2020.11.14 台灣人工智慧年會

Beyond the Prediction
(預測之外：跨越預測與決策間的鴻溝)

Dr. Chia-Yen Lee (李家岩 博士）

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

經常在**企業**與**學校**從事教育工作的這些年…

有數據就有AI？

好像對、但也不完全對…

數據是AI的”必要條件”之一，但非充分條件

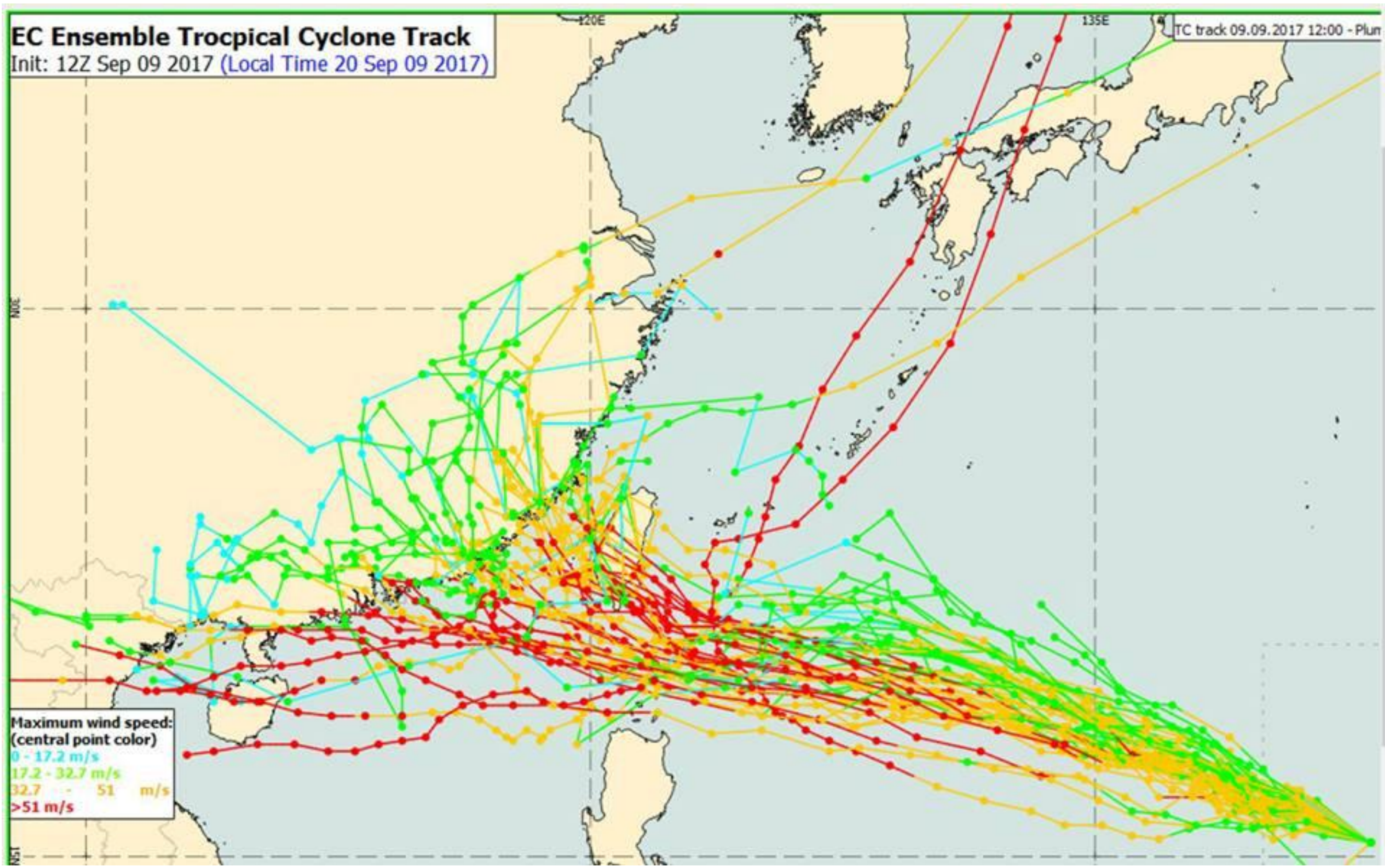
實務上常發生的是...

