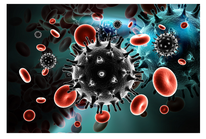
Combining Kernel and Model-Based Reinforcement Learning for HIV Therapy Selection

Sonali Parbhoo[1], Jasmina Bogojeska[2], Volker Roth[1] and Finale Doshi-Velez[3]

Department of Mathematics and Computer Science, University of Basel, Switzerland[1]
IBM Research, Zurich, Switzerland[2]
Harvard University, John A. Paulson School of Engineering and Applied Sciences, Cambridge, MA.[3]

Motivation

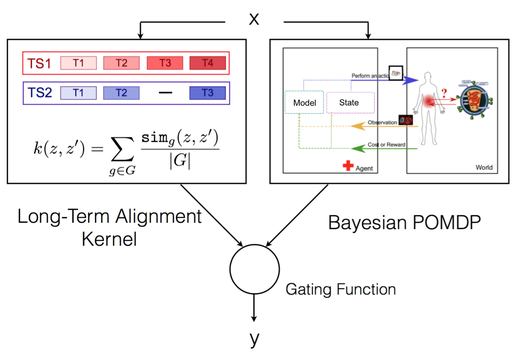
Human Immunodeficiency Virus (HIV) infects immune cells and causes AIDS. The virus is highly mutagenic, so choosing a therapy is difficult.



Combination therapy is prescribed to overcome drug resistance. Existing therapy selection approaches based on regression are problematic:

- Do not account for **long-term** effects of therapies.
- Do not account for **patient heterogeneity.**

Trade-off between model generality vs specificity.



We propose a **Mixture-of-Experts** (MoE) model to overcome this.
Goal: Automatic optimal therapy selection

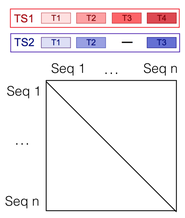
The History Alignment Kernel

Patients with similar treatment histories respond similarly.

- Alignment of therapies [1]:

$$k(z,z') = \sum_{g \in G} \frac{\mathtt{sim}_g(z,z')}{|G|}; \quad \mathtt{sim}_g(z,z') = \frac{u_{zg}^\top u_{z'g}}{\max(||u_{zg}||^2, ||u_{z'g}||^2)}$$

- Alignment of therapy sequences with Needleman Wunsch.



More common drugs/mutations $\implies$ *more similar treatments.*
Better alignment $\implies$ *more similar histories.*

Short-Term Therapy Success:
Viral RNA < 400 copies/mL

Long-Term Therapy Success

- Immediate reward criterion:

$$r_t = \begin{cases} -0.7 \log V_t + 0.6 \log T_t - 0.2|M|, & \text{if } V_t \text{ is above detection limits} \\ 5 + 0.6 \log T_t - 0.2|M|, & \text{otherwise} \end{cases}$$

Long-term success sums criterion over patient's future history.

Bayesian POMDP

Models **sequential decision-making** explicitly.

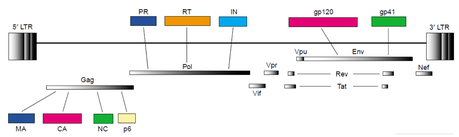
- 6 month time increments
- 7 hidden physiological states; 70 mutations and binned viral RNA values
- Transition and observation parameters drawn from Dirichlet
- Sample m models and update beliefs accordingly.

How do we know what action to choose?

- Build forward search tree: Search therapies to optimise outcomes over 5 years.
- Evaluate policies with off-policy evaluation.

Data

- 32 960 patients' Therapy Change Episode (TCE) and clinical data
- 312 most frequently occurring drug combinations
- Genotype resistance data from *Pol* and *Env* regions [2].



The viral genome.

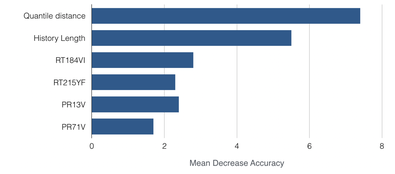
Experiments

	Doubly Robust	IS	Weighted IS
Short-term Alignment	2.17 ±1.47	2.14 ±1.22	2.15 ±1.16
Long-term Alignment	9.48 ± 1.90	5.42 ± 1.93	6.74 ± 1.89
POMDP	6.34 ± 2.15	4.36 ± 2.38	6.76 ± 2.24
MoE	**11.47 ± 1.38**	**12.25 ± 1.41**	**11.23 ± 1.40**
Truncated MoE	4.61 ± 2.35	4.73 ± 2.49	4.71 ± 2.18

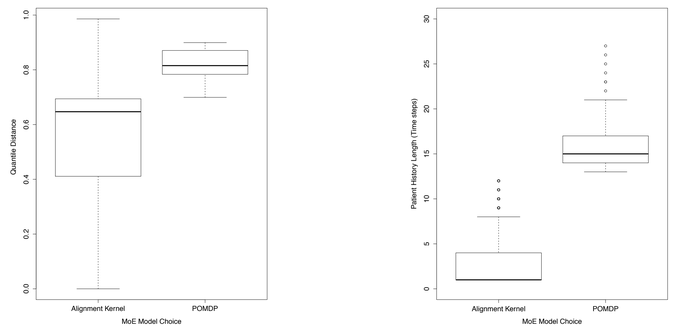
Off-policy evaluation for therapy selection models.

- Mixture-of-Experts selects therapy policy over 5 year horizon.
- Mixture-of-Experts chooses the POMDP 26% vs. the Long-Term Alignment 74%.
- Off-policy evaluation corrects distributional mismatch between data policy and learned policy.

When does the Mixture-of-Experts pick each model?



Feature importance for Mixture-of-Experts.



Mixture-of-Experts model choice over a) distances to closest neighbour b) history lengths.

Conclusions

- Mixture-of-experts addresses trade-off between model **generality** and **specificity**.
- **Different policies** produced when optimising over long-term vs. short-term.
- Mixture-of-Experts yields **best** policies.
- Expert chosen based on **clustering characteristics**.

References

[1] Jasmina Bogojeska, Daniel Stöckel, Maurizio Zazzi, Rolf Kaiser, Francesca Incardona, Michal Rosen-Zvi, and Thomas Lengauer. History-alignment models for bias-aware prediction of virological response to hiv combination therapy. In *AISTATS*, pages 118–126, 2012.

[2] Annemarie M Wensing, Vincent Calvez, Huldrych F Günthard, Victoria A Johnson, Roger Paredes, Deenan Pillay, Robert W Shafer, and Douglas D Richman. Special contribution 2014 update of the drug resistance mutations in hiv-1. *Topics in antiviral medicine*, page 642, 2014.

Acknowledgement: We thank the Swiss National Science Foundation for financial support and EuResist Integrated Database for providing the data used in this research.