From Predictive Analytics to Optimization-Guided Learning (OGL)

Chia-Yen Lee, Ph.D. (李家岩 博士)

Department of Information Management (資訊管理學系)
National Taiwan University (國立台灣大學)

Education

- Ph.D, ISE, Texas A&M University, USA (Major: Operations Research)
- M.S., IEEM, National Tsing Hua University, Taiwan
- B.S. & B.B.A., Mathematical Science and MIS, National Chengchi University, Taiwan

Experience

- Prof., Dept of Information Management, National Taiwan University (2020-now)
- Deputy Director, CPO, tsmc (2024-now).
- Director, Institute of Manufacturing Information and Systems, NCKU (2018-2020)
- AE, IEEE Transactions on Semiconductor Manufacturing (2023-2025)
- AE, IEEE Transactions on Automation Science and Engineering (2020-2022)
- Grants: 50+ industry-academia cooperation projects (manufacturing focus)
  - Applications: Semiconductor manufacturing/packaging, panel, motor drier, fasteners, machine tools, petrochemical, plant factory, educational process, hospitals, etc.
- Committee: National Quality Award, National Science & Technology Council, etc.
- Consultant: Semiconductor, TFT-LCD, AutoML Startup, Taiwan AI Academy, etc.

Award

- IE Award, CIIE (2023); Outstanding Research Award, MOST (2022); IEEE Senior Member (2021); Micron Teacher Award (2018); Ta-You Wu Memorial Award of Distinguished Young Scholars, MOST (2017)

Research Interest

- Manufacturing Data Science, Intelligent Manufacturing Systems, Productivity and Efficiency Analysis, Multi-Objective Stochastic Optimization

Agenda

- Five-Phase Analytics
- From Predictive to Prescriptive Analytics
- From Prescriptive to Optimization-Guided Learning
- Takeaway

Five-Phase Analytics

- FIVE-Phase Analytics: *A way from POINT to PLANE*

What	Why	When	How/Who	Where
What happened?	Why did it happen?	When/What will happen next?	How can we optimize decision? Who takes it?	Where should build automatic routines?
Descriptive Analytics	Diagnostic Analytics	Predictive Analytics	Prescriptive Analytics	Automating Analytics
Data Collection Visualization Niche Identification	Root Cause Interpretation Causality	Variability Information Diversity Trend Pattern	Risk Assessment Resource Leveraging Decision Tradeoff	Real-time Monitoring Domain Adaptation Financial Dashboard
見	識	謀	斷	行

U N I S O N

Lee, C.-Y., and Chien, C.-F., 2022. Pitfalls and protocols of data science in manufacturing practice. *Journal of Intelligent Manufacturing*, 33, 1189–1207.

One Example…

- We typically build AI models for prediction or scenario analysis
  - CNN, LSTM, SVM, Random Forest, Boosting, PLS, …
  - Then…Which model is better? What's the next step after prediction?
- Prediction is Risky!
  - How about the potential risk (i.e. loss) after decision-making?
- Example
  - Model A with accuracy 95%, however, inaccurate prediction could lead to big loss.
  - Model B with accuracy 90%, however, inaccurate prediction could lead to small loss.
  - Which model do you prefer?
- Predictive Thinking → Prescriptive Decision

POLab

Confusion Matrix for Binary Classification

- Two risks: false alarm (type I) and miss rate (type II) → Prescriptive
- Trade-off between two misclassified errors → cost sensitive

Model A		Predict: FAIL	Predict: PASS
True	FAIL	61	7
True	PASS	29	31

Model B		Predict: FAIL	Predict: PASS
True	FAIL	47	21
True	PASS	7	53

	Testing	
	Accuracy	AUC
Model A	71.9%	70.2%
Model B	78.1%	78.9%

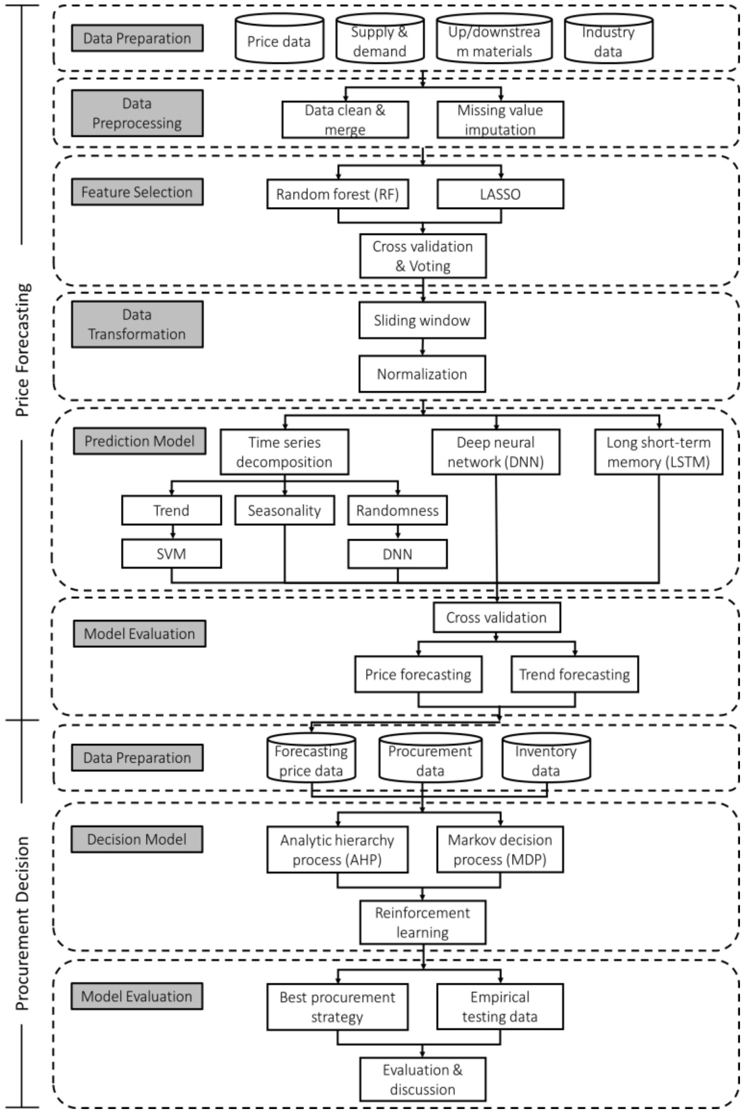
AUC: Area under the Curve of ROC

Lee, C.-Y., and Chien, C.-F., 2022. Pitfalls and protocols of data science in manufacturing practice. *Journal of Intelligent Manufacturing*, 33, 1189–1207.

Decisions take into account the RISKS associated with the realization of uncertain events.

Raw Material Price Prediction and Procurement

◻ Price Forecast (Predictive Analytics)



Feature Extraction

Crude Oil

	Variable	r.c.	c.c.
	X01	17.9591	0.5195
★	X02	18.3465	0.5229

Upstream

	Variable	r.c.	c.c.
★	X03	2.2597	0.8511
★	X04	0.8613	0.1990
	X05	1.6192	0.5878
	X06	-0.3626	-0.1536
	X07	2.0611	0.5830
★	X08	-0.4406	-0.4160
	X09	0.0002	0.4336

Butadiene

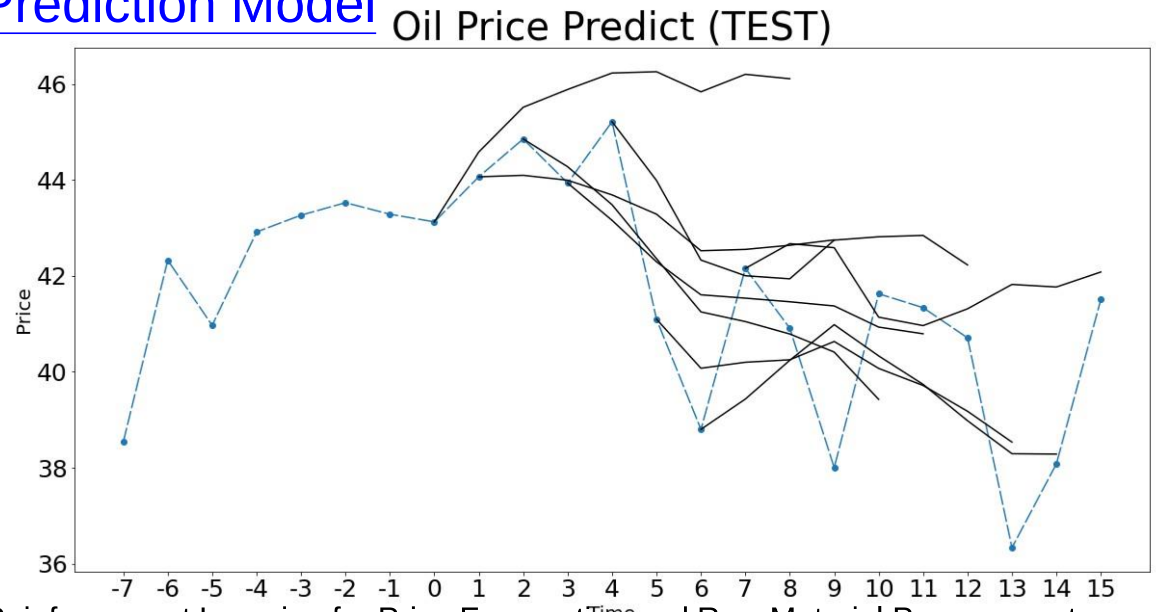
	Variable	r.c.	c.c.
	Y		
	X10	0.9936	0.9997
	X11	1.0051	0.9997
	X12	0.8651	0.8741
	X13	0.9153	0.8584
★	X14	0.8382	0.9401
	X15	0.8153	0.9406
	X16	0.8615	0.9390
	X17	0.8751	0.8425
	X18	0.8802	0.9512
	X19	0.2148	0.0719
	X20	0.3655	0.1411
	X21	0.0012	0.4902

Downstream

	Variable	r.c.	c.c.
★	X22	5.6786	0.7958
	X23	0.7643	0.9258
	X24	0.8698	0.9424
	X25	0.9213	0.7151
	X26	0.8003	0.9272
	X27	0.8897	0.9313
	X28	0.0008	0.5036
	X29	0.0005	0.4319
★	X30	1.0112	0.9810
	X31	1.3185	0.9753
	X32	0.1570	0.1345
★	X33	0.7595	0.3052
	X34	1.4520	0.6067

Legend	
(pink)	Butadiene(BD)
(yellow)	r.c. > 1.5
(blue)	r.c. < 0
★	Important variables identified by feature selection

Prediction Model



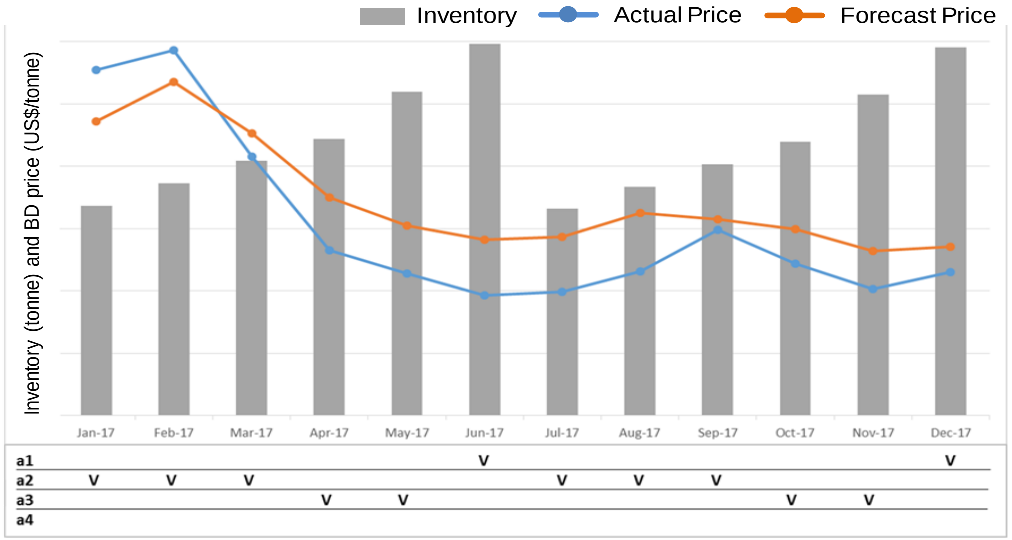
Lee, C.-Y., Chou, B.-J., and Huang, C.-F. 2021. Data Science and Reinforcement Learning for Price Forecasting and Raw Material Procurement in Petrochemical Industry. *Advanced Engineering Informatics*, 51, 101443.

