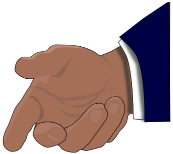
Collecting the right data:
A machine learning theory perspective on A/B testing

Wouter M. Koolen

AI020, April 24 2024

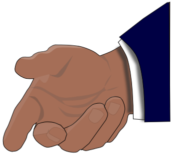
Welcome! A bit about me



Senior Researcher, Machine Learning group, Centrum Wiskunde & Informatica

Professor of Mathematical Machine Learning, University of Twente

Welcome! A bit about me



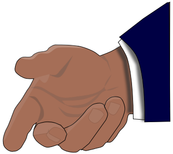
Senior Researcher, Machine Learning group, Centrum Wiskunde & Informatica

Professor of Mathematical Machine Learning, University of Twente

Lecturer *Machine Learning Theory* for MasterMath

Lecturer *Graphical Models and Causality* at UT

Welcome! A bit about me



Senior Researcher, Machine Learning group, Centrum Wiskunde & Informatica

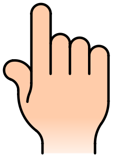
Professor of Mathematical Machine Learning, University of Twente

Lecturer *Machine Learning Theory* for MasterMath

Lecturer *Graphical Models and Causality* at UT

Keywords: Machine Learning, Online Learning, Statistics, Game theory, Optimisation

The idea of this workshop



I'll sketch a PhD project trajectory.

Please interrupt!

What is a PhD project

- Typically 4 years
- One or more supervisors
- Training to be an independent researcher …
- … by doing actual research
- Large academic freedom
- Join community: conferences, workshops, summer schools, internship

How to start a PhD project

Need a supervisor with an open position

Don’t wait for the perfect vacancy. Engage!

Check out the PhD program of the European Laboratory for Learning and Intelligent Systems (ELLIS). They do central recruiting for top AI/ML in Europe.

ellis
European Laboratory for Learning and Intelligent Systems

It all starts with a question

Suppose we are excited about autonomous driving.

It all starts with a question

Suppose we are excited about autonomous driving.

As you may know, training AI systems (e.g. deep neural networks) takes a lot of data.

It all starts with a question

Suppose we are excited about autonomous driving.

As you may know, training AI systems (e.g. deep neural networks) takes a lot of data.

Yet not all data are equally valuable/useful.

It all starts with a question

Suppose we are excited about autonomous driving.

As you may know, training AI systems (e.g. deep neural networks) takes a lot of data.

Yet not all data are equally valuable/useful.

Let's optimise and automate the data collection.

It all starts with a question

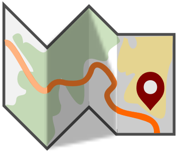
Suppose we are excited about autonomous driving.

As you may know, training AI systems (e.g. deep neural networks) takes a lot of data.

Yet not all data are equally valuable/useful.

Let's optimise and automate the data collection.

Where should I send my prototype for training?

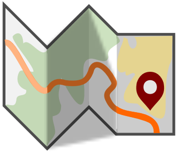


Dragons everywhere

Where should I send my prototype for training?

- AI system is huge parameterised model
- Lots of possible environments to drive in
- Multiple objectives (safety, efficiency, …)
- Feedback (crash/intervention) is very one-sided

Dragons everywhere



Where should I send my prototype for training?

- AI system is huge parameterised model
- Lots of possible environments to drive in
- Multiple objectives (safety, efficiency, …)
- Feedback (crash/intervention) is very one-sided

Distilled goal:

- Identify parameters that cause fewest crashes in natural environment mix.

Simplify

Close-by parameters in close-by environments result in close-by outcomes

Simplify

Close-by parameters in close-by environments result in close-by outcomes

Often true, but too complicated. Let’s discretise!

Simplify

Close-by parameters in close-by environments result in close-by outcomes

Often true, but too complicated. Let's discretise!

Environments

Simplify

Close-by parameters in close-by environments result in close-by outcomes

Often true, but too complicated. Let's discretise!

Environments

Parameters

Simplify

Close-by parameters in close-by environments result in close-by outcomes

Often true, but too complicated. Let’s discretise!

