Avoiding Undesired Future with Minimal Cost in Non-Stationary Environments

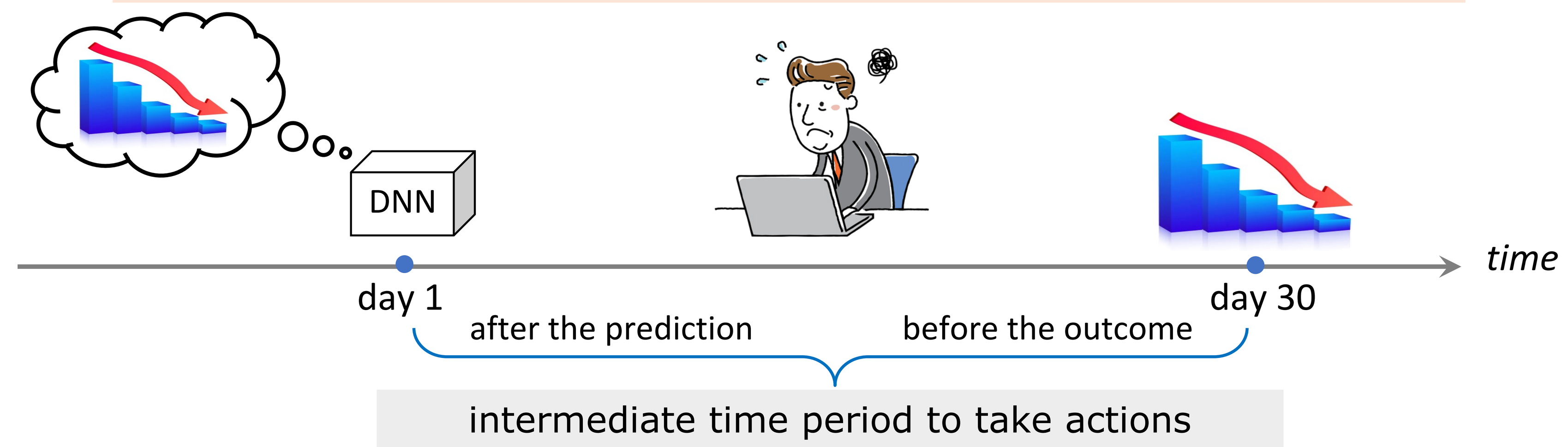
Wen-Bo Du Tian Qin Tian-Zuo Wang Zhi-Hua Zhou

National Key Lab for Novel Software Technology, Nanjing University, China

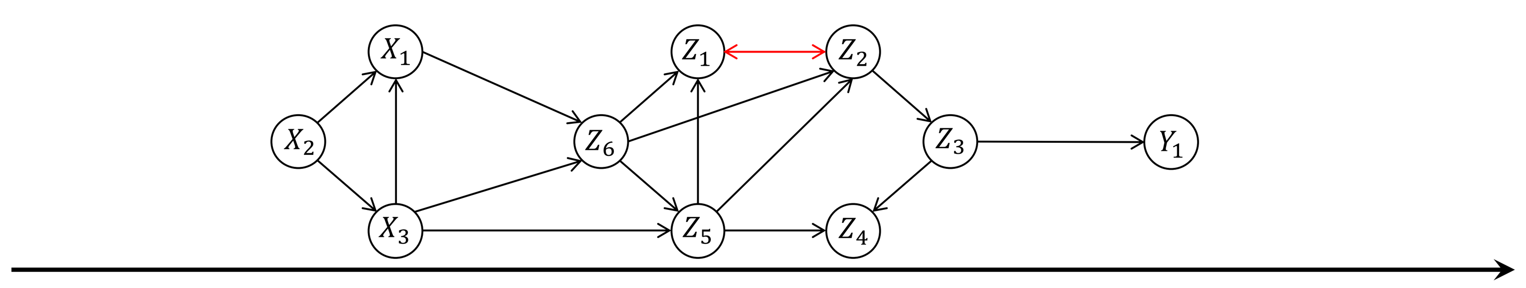
{duwb, qint, wangtz, zhouzh}@lamda.nju.edu.cn

Motivation

How to Avoid Undesired Future (AUF) if an ML model predicts that bad things will happen?