Raw Material Price Prediction and Procurement

Reinforcement Learning (Prescriptive Analytics)

Human Judgment Current Policy

Optimal Policy Reinforcement Learning



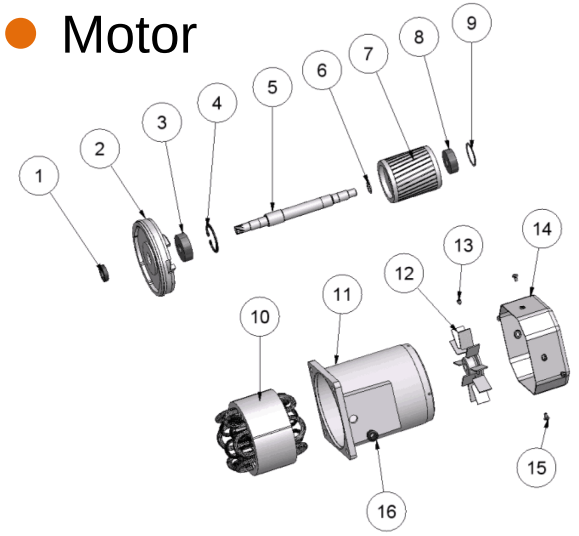
	Current policy	(s,S) policy	Optimal policy
Average inventory (tonne)	3112	1812	3197
Standard deviation of inventory (tonne)	743	302	489
Amount purchased (tonne)	25,301	35,430	36,835
Total cost (US$)	44,596,113	42,324,694	39,091,618

Lee, C.-Y., Chou, B.-J., and Huang, C.-F. 2021. Data Science and Reinforcement Learning for Price Forecasting and Raw Material Procurement in Petrochemical Industry. *Advanced Engineering Informatics*, 51, 101443.

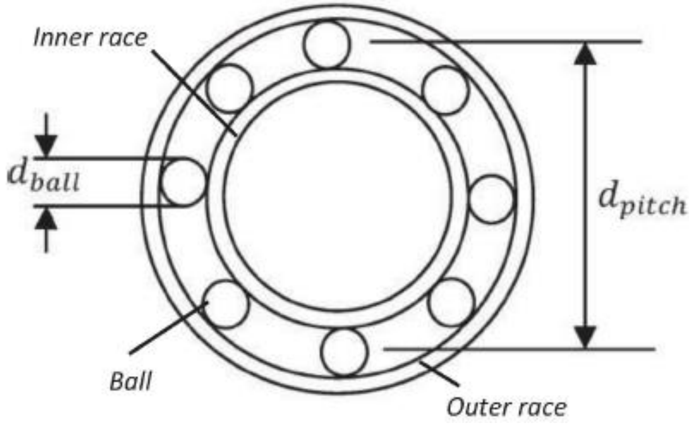
Prognostics and Health Management (PHM)

Remaining Useful Life (RUL) (Predictive Analytics)

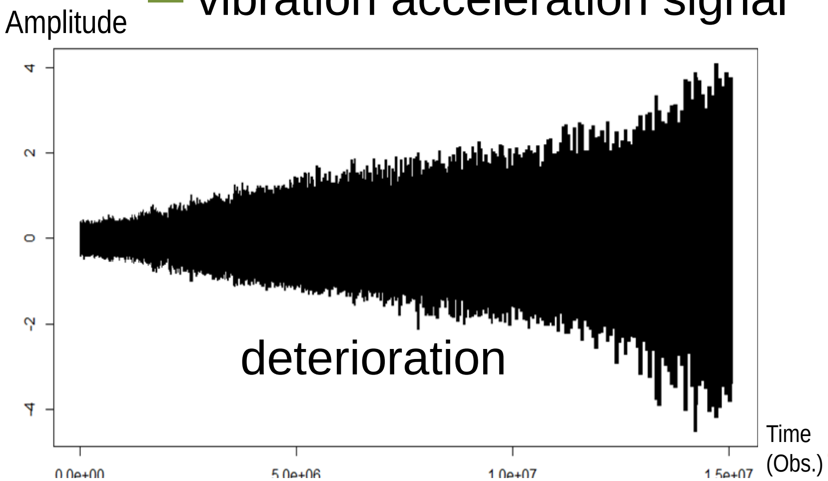
- Motor

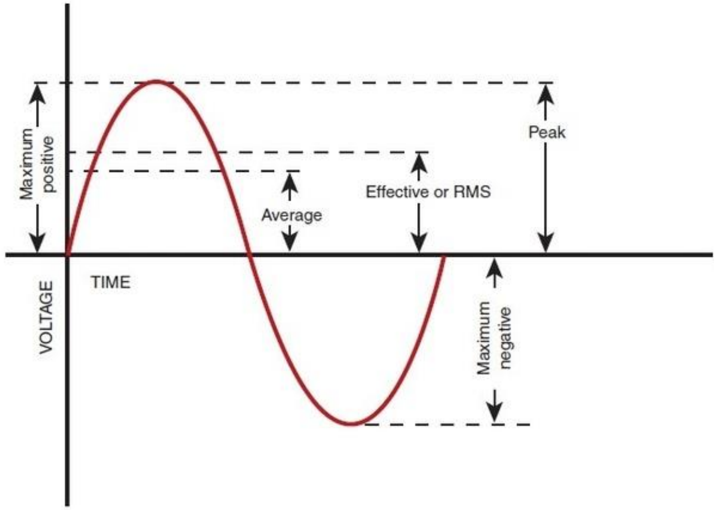


- Bearing



- Data Source
  - vibration acceleration signal

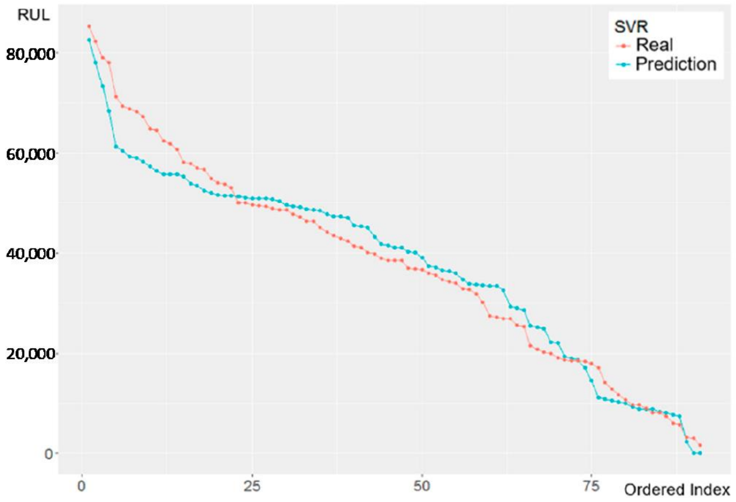




- Feature Engineering
  - Time Domain, Frequency Domain, Time-frequency Domain

Feature	Equation
Maximum Value	$f_{1i} = max(X(i))$
Mean Value	$f_{2i} = \frac{1}{N}\sum_{i=1}^{N} X(i)$
Minimum Value	$f_{3i} = min(X(i))$
Standard Value	$f_{4i} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X(i) - f_{1i})^2}$
Peak to Peak Value	$f_{5i} = f_{1i} - f_{3i}$
Mean Amplitude	$f_{6i} = \frac{1}{N}\sum_{i=1}^{N} \lvert X(i)\rvert$
Root Mean Square Value	$f_{7i} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} X(i)^2}$
Skewness Value	$f_{8i} = \frac{1}{N}\sum_{i=1}^{N} X(i)^3$
Kurtosis Value	$f_{9i} = \frac{1}{N}\sum_{i=1}^{N} X(i)^4$
Waveform Indicator	$f_{10i} = \frac{f_{7i}}{f_{6i}}$
Pulse Indicator	$f_{11i} = \frac{f_{1i}}{f_{6i}}$
Kurtosis Index	$f_{12i} = \frac{f_{9i}}{f_{7i}}$
Peak Index	$f_{13i} = \frac{f_{1i}}{f_{7i}}$
Square Root Amplitude	$f_{14i} = (\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lvert X(i)\rvert})^2$
Margin Indicato	$f_{15i} = \frac{f_{1i}}{f_{14i}}$
Skewness Indicator	$f_{16i} = \frac{f_{8i}}{f_{7i}^4}$

