Boosted Generative Models

Aditya Grover  \quad Stefano Ermon
Stanford University
{adityag, ermon}@cs.stanford.edu

ABSTRACT

We propose a new approach for using unsupervised boosting to create an ensemble of generative models, where models are trained in sequence to correct earlier mistakes. Our meta-algorithmic framework can leverage any existing base learner that permits likelihood evaluation, including recent latent variable models. Further, our approach allows the ensemble to include discriminative models trained to distinguish real data from model-generated data. We show theoretical conditions under which incorporating a new model in the ensemble will improve the fit and empirically demonstrate the effectiveness of boosting on density estimation and sample generation on synthetic and benchmark real datasets.

LEARNING IN GENERATIVE MODELS

Age of big unlabeled data!

Given \( X = \{ x_i \in \mathbb{R}^d \}_{i=1}^n \sim P \), we want to learn a generative model \( Q \) from a model family \( \mathcal{Q} \) that minimizes the KL-divergence w.r.t. \( P \)

\[
\min_{Q \in \mathcal{Q}} D_{KL}(P || Q)
\]

Small \( Q \)? Model mismatch
Large \( Q \)? Poor generalization, difficult to optimize.

SUPERVISED BOOSTING

\[
H(x) = \text{sign} \left( \sum_{t=1}^{T} a_t h_t(x) \right)
\]

When does this work? (Informally) Every classifier makes a better-than-random prediction [1].

UNSUPervised BOOSTING

No labels! Interested in a density \( q_T \) after \( T \) rounds of boosting.

- **Learning objective**
  \[
  \min_{Q \in \mathcal{Q}} D_{KL}(P || Q_T)
  \]

- **Product of experts**
  \[
  q_T = \frac{\prod_{t=0}^{T-1} h_t}{Z_T}
  \]
  \[
  Z_T = \int \prod_{t=0}^{T-1} h_t \, d x
  \]

When does this work?

At round \( t \) of boosting, define the reduction in KL divergence , \( \delta_{KL}(h_t, a_t) = D_{KL}(P || Q_{t-1}) - D_{KL}(P || Q_t) \).

**Theorem.** If \( \delta_{KL}(h_t, a_t) \geq 0 \), then \( E_{Q_{t-1}}[\log h_t] \geq E_{Q_{t-1}}[\log h_{t-1}] \) for all \( a_t \in [0, 1] \).

ALGORITHMIC FRAMEWORK

- **GenBGM**: Training ensemble of generative models after reweighting.
- **DiscBGM**: Training ensemble of discriminative models for unsupervised-as-supervised learning.

<table>
<thead>
<tr>
<th>Model</th>
<th>Test negative log-likelihood (in nats)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base model</td>
<td>4.69 ± 0.01</td>
</tr>
<tr>
<td>Add model [3]</td>
<td>4.64 ± 0.02</td>
</tr>
<tr>
<td>GenBGM</td>
<td>4.58 ± 0.10</td>
</tr>
<tr>
<td>DiscBGM</td>
<td>4.42 ± 0.01</td>
</tr>
</tbody>
</table>

Avg. test negative log-likelihoods (with std. error) for instances of mixture of Gaussians.

DENSITY ESTIMATION

GenBGM

- Start with uniform weights for training data
- Fit maximum likelihood model to data
- Repeat for \( t = 1, 2, ..., T \) rounds:
  - Reweight the training data
  - Refit model \( h_t \) to the reweighted data

DiscBGM

- Fit maximum likelihood model to training data
- Repeat for \( t = 1, 2, ..., T \) rounds:
  - Train binary classifier \( c_t \) to distinguish train data and model-generated data
  - Specify model \( h_t \propto c_t^{\alpha_t} \) \[2, 4\]

REFERENCES