➢ **Structural rehearsal model [1, 2]**



$\mathbf{V} := f(\mathbf{PA}^G;\beta) + \boxed{\varepsilon}$ noise, e.g. $\varepsilon \sim \mathcal{N}(0,\Sigma)$

rehearsal graph G

structural equation

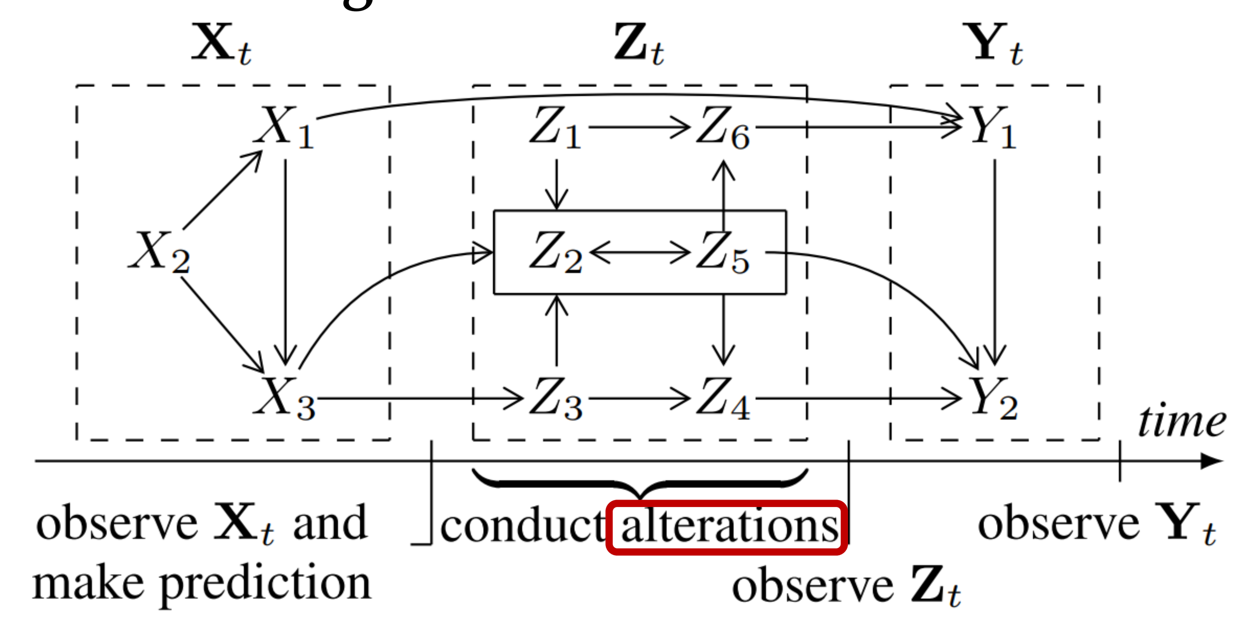
structural rehearsal model (SRM): $\langle G, \boldsymbol{\theta}\rangle$

[1] Z.-H. Zhou. Rehearsal: Learning from prediction to decision. Frontiers of Computer Science, 2022.

[2] T. Qin, T.-Z. Wang, Z.-H. Zhou. Rehearsal learning for avoiding undesired future. NeurIPS, 2023.

Problem formulation

➢ **Rehearsal learning framework**



➢ **AUF with minimal cost in non-stationary environments**

$$\min_{\mathbf{z}_t^{\xi_t}} \quad \left(\mathbf{z}_t^{\xi_t} - \mathbf{z}_t^0\right)^\top \mathbf{W}\left(\mathbf{z}_t^{\xi_t} - \mathbf{z}_t^0\right) \qquad (\star)$$
$$\text{s.t.} \quad \mathbb{P}\left(\widehat{\mathbf{Y}}_t \in \mathcal{S} \mid \widehat{\boldsymbol{\theta}}_t, \mathbf{x}_t, Rh(\xi_t)\right) \geq \tau$$

- How can optimization be performed under a probabilistic constraint?
- Using the QCQP transformation to realize *efficient alteration selection*.

- How to accurately estimate $\boldsymbol{\theta}_t$ in non-stationary environment?
- Adopting online-ensemble to obtain the *estimation of dynamic influence*.