- RUL Prediction

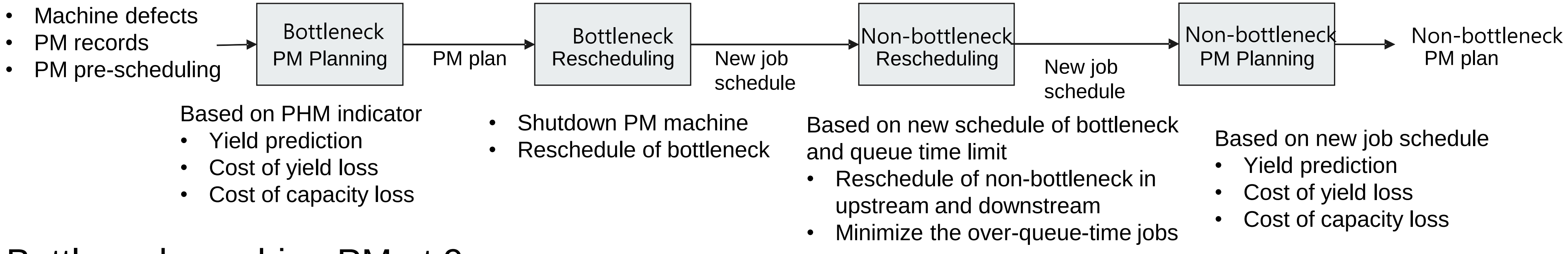


http://www.li-ming.com.tw/

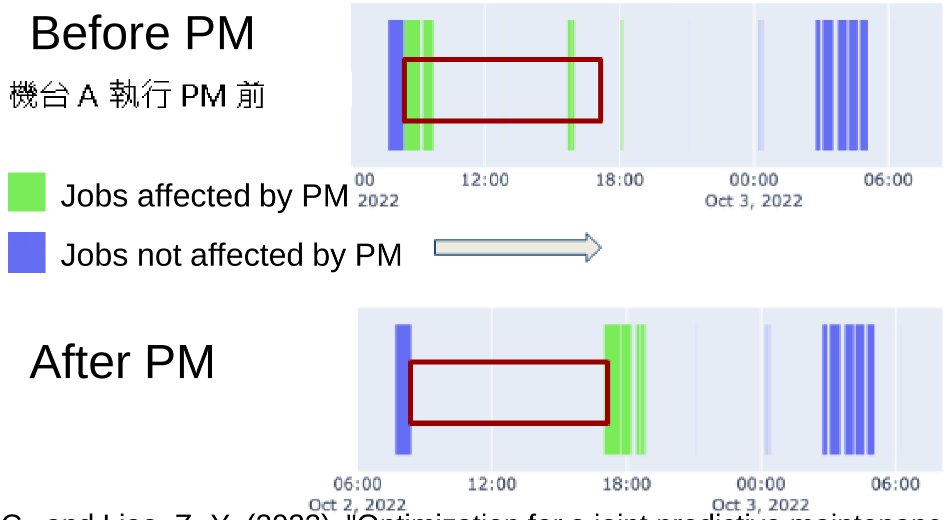
Lee, C.-Y., T.-S. Huang, M.-K. Liu, and C.-Y. Lan. 2019. Data Science for Vibration Heteroscedasticity and Predictive Maintenance of Rotary Bearings. Energies., 12 (5), 801.
Jiang, W., Hong, Y., Zhou, B., He, X. and Cheng, C. 2019. "A gan-based anomaly detection approach for imbalanced industrial time series," IEEE Access, vol. 7, pp. 608–619.

Predictive Maintenance (PdM) Scheduling (Prescriptive Analytics)

- The proposed four-stage PM (preventive maintenance) algorithm provides a tradeoff between machine workload (capacity loss) and condition (PHM indicator; yield loss), and integrates non-bottleneck machines in upstream and downstream of the bottleneck.



Bottleneck machine PM at 9am



- The system considers bottleneck and connects them to upstream and downstream;
- Supports machine yield prediction;
- Estimates the costs of yield loss and capacity loss;
- Incorporates queue time limit and maintenance resources (available labor hour) into the model;
- Considers the production uncertainty for developing stochastic dynamic programming;
- And recommends the priority of machine PM.

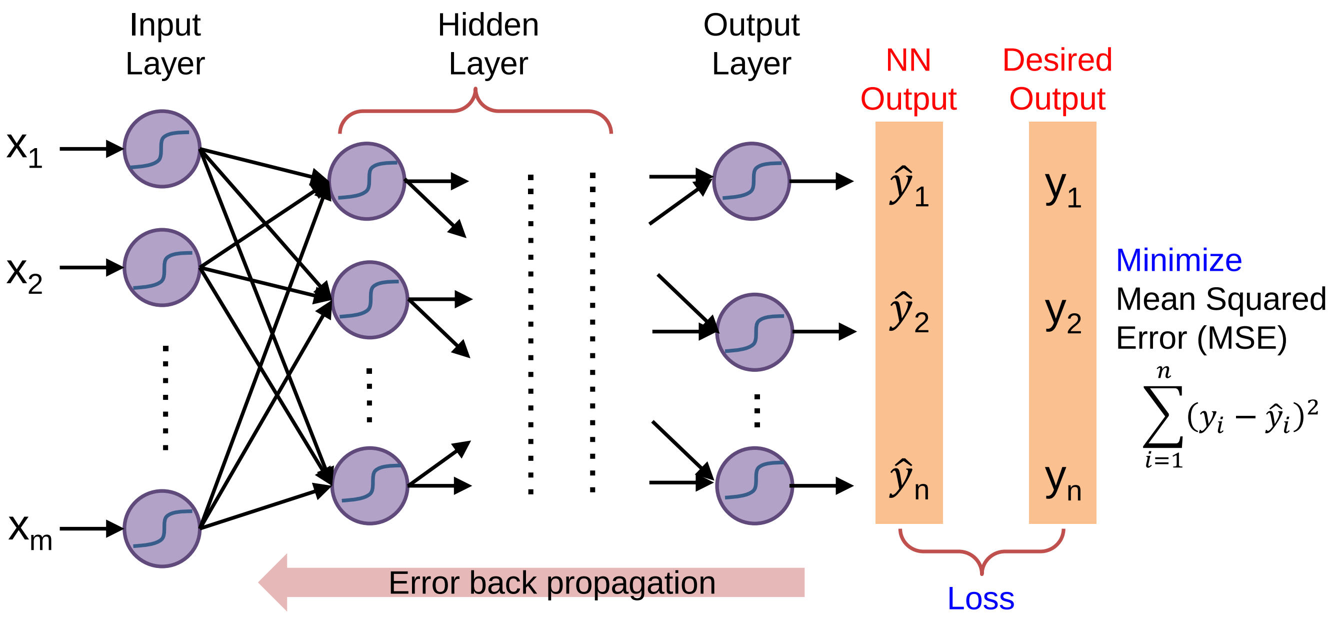
Kung, L.-C., and Liao, Z.-Y. (2022). "Optimization for a joint predictive maintenance and job scheduling problem with endogenous yield rates", IEEE Transactions on Automation Science and Engineering, 19(3), 1555-1566.

Workflow: From Predictive to Prescriptive Analytics

- 1st stage: Predictive Analytics
  - Estimation
  - Prediction or forecast is difficult
    - because it's about the FUTURE
  - Estimation or imputation is reasonable
    - Use known information to estimate UNKNOWN information
- 2nd stage: Prescriptive Analytics
  - Optimization
  - Decision-maker's preference structure
    - Multi-objective decision analysis
  - Resource allocation optimization
    - 8M1I: 人(Man)、機(Machine)、料(Material)、方法(Method)、測量(Measure)、時間(Minutes)、資金(Money)、環境(Mother nature/environment)、資訊(information)
  - Risk assessment & diversification

- Example

1st Stage



2nd Stage

$$
\begin{aligned}
\min\ & \boldsymbol{C}^T\boldsymbol{x} \\
\text{s.t.}\ & \boldsymbol{Ax} \geq \boldsymbol{b} \\
& \tilde{\boldsymbol{T}}\boldsymbol{x} \geq \tilde{\boldsymbol{r}} \\
& \boldsymbol{x} \geq \boldsymbol{0}
\end{aligned}
$$

Prediction is the Process; Decision is the Purpose.

So…

1. Error/MSE focus
2. Data-driven
3. Causality
4. Find the change in unchanging env.
   - 在不變中找變 (infor. content)

From Predictive to Prescriptive Analytics

1. Objective/KPI focus
2. Decision-oriented
3. Resource Allocation
4. Find the unchange in changing env.
   - 在變中找不變 (robustness)

Stock Price Forecast

Predictive Analytics

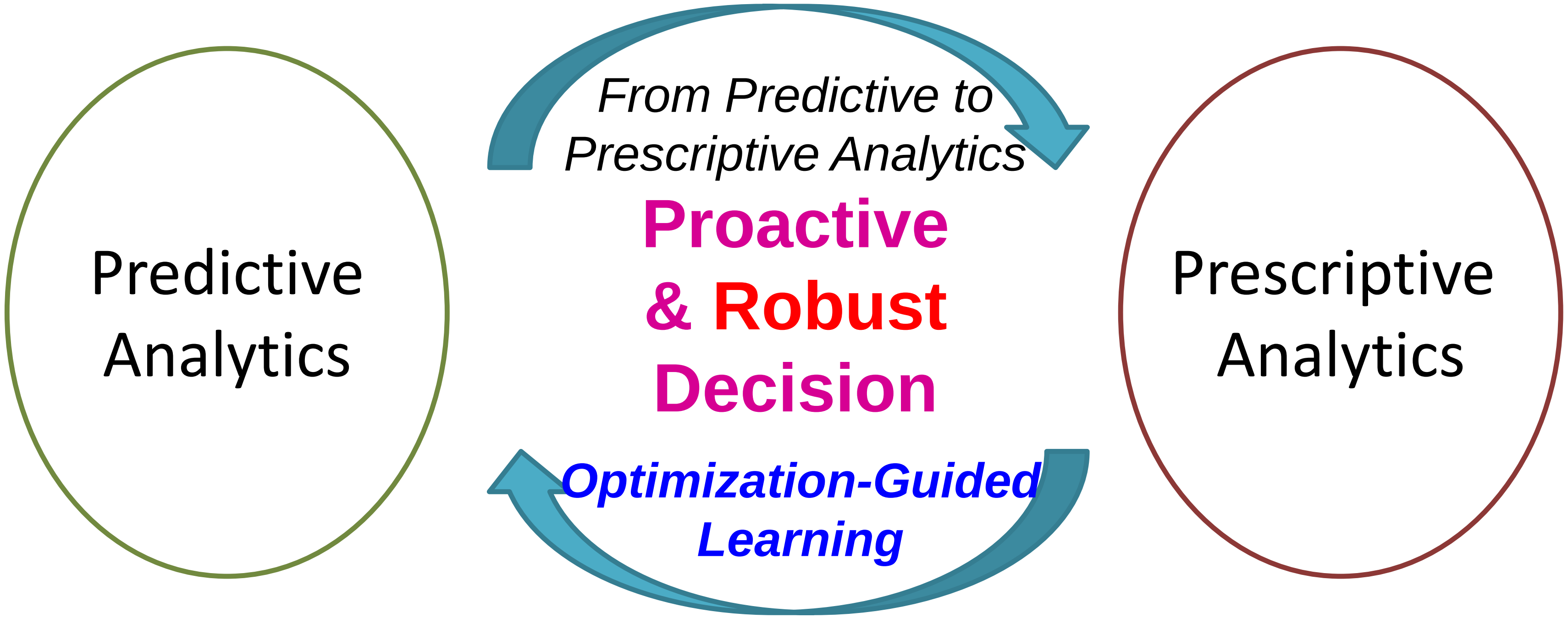
Proactive Decision

Prescriptive Analytics

Portfolio Optimization

Robustness?