- In reality, if you provide the ML model to BOSS..
  - Engineer: I built a XXX model with excellent prediction accuracy XXX
  - Boss: mmm….ok….so……
    - (Too good to be true?….)
    - (Counterintuitive…I should believe you (AI), or…believe myself?...)
  - Boss: OK... Maybe we try this way…Please simulate some scenarios/experiments and let me see the results…

My proposed model is excellent!

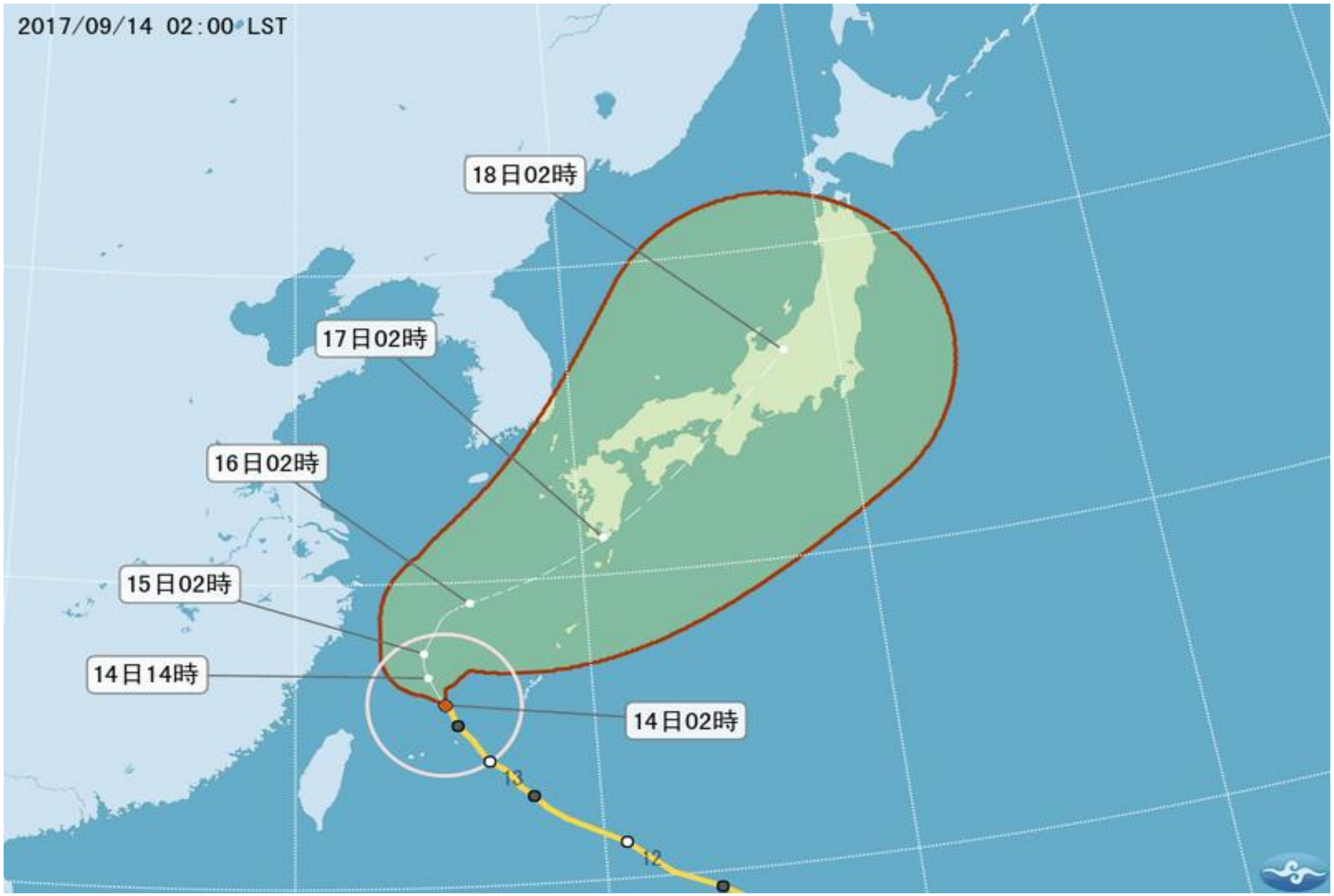
Scenario Analysis

- 2017/09/09 European Centre for Medium-Range Weather Forecasts (ECMWF) showed 50 predictive scenarios of typhoon Talim..



