Engstrom L., Turner A. and Madry A. (2018). Robustness may be at odds with accuracy. arXiv:1805.12152.

Wu T.-H. and Lau R. (1992). A class of self-healing ring architectures for SONET network applications. *IEEE Transactions on Communications* **40**(11), 1746–1756.

Cite this article: Church KW (2019). Emerging trends: GANs vs. good enough. *Natural Language Engineering* **25**, 323–329. https://10.1017/S1351324919000020

https://doi.org/10.1017/S1351324919000020 Published online by Cambridge University Press