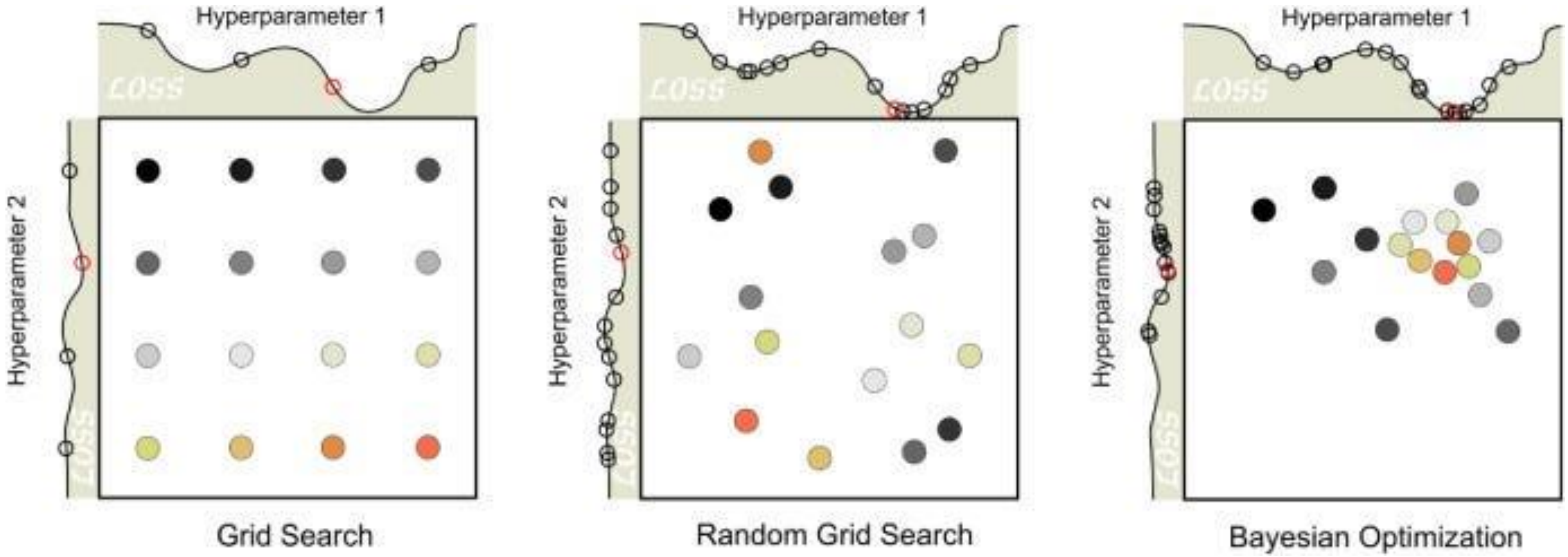
Optimization-Guided Learning (OGL)



- Hyperparameter Optimization in Learning Algorithm
  - Grid Search, Random Search (eg. tabu search, genetic algorithm)
  - Bayesian Optimization; Optimal Computing Budget Allocation (OCBA)
- Optimization-Guided Learning (OGL)
  - Genetic Algorithm embedded with Reinforcement Learning (GAeRL)
  - Reinforcement Learning embedded with Robust Optimization (RLeRO)

Optimization-Guided Learning (OGL)

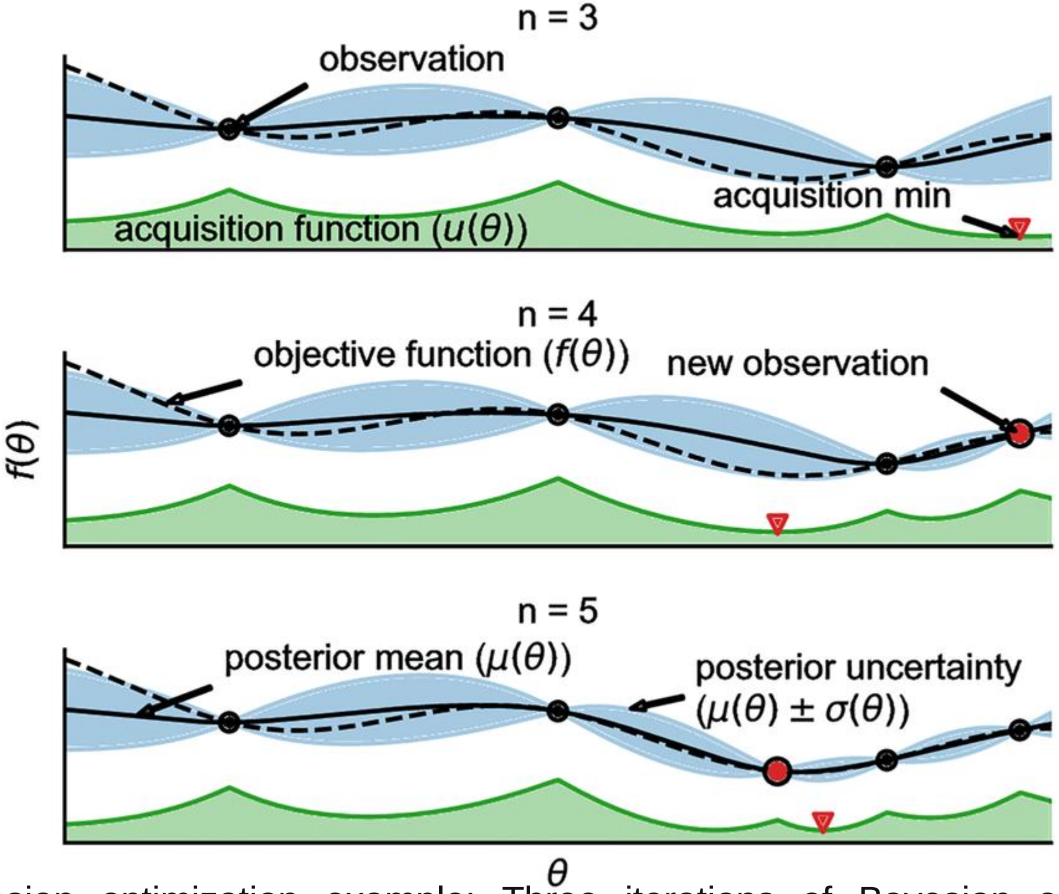
- Hyperparameter Optimization in Learning Algorithm



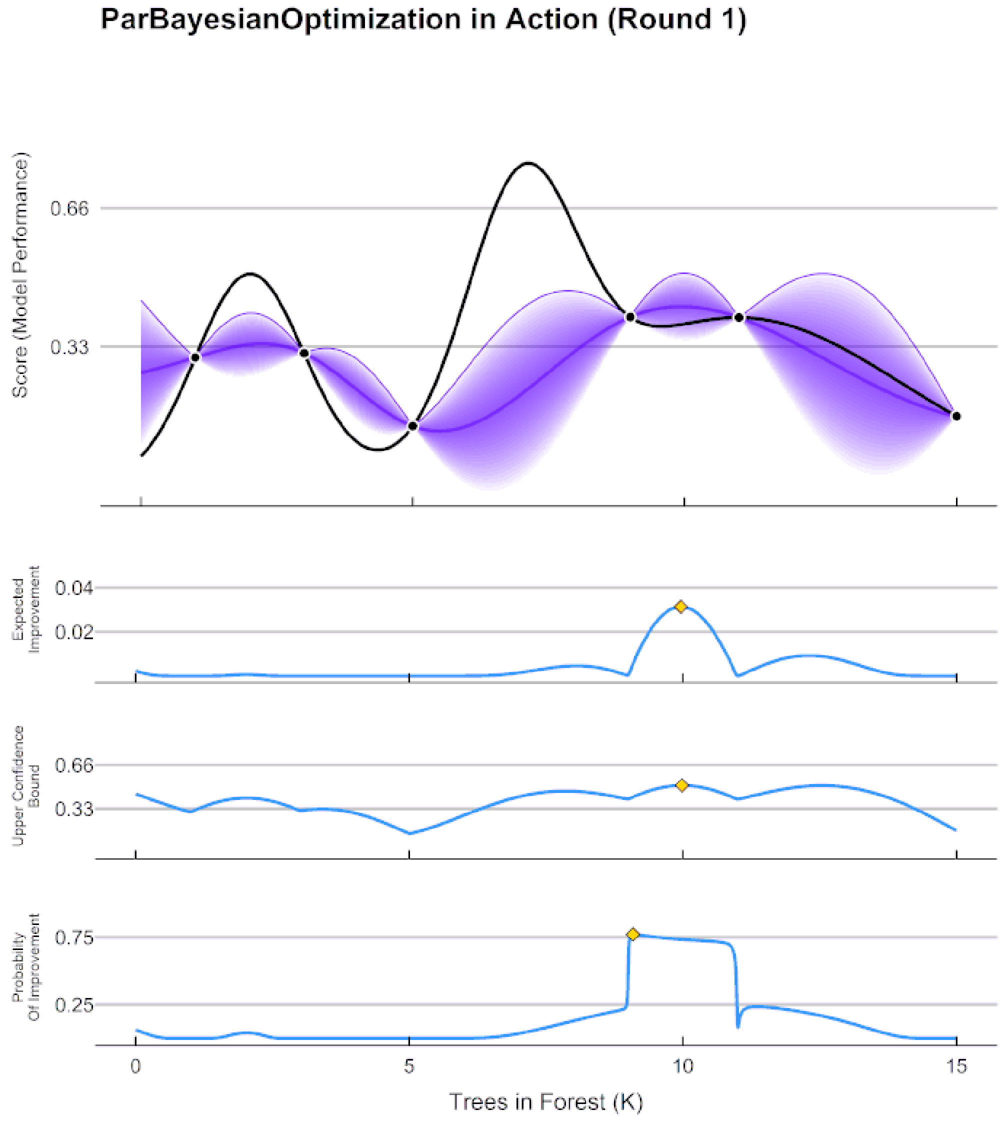
https://pub.aimind.so/understanding-hyperparameter-optimization-techniques-4a39d0494612

Optimization-Guided Learning (OGL)

☐ Bayesian Optimization



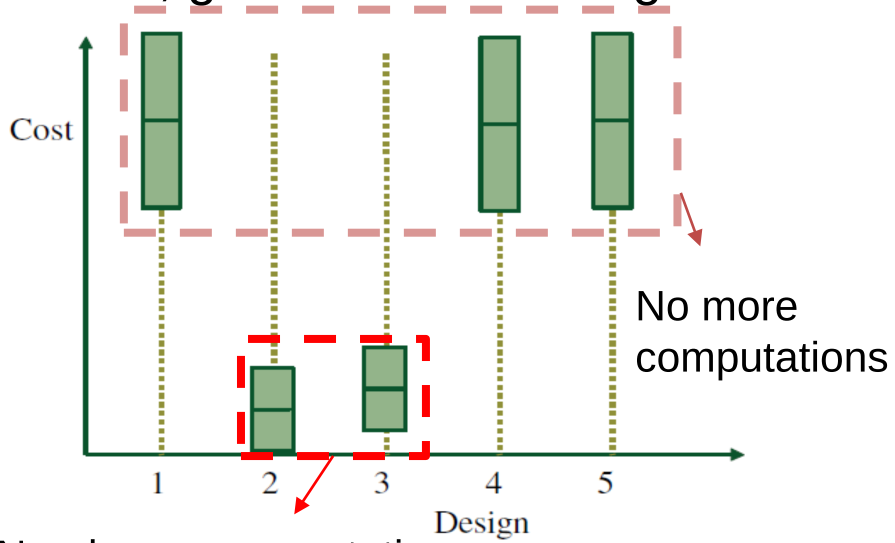
Bayesian optimization example: Three iterations of Bayesian optimization minimizing a 1D function. The figure shows a Gaussian process (GP) approximation (solid black line and blue shaded region) of the underlying objective function (dotted black line). The figure also shows the acquisition function (green). The acquisition function (GP-LCB, lower confidence bound) is the difference of the mean and variance of the GP (multiplied by a constant), which Bayesian optimization minimizes to determine where to sample next.



https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006606
https://en.wikipedia.org/wiki/Bayesian_optimization

Optimal Computing Budget Allocation (OCBA) (Chen and Lee, 2011)

- Given a limited computing power, OCBA finds the best alternative (i.e. parameter design) by maximizing the probability of correct selection (PCS). OCBA maximizes PCS, given a limited budget.



