Cheap calibration of the normalizing flow using arbitrary classifier

by Robert Drynkin, Maxim Artemev, Denis Derkach

April 2020

Content

- Assumptions
- Idea
- Implemetation details
- Pictures
- Current results
- FAQ

Assumptions

Normalizing flows

- - Long time to train
- - Unstable training
- - Have an artefacts due to continuity
- + Generative model

Binary classifier

- + Fast training
- + A lot of libraries
- + Well studied
- - Not a generative model

Can we use binary classifier to improve normalizing flow with the preservation of useful properties?

Idea

$Q(x)$ – density of the real distribution
$P(x)$ – density of the distribution given by nf
$clf(x)$ – prediction of the perfect discriminator (binary classifier which trained to distinguish samples from Q and P)

We can express $Q(x)$ using only $P(x)$ and $clf(x)$

$$clf - \text{perfect} \Rightarrow clf(x) = \frac{Q(x)}{P(x) + Q(x)} \Rightarrow \frac{clf(x)}{1 - clf(x)} = \frac{Q(x)}{P(x)}$$

$$Q(x) = P(x)\frac{clf(x)}{1 - clf(x)}$$

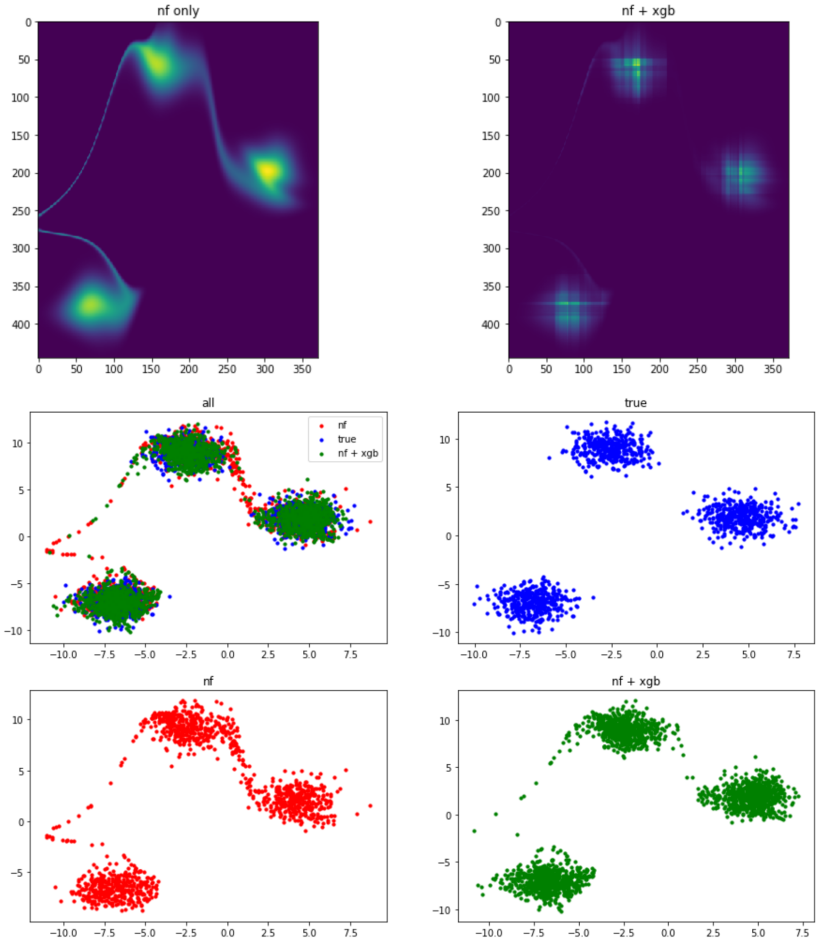
Implemetation details

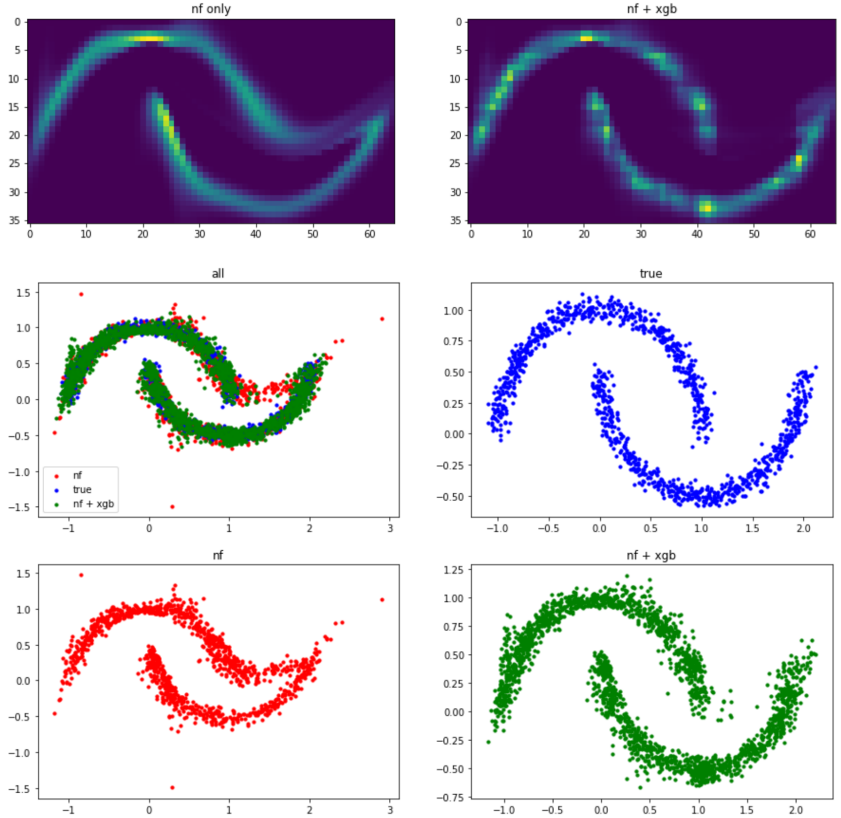
For the numerical stability:

$$\log Q(x) = \log P(x) + clf\_logit(x)$$

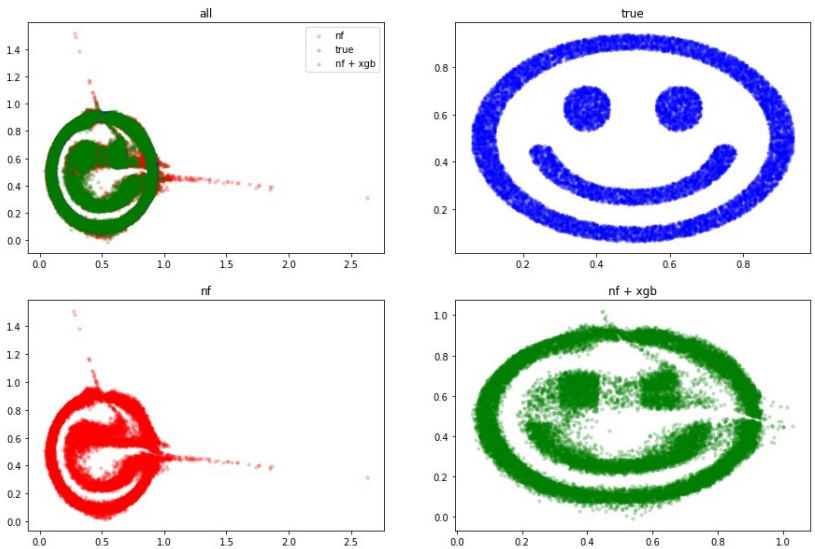
Sampling: rejection sampling from $Q(x)$ using $c \cdot P(x)$ as a major distribution. Acceptance rate equals to $1/c$, where $c = \max \exp clf\_logit(x)$

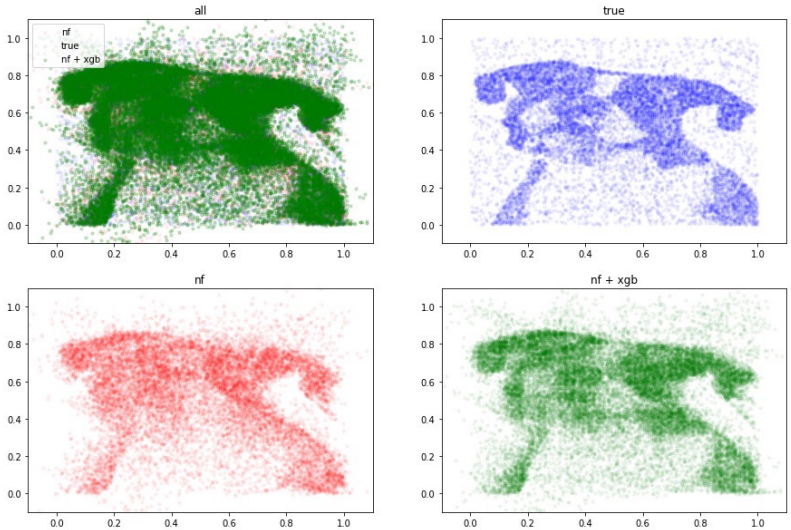
Pictures


nf only
nf + xgb
all
true
nf
nf + xgb
nf
true
nf + xgb


nf only
nf + xgb
all
true
nf
nf + xgb
nf
true
nf + xgb

Pictures


all
nf
true
nf + xgb
true
nf
nf + xgb


all
nf
true
nf + xgb
true
nf
nf + xgb

Current results

GAS dataset

- SOTA: maf(10) $\rightarrow$ 10.08 log likelihood
- Ours: maf(5) $\rightarrow$ 9.48 log likelihood $\rightarrow$ 10.23 ll after calibration by Catboost(2k trees)

Future work

- KL finetuning nf
- Conditional sampling

FAQ

- What happens if $P = Q$ (perfect nf)?
  Then $\forall x\ clf\_logit(x) = 0$, acceptance rate of rejection sampling $= 1$.
- What happens if P is trivial (e. g. uniform)?
  Then $\exists x\ clf\_logit(x) \to \infty$, acceptance rate of rejection sampling $\to 0$.

The End