Environments





Parameters




Outcomes


World model (Stochastic Bandit)

The world simplifies to vector + table:

World model (Stochastic Bandit)

The world simplifies to vector + table:

Known natural environment mix

$\mathbb{P}\Big($  $\Big) = 40\%$

$\mathbb{P}\Big($  $\Big) = 30\%$

$\mathbb{P}\Big($  $\Big) = 15\%$

$\mathbb{P}\Big($  $\Big) = 10\%$

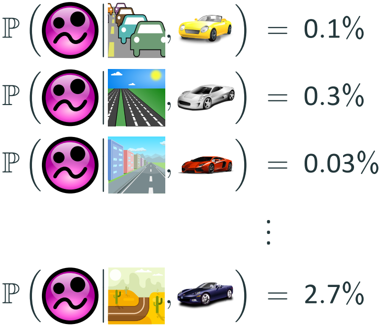
$\mathbb{P}\Big($  $\Big) = 5\%$

World model (Stochastic Bandit)

The world simplifies to vector + table:

Known natural environment mix

Unknown crash probabilities

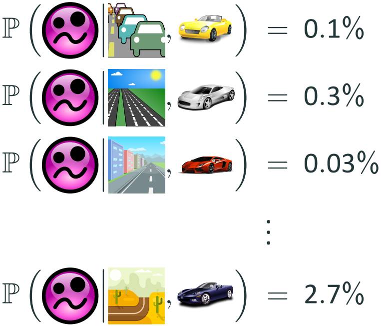


World model (Stochastic Bandit)

The world simplifies to vector + table:

Known natural environment mix

Unknown crash probabilities



Together these determine the best parameter on average. Say .

Learning the best Parameter by Driving: Example Interaction

Learning the best Parameter by Driving: Example Interaction

1   

Learning the best Parameter by Driving: Example Interaction

1

2

Learning the best Parameter by Driving: Example Interaction

Learning the best Parameter by Driving: Example Interaction

1
2
3
4

Learning the best Parameter by Driving: Example Interaction

Learning the best Parameter by Driving: Example Interaction

1
2
3
4
5
⋮
65535
65536

Learning the best Parameter by Driving: Example Interaction

Design

So how do we build that learning algorithm?

- Reliable
- Data efficient

Theory: Characteristic Time and Oracle Weights

Answering correctly in world μ requires data rejecting all worlds with a different answer.

Theory: Characteristic Time and Oracle Weights

Answering correctly in world $\boldsymbol{\mu}$ requires data rejecting all worlds with a different answer.

Theorem (Garivier and Kaufmann, 2016; Russac et al., 2021)

Any δ-correct testing algorithm must, for any world $\boldsymbol{\mu}$, take samples at least

$$\boxed{\textit{samples}(\boldsymbol{\mu})} \geq \ln \frac{1}{\delta} \cdot \boxed{\frac{1}{\max\limits_{\text{par+env proportions } \boldsymbol{w}} \; \min\limits_{\substack{\text{world } \boldsymbol{\lambda} \text{ with answer} \\ \text{different from that of } \boldsymbol{\mu}}} \; \sum\limits_{\text{par } p,\text{ env } e} w_{p,e} \, \mathrm{KL}(\mu_{p,e}, \lambda_{p,e})}}$$

Theory: Characteristic Time and Oracle Weights

Answering correctly in world $\boldsymbol{\mu}$ requires data rejecting all worlds with a different answer.

Theorem (Garivier and Kaufmann, 2016; Russac et al., 2021)

Any δ-correct testing algorithm must, for any world $\boldsymbol{\mu}$, take samples at least

$$\boxed{\textit{samples}(\boldsymbol{\mu})} \geq \ln \frac{1}{\delta} \cdot \boxed{\frac{1}{\max\limits_{\text{par+env proportions } \boldsymbol{w}} \min\limits_{\substack{\text{world } \boldsymbol{\lambda} \text{ with answer} \\ \text{different from that of } \boldsymbol{\mu}}} \sum\limits_{\text{par } p,\text{ env } e} w_{p,e} \operatorname{KL}(\mu_{p,e}, \lambda_{p,e})}}$$

Why should we care?

- Characterises* complexity of each world $\boldsymbol{\mu}$
- Optimal testing algorithm must sample with proportions $\arg\max_{\boldsymbol{w}}$

What are those oracle weights $w^*(\mu)$

$w^*_{e,p}(\mu)$				
	0.05	0.01	0.13	...
				