Notations:
K: a set of designs (alternatives)
$b \in K$: the best design
n_i: # of simulation allocated for design $i \in K$
μ_i: the mean of fitness value for design i
T: the total computing budget

OCBA Model

$$\max_{N_1,\ldots,N_{|K|}} PCS = 1 - \underbrace{\textstyle\sum_{i\in K, i\neq b} P\{\tilde{\mu}_b > \tilde{\mu}_i\}}_{\text{Approximate Probability of Correct Selection(APCS)}}$$

$$\text{s.t. } \sum_{i\in K} n_i \le T,$$
$$n_i \ge 0, \forall i \in \mathrm{K}.$$

PCS ≥APCS

Let σ_i be the variance for design i. PCS can be asymptotically maximized when the relationship between two non-best design i and j, where $i \neq j \neq b$, in the lth iteration.

$$\begin{cases} \dfrac{n_i^{l+1}}{n_j^{l+1}} = \left(\dfrac{\sigma_i/(\tilde{\mu}_b - \tilde{\mu}_i)}{\sigma_j/(\tilde{\mu}_b - \tilde{\mu}_j)}\right)^2 \\ n_b^{l+1} = \sigma_b \sqrt{\sum_{i\in K, i\neq b} \left(n_i^{l+1}/\sigma_i\right)^2} \end{cases}$$

# of simulation replications for the best design

OCBA should be the best Ranking and Selection process (Branke et al., 2007)

Chen, C. H. and Loo H. Lee. Stochastic simulation optimization an optimal computing budget allocation. Singapore Hackensack, NJ: World Scientific, 2011.

Optimization-Guided Learning (OGL)

- Genetic Algorithm embedded with Reinforcement Learning (GAeRL)

Lee, C.-Y., Ho, C.-Y., Hung, Y.-H., and Deng, Y.-W., 2024. Multi-objective genetic algorithm embedded with reinforcement learning for petrochemical melt-flow-index production scheduling. ***Applied Soft Computing***, 159, 111630.

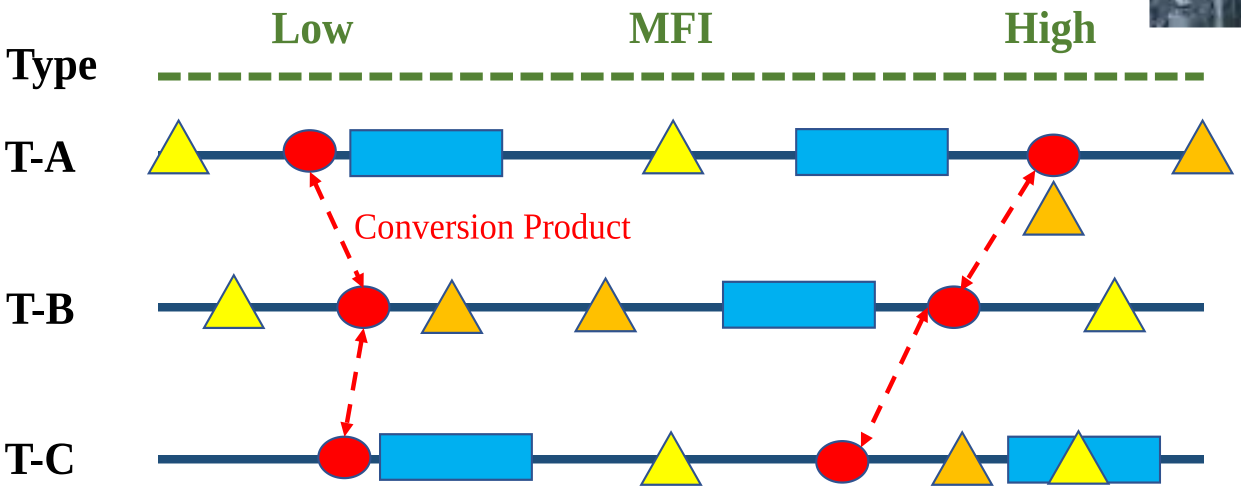
- Reinforcement Learning embedded with Robust Optimization (RLeRO)

Lee, C.-Y., Huang, Y.-T., and Chen, P.-J., 2024. Robust-optimization-guiding deep reinforcement learning for chemical material production scheduling. ***Computers and Chemical Engineering***, 187, 108745.

Genetic Algorithm embedded with Reinforcement Learning (GAeRL)

Petrochemical Production Scheduling (化工廠排程**特性**)

- Objective Functions
  - Minimize tardiness (satisfying due date)
    - Total Tardiness $= \sum_{j \in J} T_j = \sum_{j \in J} \max\{C_j - D_j, 0\}$
  - Minimize # of conversion, transition time, volume of transition product
    - Transition Prodcuts $= \sum_{t \in T} (MFI_{t+1} - MFI_t)^2$
- Constraints
  - Type Conversion Constraint
  - Specific Group Constraint
  - Melt-flow-index (MFI) Slowly-Rise-and-Fall Constraint
  - Sequence-Dependent Transition Time Constraint



Blömer, F., & Günther, H.-O. (1998). Scheduling of a multi-product batch process in the chemical industry. Computers in industry, 36(3), 245-259.
Blomer, F., & Gunther, H.-O. (2000). LP-based heuristics for scheduling chemical batch processes. International Journal of Production Research, 38(5), 1029-1051.

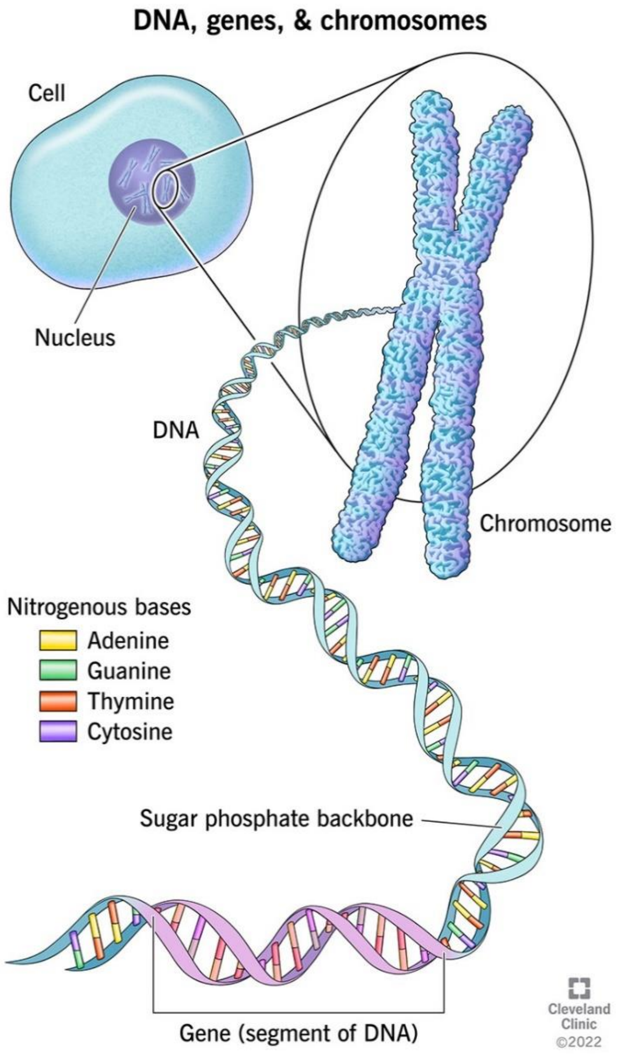
Scheduling Methodology

- JSP is among the hardest combinatorial optimization problems.
  - NP-hard problem
- Heuristic Method (Priority Rule)
  - Shortest processing time (SPT), earliest due date (EDD), etc.
  - Pros: easy to understand
  - Cons: poor performance for complicated production line
- Meta-Heuristic Algorithm (Tabu, Simulated Annealing, **Genetic Algorithm**)
  - Approximated-optimization approach
  - Pros: provide a good solution efficiently
  - Cons: cannot guarantee the global optimum
- Reinforcement Learning
  - Optimal control approach to take actions in a dynamic environment
  - Pros: consider decision over time for dynamic flexible job shop scheduling (DFJSS)
  - Cons: convergence issue in a large state space and action space
- Mathematical Programming
  - Optimization-based approach formulated by mixed integer programming
  - Pros: Guarantee global optimum
  - Cons: computational burden for large-scale problem (not suitable for frequent rescheduling)

Local optimum
Short running time

Global optimum
Long running time

Genetic Algorithm (GA)



Encoding
Chromosomes represent potential solutions

Initial population
Generate a set of chromosomes

Crossover and mutation
Generate offspring from parents

Fitness computation
Evaluate the chromosome quality

Selection mechanism
Solution comparison and random selection

New population

Termination condition

No

Yes

Best solution

Two-point order crossover mechanism

Crossover point

Parent 1	A	B	C	D	E	F	G	H
Parent 2	H	E	D	A	B	C	F	G
Offspring 1	E	F	D	A	B	C	G	H
Offspring 2	H	A	C	D	E	F	B	G

Arbitrary multiple-point shift mutation mechanism

Exchange order C←A←D

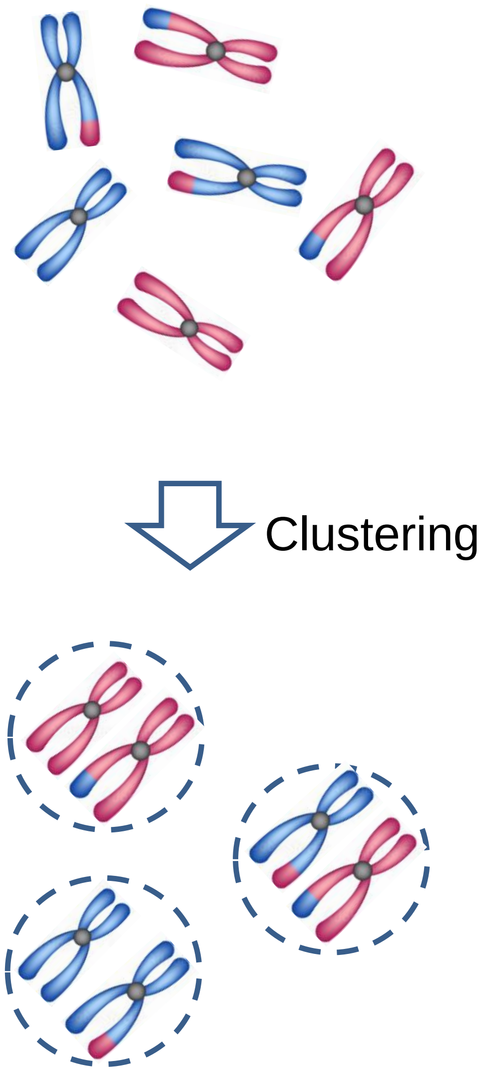
Position	A	B	C	D	E	F
Chromosome	B	D	C	E	A	F
New Chromosome	E	D	B	C	A	F

Holland, J. H. (1975). “Adaptation in Natural and Artificial Systems,” University of Michigan Press, Ann Arbor.