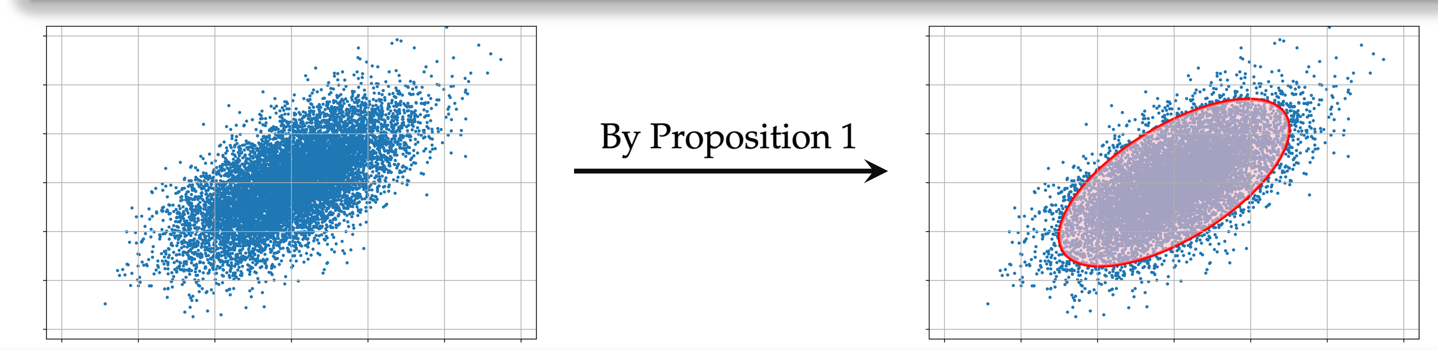
Efficient alteration selection

➢ **Probability region for $Pr_{AUF} = \tau$**

Proposition 1 *The following set $\mathcal{P}$ is a probability region which satisfies the fact that $\mathbb{P}(\mathbf{Y}_t \in \mathcal{P} \mid \boldsymbol{\theta}_t, \mathbf{x}_t, Rh(\xi_t)) = \tau$:*

$$\mathcal{P} = \left\{\boldsymbol{\mu}_{\mathbf{y}_t} + \left(\chi^{-1}(\tau)\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\top\right)^{\frac{1}{2}}\mathbf{u} \;\middle|\; \|\mathbf{u}\|_2 \leq 1\right\},$$

where $\boldsymbol{\mu}_{\mathbf{y}_t} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{z}_t^{\xi}$, $\mathbf{u}$ is an arbitrary point in the unit sphere in $\mathbb{R}^{|\mathbf{Y}_t|}$, and $\chi^{-1}(\cdot)$ denotes the quantile function of the χ^2_λ distribution with $\lambda = |\mathbf{Y}_t|$.



Theorem 2 *By using the suggested alteration ξ_t from:*

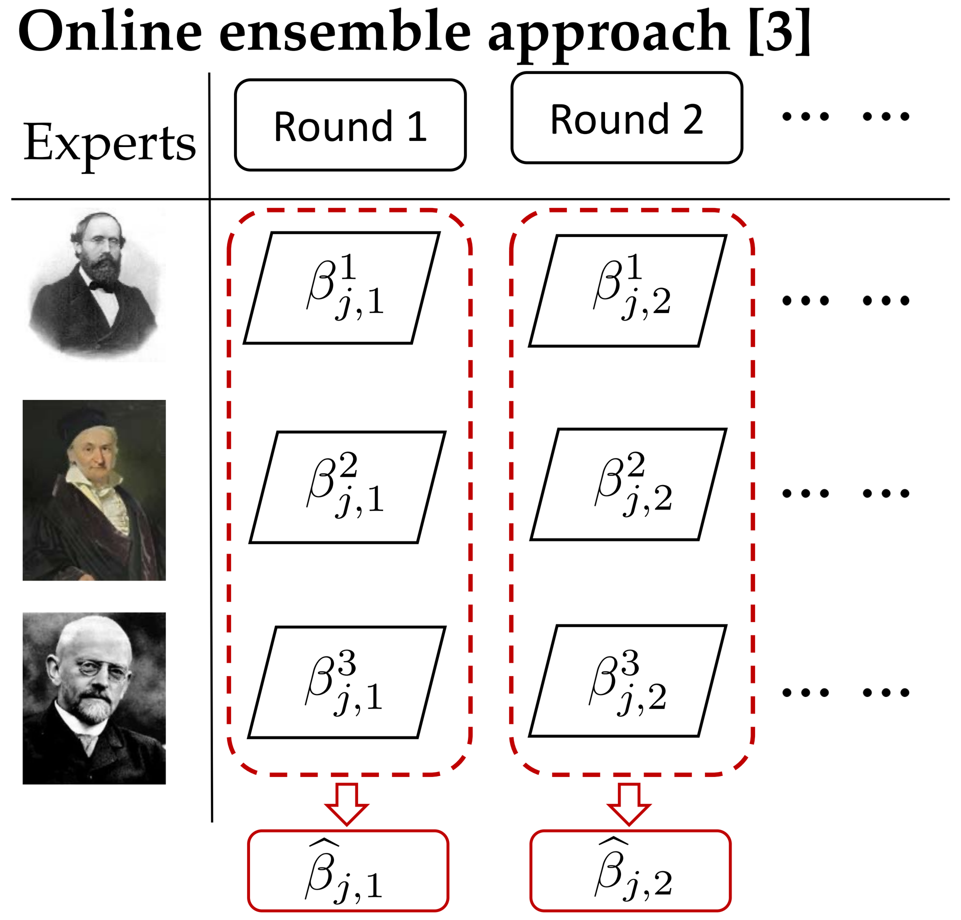
$$\min_{\mathbf{z}_t^{\xi_t}} \quad \left(\mathbf{z}_t^{\xi_t} - \mathbf{z}_t^0\right)^\top \mathbf{W}\left(\mathbf{z}_t^{\xi_t} - \mathbf{z}_t^0\right)$$
$$\text{s.t.} \quad \mathbf{M}\mathbf{A}\mathbf{x}_t + \mathbf{M}\mathbf{B}\mathbf{z}_t^{\xi_t} + \|\mathbf{M}\mathbf{P}\|_{2,row} \leq \mathbf{d};$$