				
				
			...	0.08

Instance-Optimal Algorithms

Sample complexity lower bound at world $\boldsymbol{\mu}$ governed by max-min problem:

$$\max_{\text{par+env proportions } \boldsymbol{w}} \quad \min_{\substack{\text{world } \boldsymbol{\lambda} \text{ with answer} \\ \text{different from that of } \boldsymbol{\mu}}} \quad \sum_{\text{par } p, \text{ env } e} w_{p,e} \operatorname{KL}(\mu_{p,e}, \lambda_{p,e})$$

Main challenge: driving with proportions $\boldsymbol{w}^*(\boldsymbol{\mu}) = \arg\max_{\boldsymbol{w}}$ without knowing world $\boldsymbol{\mu}$.

Instance-Optimality: Iterative Saddle Point Approach

CWI specialty

Approx. solve saddle point problem iteratively: $w_1, w_2, \ldots \to w^*(\mu)$

Instance-Optimality: Iterative Saddle Point Approach

CWI specialty

Approx. solve saddle point problem iteratively: $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots \to \boldsymbol{w}^*(\boldsymbol{\mu})$

Main pipeline:

- Get current $\boldsymbol{w}_t$ from saddle point solver.
- Pick parameter and environment $(P_t, E_t) \sim \boldsymbol{w}_t$, see outcome X_t
- Update estimate $\hat{\boldsymbol{\mu}}_t$ of world.
- Advance the saddle point solver one iteration.
- Add optimism to gradients to induce exploration ($\hat{\boldsymbol{\mu}}_t \to \boldsymbol{\mu}$).
- Regret bounds + concentration + optimism $\Rightarrow$ finite-time guarantee:

Instance-Optimality: Iterative Saddle Point Approach

CWI specialty

Approx. solve saddle point problem iteratively: $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots \to \boldsymbol{w}^*(\boldsymbol{\mu})$

Main pipeline:

- Get current $\boldsymbol{w}_t$ from saddle point solver.
- Pick parameter and environment $(P_t, E_t) \sim \boldsymbol{w}_t$, see outcome X_t
- Update estimate $\hat{\boldsymbol{\mu}}_t$ of world.
- Advance the saddle point solver one iteration.
- Add optimism to gradients to induce exploration ($\hat{\boldsymbol{\mu}}_t \to \boldsymbol{\mu}$).
- Regret bounds + concentration + optimism $\Rightarrow$ finite-time guarantee:

Theorem (Degenne, Koolen, and Ménard, 2019)

For every $\delta \in (0,1)$ and world $\boldsymbol{\mu}$, the above scheme takes samples bounded by

$$\textit{samples}(\boldsymbol{\mu}) \leq \boxed{\textit{char. time}} \cdot \ln \frac{1}{\delta} + o(\ln \frac{1}{\delta})$$

Lessons

Content:

- Optimal data collection can be achieved by learning algorithms
- It is inefficient follow the natural environment mix
- It will take many samples to see small differences between good parameters
- Discretising parameters finer makes learning harder …
- …while finer discretisation of environments can help

Lessons

Content:

- Optimal data collection can be achieved by learning algorithms
- It is inefficient follow the natural environment mix
- It will take many samples to see small differences between good parameters
- Discretising parameters finer makes learning harder …
- …while finer discretisation of environments can help

Meta:

- It will go deep
- Learning/AI/ML will require mix of algorithms, statistics, game theory, optimisation
- Need to zoom out, scale up and iterate. This is hard!

References

Degenne, R., W. M. Koolen, and P. Ménard (Dec. 2019). **"Non-Asymptotic Pure Exploration by Solving Games".** In: *Advances in Neural Information Processing Systems (NeurIPS) 32*, pp. 14492–14501.

Garivier, A. and E. Kaufmann (2016). **"Optimal Best arm Identification with Fixed Confidence".** In: *Proceedings of the 29th Conference On Learning Theory (COLT)*.

Russac, Y., C. Katsimerou, D. Bohle, O. Cappé, A. Garivier, and W. M. Koolen (Dec. 2021). **"A/B/n Testing with Control in the Presence of Subpopulations".** In: *Advances in Neural Information Processing Systems (NeurIPS) 34*.