How does RL guide GA?

Population Similarity

- Population similarity: cluster the population with DBSCAN, and then use the Spearman rank correlation coefficient to estimate the correlation within the groups and between the groups.
- Correlation: total - within group = between groups
- Population similarity as the correlation within the groups divided by the number of the clusters.
- Guide the mutation and crossover to balance exploration and exploitation.
  - if chromosomes appear similar, the population might be premature: need exploration and increase mutation and crossover rates
  - if chromosomes appear dissimilar, the population are not converged: need exploitation and decrease mutation and crossover rates.



Phase of Iterations

- Phase of iteration: the number of iterations that GA did not find the better solutions.

State Space in RL

- Discretize the two states by assigning quartiles to each value (i.e., the state space consists of the $4 \times 4 = 16$ states).

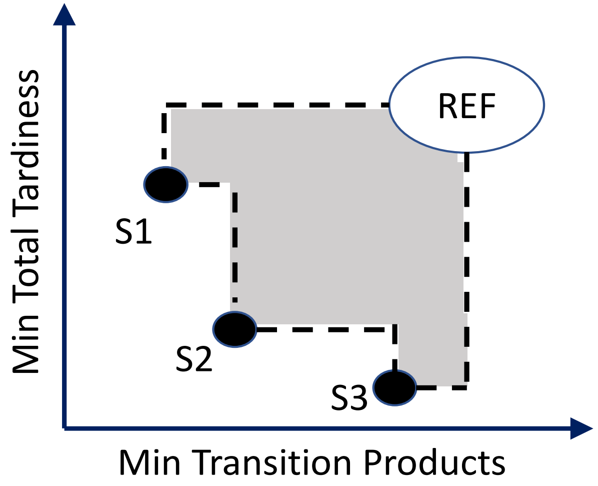
Reinforcement Learning Elements

Action

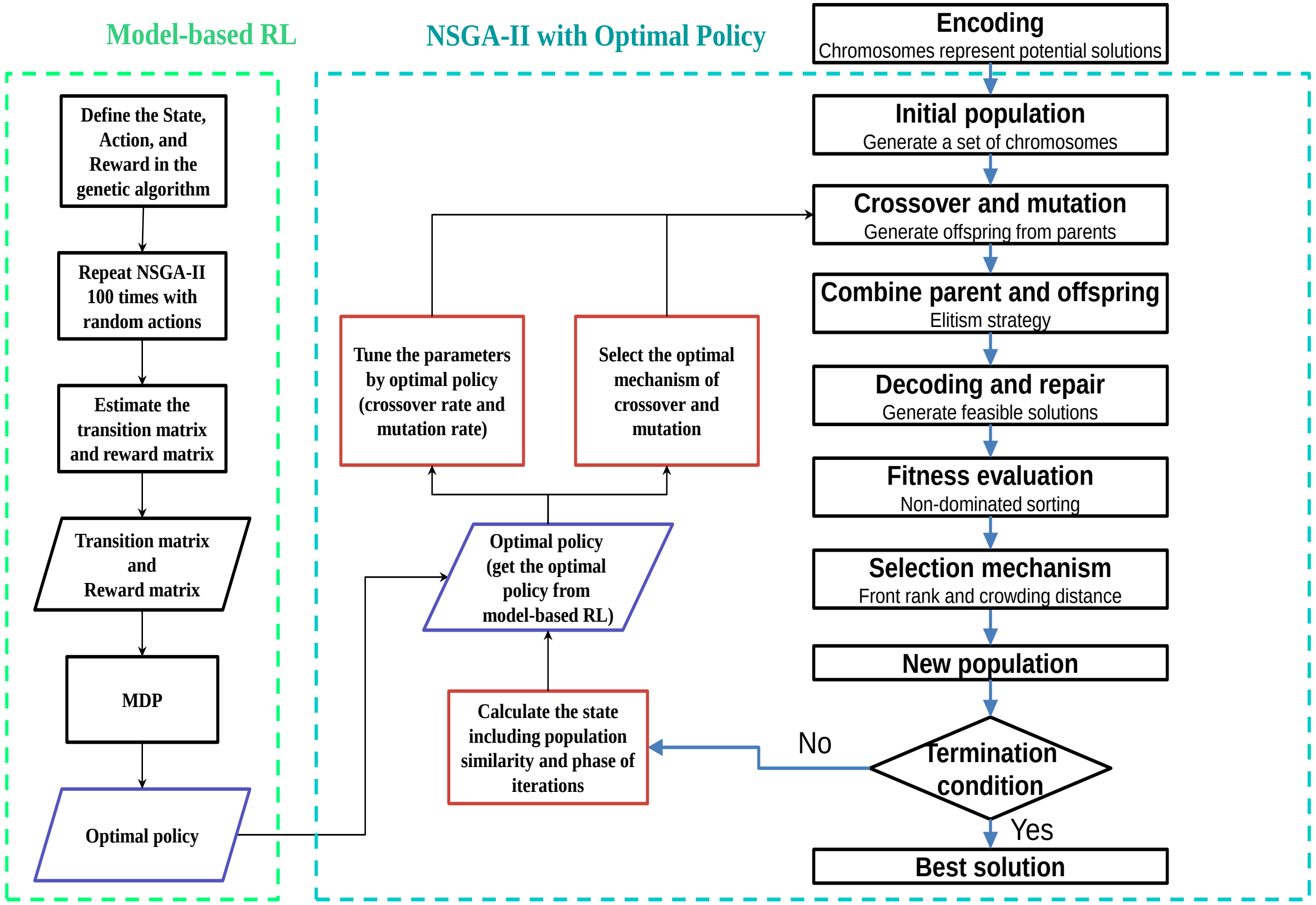
- Rate tuning and mechanism design related to crossover and mutation
- Rate tuning: "fix", "increase", and "decrease" by multiplying the original rate by 1, 1.02, and 0.98.
- Combining the crossover and mutation rates to generate 3×3=9 actions.
- Crossover mechanism: "one-point order crossover", "two-point order crossover", and "position-based order crossover"
- Mutation mechanism: "adjacent two-point change mutation", "arbitrary multiple-point shift mutation", and "shift change mutation".
- Combining the crossover and mutation mechanisms to generate 3×3=9 actions

Reward

- Two-objective reward: minimization of (1) transition products and (2) total tardiness.
- Hypervolume as the volume surrounded by the solutions and a reference point (REF) (i.e., the poorest solution having the highest limit of each objective).
- Goal: maximize the hypervolume: If the present hypervolume indicator is better than last time, the reward is with +1; otherwise, if it is worse than last time, the reward is with -1.



GA embedded with RL (GAeRL)



Pseudocode of GAeRL

```
Begin
    Model-based RL
    While not maximal number of iterations
            Repeat NSGA-II with random action
            Collect data (state, action, reward)
    End
    Estimate the transition matrix and reward matrix
    MDP (value iteration)
    Obtain optimal policy (adjustment of mechanism or rate of crossover and mutation)

    NSGAw/OP
    Input: Optimal policy and initial parameters
            Generate a population from EEH or generate a population randomly
            While not termination condition
                    Generate offspring by crossover and mutation
                    Combine offspring and parent population
                    For each chromosome from offspring and parents
                            Compute the fitness function
                    Non-dominated sorting
                    Select chromosomes based on non-domination front rank and crowding distance
                    Build new population
                    Compute the state according to new population
                    Choose optimal policy according to the state
                    Tune rate or select mechanism of mutation and crossover
            End
            Decode to the petrochemical production schedule
End
```


Lee, C.-Y., Ho, C.-Y., Hung, Y.-H., and Deng, Y.-W., 2024. Multi-objective genetic algorithm embedded with reinforcement learning for petrochemical melt-flow-index production scheduling. *Applied Soft Computing*, 159, 111630.

Empirical Study of Petrochemical Scheduling

Petrochemical Factory

- Leading manufacturer & supplier of polypropylene in Taiwan
  - Product portfolio consists primarily of SBS, SIS, SEBS, and SEP, including compound materials for footwear, modified asphalt, waterproofing membranes, adhesives, and plastics modification.

- Data Source: manufacturing execution system (MES)
- Time: First half of 2019
- Data Size
  - 38 orders including 199 batches, 4 types of catalyst, 4 types of donors, and 12 precedence groups.
  - Transform the data for proprietary information protection without loss of generality.
- Results
  - Reduce transition products in the petrochemical production line by more than 10% through minimizing the change of the Material Flow Index (MFI).
  - It ensures the fulfillment of customer due dates.

Empirical Study of Petrochemical Scheduling

POLab

State Space

- 16 states (combination of 4 levels of similarity and 4 levels in phase due to quantile discretization).

Action Space

- 9 actions (combination of 3 levels in crossover and 3 levels in mutation) with respect to rate tuning and mechanism selection, respectively.

Initial Parameters Settings

- Population size 20, crossover rate 0.8, and mutation rate 0.2.

Action ID	Crossover rate	Mutation rate
R1	Decrease	Decrease
R2	Fix	Decrease
R3	Increase	Decrease
R4	Decrease	Fix
R5 (baseline)	Fix	Fix
R6	Increase	Fix
R7	Decrease	Increase
R8	Fix	Increase
R9	Increase	Increase

Action ID	Crossover	Mutation
M1	One-point order crossover	Adjacent two-point change
M2	Two-point order crossover	Adjacent two-point change
M3	Position-based order crossover	Adjacent two-point change
M4	One-point order crossover	Arbitrary multiple-point shift
M5 (baseline)	Two-point order crossover	Arbitrary multiple-point shift
M6	Position-based order crossover	Arbitrary multiple-point shift
M7	One-point order crossover	Shift change mutation
M8	Two-point order crossover	Shift change mutation
M9	Position-based order crossover	Shift change mutation

POLab

Performance Comparison (with 30 replications)

- Six Algorithms: (1) Engineering Experience Heuristic (EEH), (2) NSGA-II, (3) NSGA-II with random action (NSGAw/RA) for rate tuning of crossover and mutation, (4) NSGA-II with random action (NSGAw/RA) for mechanism selection of crossover and mutation, (5) NSGAeRL for rate tuning, (6) NSGAeRL for mechanism selection. (Note: NSGAw/RA has the same set of actions with NSGAeRL, but with equal probability of selecting actions rather than the optimal policy.)