where $\mathbf{P} = \left(\chi^{-1}(\tau)\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\top\right)^{\frac{1}{2}}$, it can be guaranteed that $\mathbb{P}\left(\mathbf{Y}_t \in \mathcal{S} \mid \widehat{\boldsymbol{\theta}}_t, \mathbf{x}_t, Rh(\xi_t)\right) \geq \tau$.

Leveraging the probability region proposed in Prop. 1, the original AUF problem ($\star$) can be transformed into a convex QCQP in Thm. 2. Hence, the *decision alteration can be selected efficiently and effectively*.

Dynamic influence estimation

➢ **Online ensemble approach [3]**



[3] P. Zhao, Y.-J. Zhang, L. Zhang, and Z.-H. Zhou. Adaptivity and non-stationarity: Problem dependent dynamic regret for online convex optimization. Journal of Machine Learning Research, 2024.

- Time-decreasing term
- Error term reflecting problem difficulty

Theorem 1 *Denote the true parameter of β_j in time t as $\beta_{j,t}$ ($j \in [|\boldsymbol{V}|]$) and choose $\eta_j \in (0, 1/2L_j]$ as the step size used in SGD, then it holds that:*

$$\mathbb{E}\left\|\widehat{\beta}_{j,t} - \beta_{j,t}\right\|^2 \lesssim (1-\mu_j\eta_j)^{\frac{t}{m}}\left\|\widehat{\beta}_{j,0} - \beta_{j,0}\right\|^2 + \left(\frac{m\Delta_j}{\mu_j\eta_j}\right)^2 + \frac{\eta_j\sigma^2}{\mu_j},$$

where μ_j, L_j are the minimal and maximal eigenvalues of $\{\ell_{j,t}(\cdot)\}_{t=1}^T$'s Hessian Matrices, σ^2 upper-bounds $\mathrm{Var}(\widehat{g}_{j,t})$, Δ_j upper-bounds $\|\beta_{j,t+1} - \beta_{j,t}\|_2$, and m is the longest continuous rehearsal times, for most of the V_js, $m = 1$.

Proposition 2 *Assume $\{\widehat{\ell}_{j,t}(\cdot)\}_{t=1}^T$s are bounded for $\forall \beta_i \in \mathcal{B}$ and $t \in [T]$; then by choosing $\alpha = \sqrt{\ln N_j/T}$, for any $\eta \in \mathcal{H}_j$, estimations $\widehat{\beta}_{j,t}$s from Algorithm 1 satisfies that $\sum_{t=1}^T \widehat{\ell}_t\left(\widehat{\beta}_{j,t}\right) - \sum_{t=1}^T \widehat{\ell}_t\left(\beta_{j,t}^\eta\right) \leq \mathcal{O}\left(\sqrt{T\ln N_j}\right)$; where N_j is the number of base-learners, $\beta_{j,t}^\eta$ is the estimation from any expert with step size η in expert set $\mathcal{H}_j$ in Algorithm 1.*

The dynamic influence relations can be *effectively estimated* by using the presented online-ensemble based method.

Experiments