The Futility of Bias-Free Learning and Search

2019

Overview:

Imagine you are on a routine grocery shopping trip and plan to buy some bananas. You know that the store carries both good and bad bananas which you must search through. There are multiple ways you can go about your search.

You could:
1) randomly pick any ten bananas available on the shelf (unbiased search)
2) Only pick those bananas that are neither underripe nor overripe (biased search)

We mathematically prove that for a machine, the proportion of good bananas retrieved in its biased search is greater than the same proportion in an unbiased search; it was given prior knowledge about tasty bananas.

We further prove that an algorothm cannot be simulatenously biased towards all types of bananas (underripe, overripe, perfect).

The Search Framework:

We cast machine learning problems as search problems and use this framework to prove facts about the nature of machine learning.

A search problem is represented as a 3-tuple (Ω, t, F):
• Ω: finite search space (all bananas on shelf)
• t: target subset of elements in the search space (yummy bananas)
• F: external information resource (your brain)

An algorothmic search comprises of a search problem, history of elements examined, and information resource evaluations. It decides how to query elements of Ω.

[As a black-box optimization algorithm samples from Ω, it produces an associated probability distribution Pi based on the search history. When a sample ωk corresponding to location k in Ω is evaluated using the external information resource F, the tuple (ωk, F(ωk)) is added to the search history.]

Key Definitions:

Bias

We define bias to be a measure of how different the performance of one distrbution is to that of a uniform distribution.

Given a:
• k-sized fixed target t
• distribution over information resources D
• search algorithm PF
• performance of uniform random sampling p

$$\mathrm{Bias}(\mathcal{D}, \boldsymbol{t}) = \mathbb{E}_{\mathcal{D}} \left| \boldsymbol{t}^\top \overline{P}_F \right| - p$$

Note: algorithmic bias has very little to do with unfortunate human biases (racial, xenophobic, etc.). Our version of algorithmic bias is strictly defined in the domain of mathematics, and human examples of it include picking the simplest solution from a set, and choosing a particular type of learning algorithm used to solve a problem.

Target Divergence

What does algorithmic bias look like? Target divergence is a geometric interpretation of it which measures the angle between the target vector t and the averaged |Ω|-length simplex vector P_f using cosine similarity

(a) $\overline{P}_{f_1} = [0, 0.2, 0.8]^\top$, $\boldsymbol{t_1} = [0, 1, 1]^\top$, and $\theta_1 \approx 31°$. While all of the probability mass in $\overline{P}_{f_1}$ lies on the target set $\boldsymbol{t_1}$, the target divergence takes value greater than 0° because $\overline{P}_{f_1}$ is not uniform.

(b) $\overline{P}_{f_2} = [0, 1, 0]^\top$, $\boldsymbol{t_2} = [1, 0, 1]^\top$, and $\theta_2 = 90°$. Since none of the non-zero probability mass in $\overline{P}_{f_2}$ aligns with their corresponding target elements in the target set $\boldsymbol{t_2}$, the target divergence is maximized at 90°.

(c) $\overline{P}_{f_3} = [0.5, 0.5, 0]^\top$, $\boldsymbol{t_3} = [1, 1, 0]^\top$, and $\theta_3 = 0°$. Since $\overline{P}_{f_3}$ places all of its probability mass uniformly on the target set, the target divergence is minimized at 0°.

Results:

Improbability of Favorable Information Resources

Let F be a randome variable such that F ~ D and let q(t F) be the expected per-query probability of success for algorithm A on search problem (Ω, t, F). Then for any q_min \in [0,1],

$$\Pr(q(t, F) \geq q_{\min}) \leq \frac{p + \mathrm{Bias}(\mathcal{D}, \boldsymbol{t})}{q_{\min}}$$

Conservation of Bias over Distributions

$$\sum_{t \in \mathcal{T}_k} \int_{\mathcal{D}} \mathrm{Bias}(\mathcal{D}, \boldsymbol{t}) \, \mathrm{d}\mathcal{D} = 0$$

Futility of Bias-Free Search

If If Bias(D, t) = 0, then

$$\Pr(\omega \in t; \mathcal{A}) = p$$

where Pr(ω \in t; A) represents the per-query probability of successfully sampling an element of t using A, marginalized over information resources F ~ D.

Authors:

Jonathan Hayase

Julius Lauw

Dominique Macias

Akshay Trikha

Julia Vendemiatti

George Montañez

Acknowledgements:

We'd like to thank our loving families and supportive friends, Professor George for his guidance, and the Walter Bradley Center for Natural and Artificial Intelligence.

References:

- Montañ'ez, G.D., Hayase, J., Lauw, J., Macias, D., Trikha, A., Vendemiatti, J. :The Futility of Bias-Free Learning and Search. arXiv e-prints arXiv:1907.06010(jul 2019)
- Montañez, G.D. :The famine of forte: Few search problems greatly favor your algorithm. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics(SMC). pp. 477–482. IEEE (2017)
- Wolpert, D.H., Macready, W.G.: No free lunch theorems for optimization. Trans.Evol. Comp1(1), 67–82 (Apr 1997)
- Mitchell, T.D.: The need for biases in learning generalizations. In: Rutgers University: CBM-TR-117 (1980)