State (optimal policy)	Similarity-1	Similarity-2	Similarity-3	Similarity-4
Phase-1	**R1**	**R4**	R1	R4
Phase-2	**R8**	R8	R4	R9
Phase-3	R3	R6	R3	**R1**
Phase-4	R5	R4	**R8**	**R4**

State (optimal policy)	Similarity-1	Similarity-2	Similarity-3	Similarity-4
Phase-1	**M7**	**M6**	M1	M6
Phase-2	**M4**	M8	M4	M2
Phase-3	M1	M4	M6	**M8**
Phase-4	M4	M5	**M8**	**M8**

Mean (Standard Deviation)	EEH	NSGA-II	NSGAwRA for Rate Tuning	NSGAwRA for Mechanism Selection	NSGAeRL for Rate Tuning	NSGAeRL for Mechanism Selection
Transition Products	5993 (0)	7247 (950)	7603 (1283)	6974 (952)	6791 (940)	6517 (841)
Total Tardiness	672 (0)	316 (98)	292 (78)	260 (44)	256 (70)	234 (38)
# of Iterations	1 (0)	1287 (490)	872 (463)	1597 (378)	1418 (489)	1767 (292)
CPU Time (second)	3 (0)	1122 (448)	924 (442)	1474 (337)	1466 (483)	1722 (280)

Optimization-Guided Learning (OGL)

- Genetic Algorithm embedded with Reinforcement Learning (GAeRL)

Lee, C.-Y., Ho, C.-Y., Hung, Y.-H., and Deng, Y.-W., 2024. Multi-objective genetic algorithm embedded with reinforcement learning for petrochemical melt-flow-index production scheduling. ***Applied Soft Computing***, 159, 111630.

- Reinforcement Learning embedded with Robust Optimization (RLeRO)

Lee, C.-Y., Huang, Y.-T., and Chen, P.-J., 2024. Robust-optimization-guiding deep reinforcement learning for chemical material production scheduling. ***Computers and Chemical Engineering***, 187, 108745.

Petrochemical Production Scheduling

- Uncertainty
  - Demand fluctuation and yield rate
  - Polyhedral uncertainty sets encode a budget of uncertainty into cardinality constraints.
- Objective function
  - To maximize the gross profit of the chemical production schedule
  - $Max \underbrace{\boxed{\sum_{i\in I}\sum_{p\in P} V_i {A^*}_i x_{ip}}}_{\text{Sales profit}} - \underbrace{\boxed{\sum_{i\in I}\sum_{p\in P} C_i^S s_{ip}}}_{\text{Inventory cost}} - \underbrace{\boxed{\sum_{i\in I} C_i^L l_{ip}}}_{\text{Stockout cost}} - \underbrace{\boxed{\sum_{i\in I}\sum_{j\in I, j\neq i}\sum_{p\in P} C_{ij}^T z_{ijp}}}_{\text{Transition cost}}$
- Constraints
  - State initialization for subproblems

    $s_{iP_0} = S_i^I, \forall i \in I;\ x_{if} = X_{ij}, \forall i \in I, f \in F \cup \{P_0\}$
  - Mass balance constraint

    $s_{ip} = s_{i(p-1)} + A_i^* x_{ip} - D_{ip}^* + l_{ip}, \forall i \in I, p \in P$
  - Production transition identification

    $\sum_{i\in I} z_{ijp} = x_{jp}, \forall j \in I, p \in P$

    $\sum_{j\in I} z_{ijp} = x_{i(p-1)}, \forall i \in I, p \in P$
  - Machine occupancy constraint

    $\sum_{i\in I} x_{ip} = 1, \forall p \in P$
  - Variable domains

    $x_{ip} \in \{0,1\}, \forall i \in I, p \in P \cup \{P_0\}$

    $z_{ijp} \in \{0,1\}, \forall i \in I, j \in I, p \in P$

    $s_{ip} \geq 0, \forall i \in I, p \in P \cup \{P_0\}$

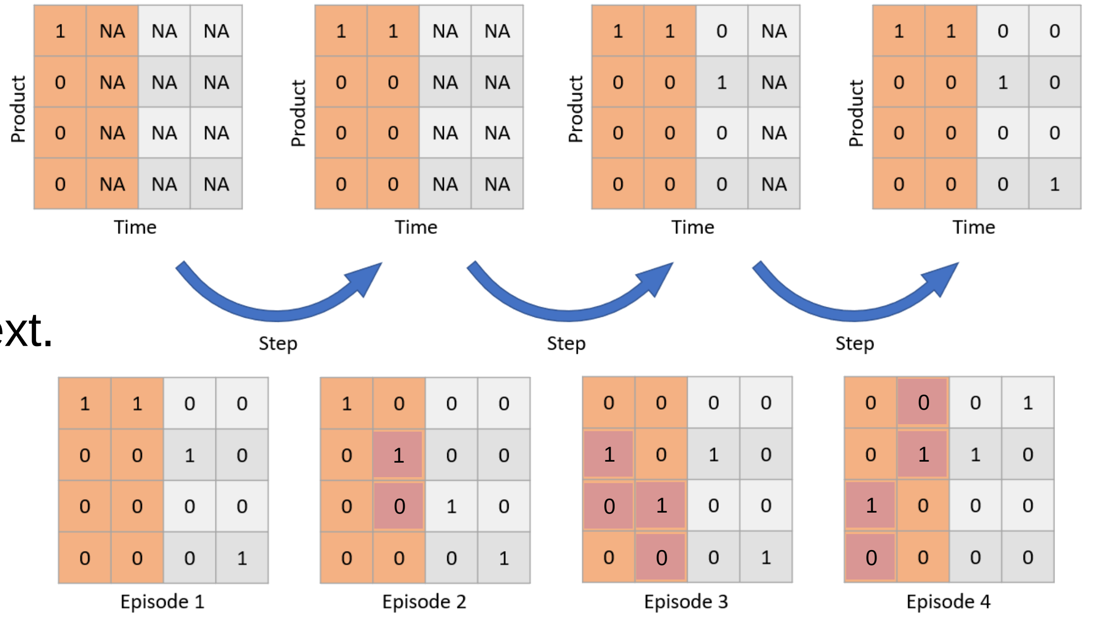
    $l_{ip} \geq 0, \forall i \in I, p \in P$

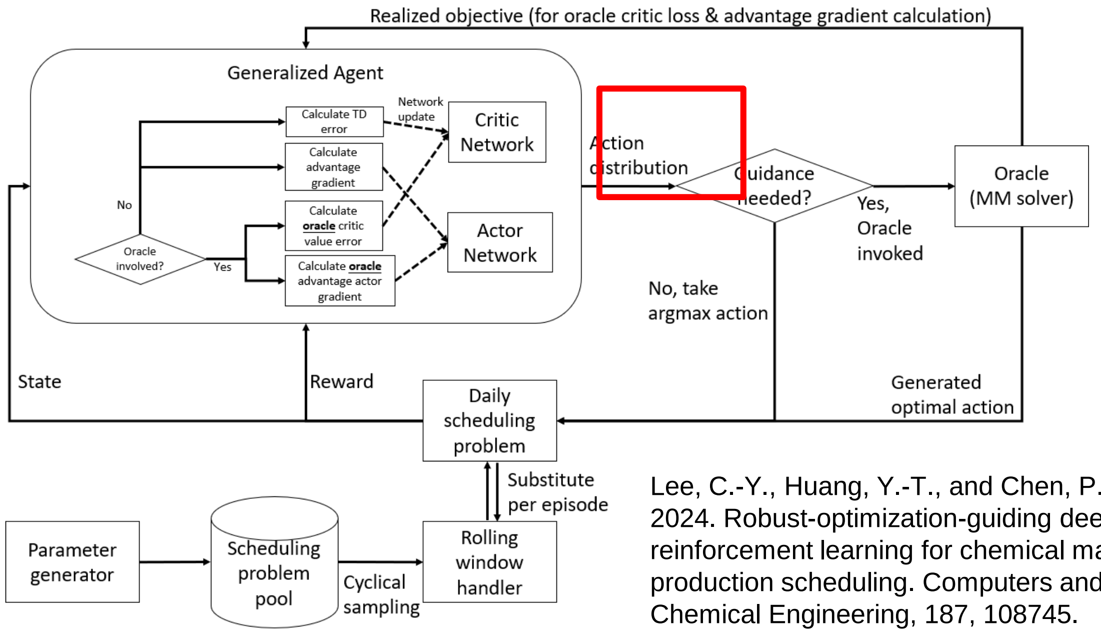
Reinforcement Learning embedded with Robust Optimization (RLeRO)

- **Action**
  - Network output discrete probability distribution A
  - Action a is sampled from A, corresponding to x_{ip}.
- **Episode**
  - Finish a complete scheduling window and rolling to the next.
- **Reward**
  - The change in objective value after a particular action
- **State encoding**

$$state_p = (\underbrace{\boxed{l_{ip}}}_{\text{Inventory level}}, \underbrace{\boxed{x_{ip}}}_{\text{Determined schedule}}, \underbrace{\boxed{\tilde{D}_{ip}}}_{\text{Predicted demand}}, \underbrace{\boxed{\tilde{A}_i x_{ip} + l_{ip} - \tilde{D}_{ip}}}_{\text{Estimated stockout}}, \underbrace{\boxed{t}}_{\text{Time counter}})$$

- **Methods**
  - Perfect information deterministic optimization (PIDO)
  - Expected value deterministic optimization (EVDO)
  - Robust optimization (RO)
  - Advantage-Actor-Critic (A2C)
  - A2C + EVDO guiding
  - A2C + RO guiding



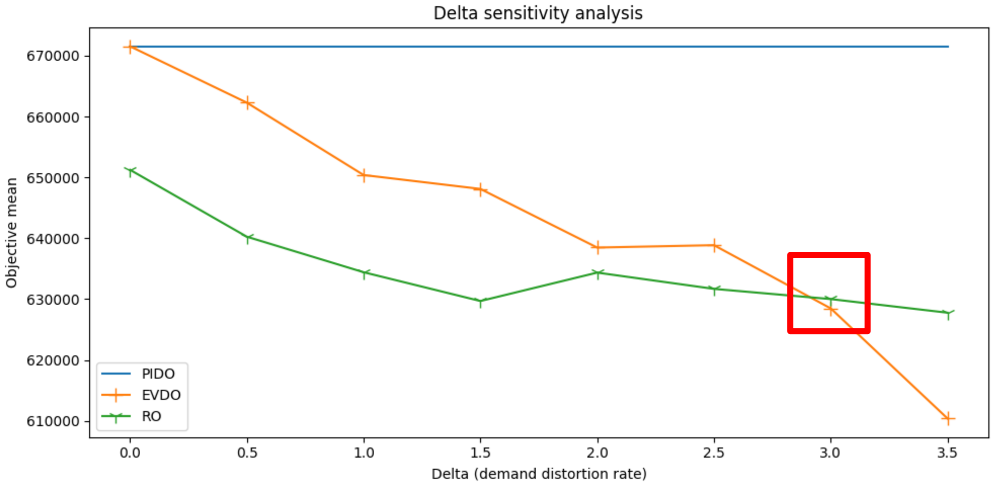


Lee, C.-Y., Huang, Y.-T., and Chen, P.-J., 2024. Robust-optimization-guiding deep reinforcement learning for chemical material production scheduling. Computers and Chemical Engineering, 187, 108745.

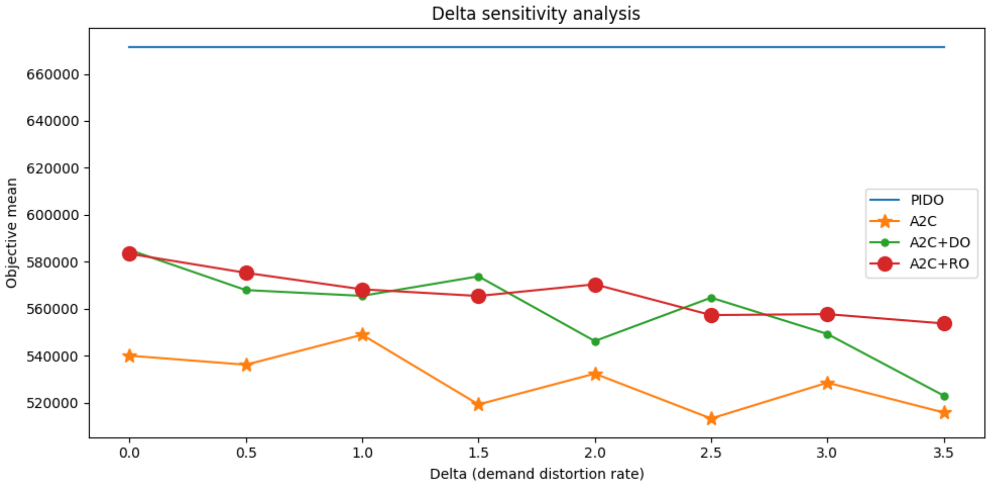
Reinforcement Learning embedded with Robust Optimization (RLeRO)

Sensitivity Analysis

- Optimization-based models

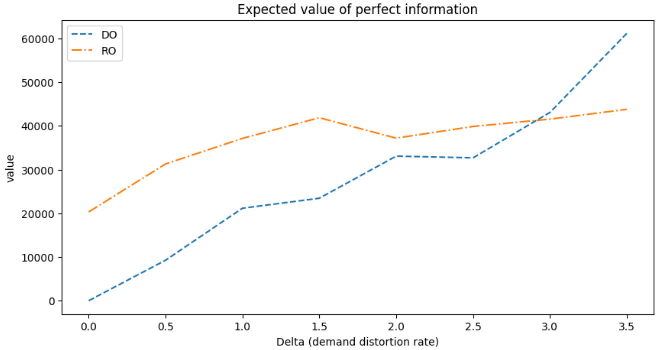


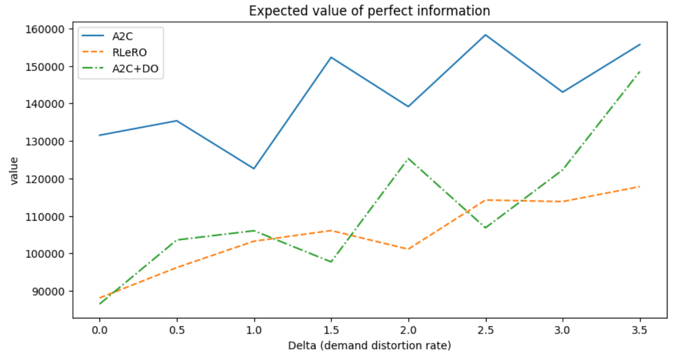
- RL-based models



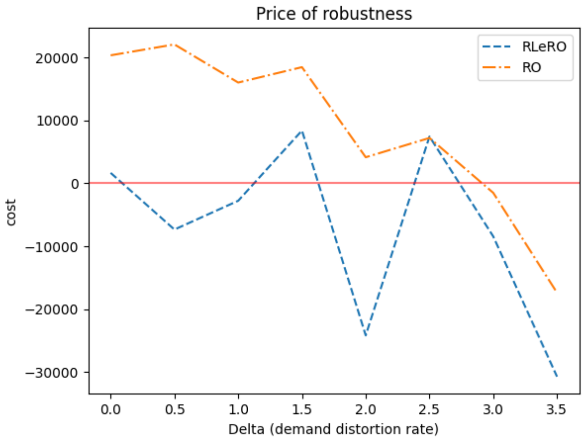
Solution Value Analysis

- Expected value of perfect information (EVPI)
  - "*How much a decision-maker would be willing to pay for perfect information when using the model*"
  - subtract the PIDO value from the target model's
  - The robust models are less needed of perfect infor.





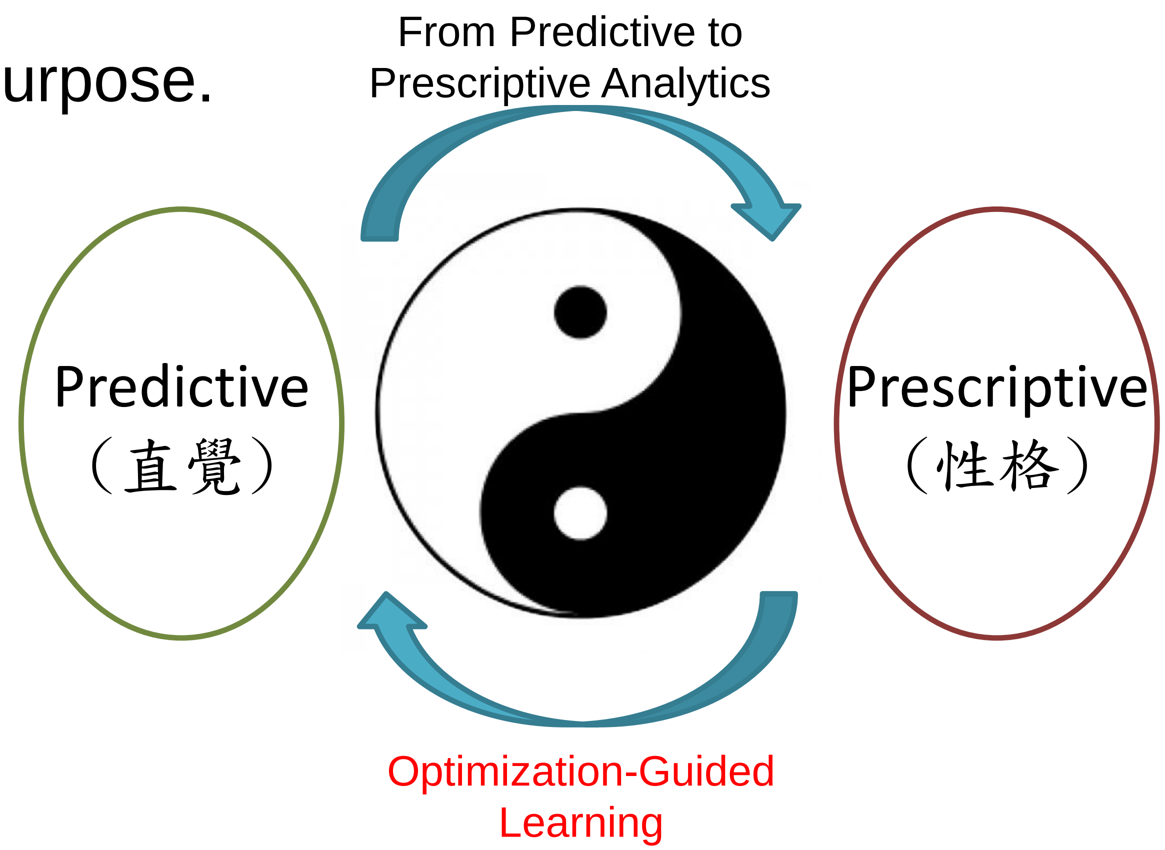
- Price of robustness
  - Distance of objective between the baseline (EVDO) and robust solution
  - "*How much it cost to apply robust solutions*"
  - Adopt conservative policies in high demand fluctuation.



Takeaway

POLab

- Prediction is the Process; Decision is the Purpose.
- From Predictive to Prescriptive Analytics
- Optimization-Guided Learning (OGL)
  - Find the "unchanged" power in "changing" env.
  - Make the learning system more "stable".



Light speed in vacuum (真空光速): 299,792,458 m/s
Planck constant (普朗克常數): $6.62607015 \times 10^{-34}$ J·s
Electron mass (電子質量): $9.10938291 \times 10^{-31}$ kg
Avogadro constant (亞佛加厥常數): $6.02214076 \times 10^{23}$ mol^{-1}
Boltzmann constant (波茲曼常數): $1.38064852 \times 10^{-23}$ J/K
Gravitational constant (重力常數): 6.67384×10^{-11} m^3/(kg · s^2)

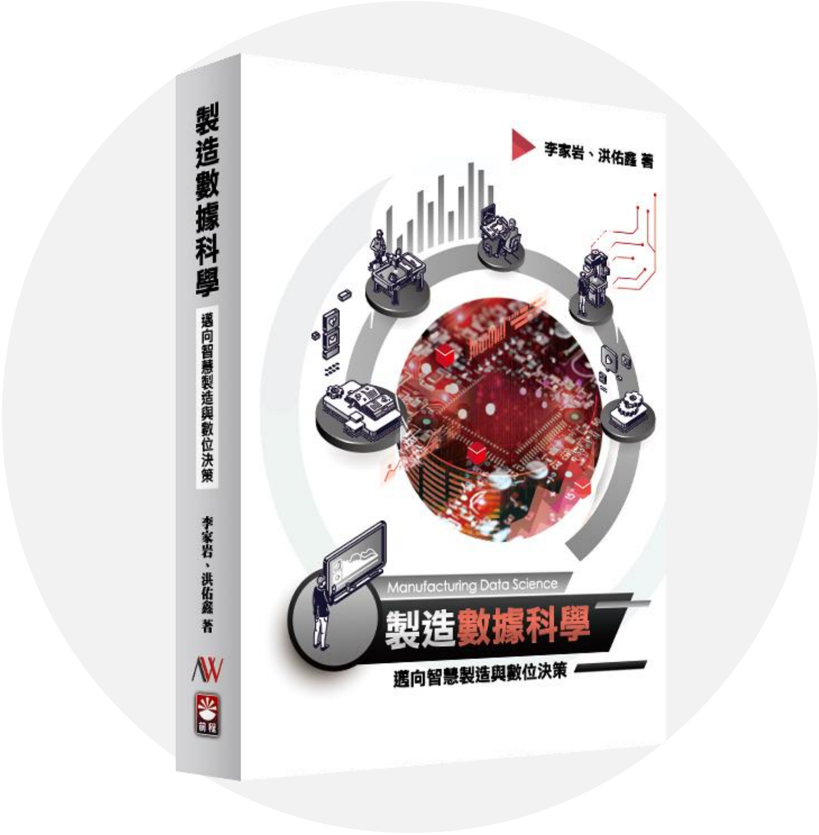
We can observe them, but cannot change them.
我們只能觀察到，但不能改變他們

https://www.youtube.com/watch?v=oBVIn2PFTYM

謝謝大家
還請多多指教


Chia-Yen Lee


ALMA MATER STUDIORUM
A.D. 1088
MATER STUDIORUM
DI BOLOGNA


李家岩、洪佑鑫 著
Manufacturing Data Science
製造數據科學
邁向智慧製造與數位決策