http://news.ebc.net.tw/news.php?nid=77898

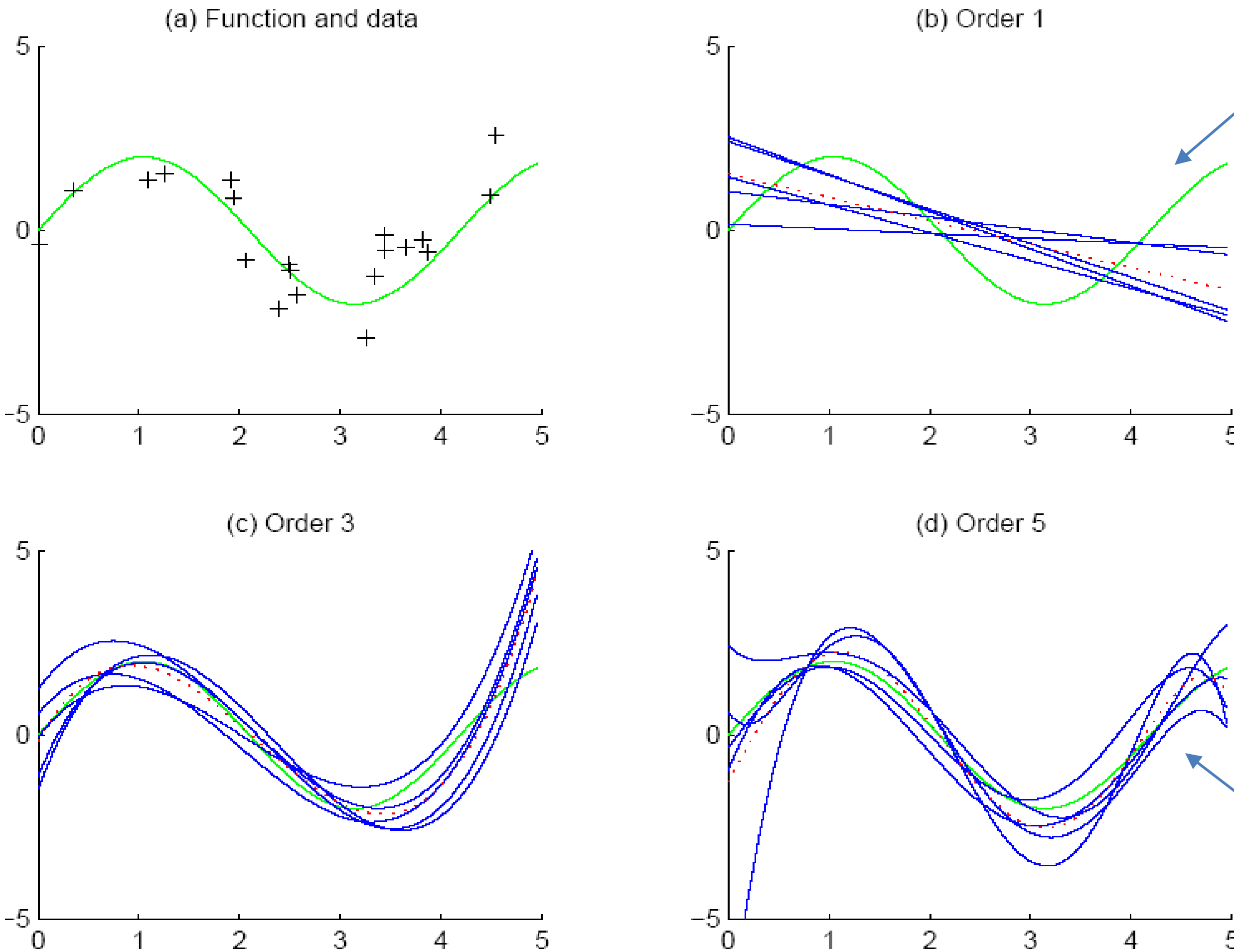
2017/09/14 typhoon Talim showed "hairpin bend" toward the north…



http://www.cna.com.tw/news/firstnews/201709140012-1.aspx

So…the insight is…

- Bias/Variance Dilemma $\mathrm{E}\left[\left(y-\hat{f}(x)\right)^2\right]=\mathrm{Bias}\left[\hat{f}(x)\right]^2+\mathrm{Var}\left[\hat{f}(x)\right]+\sigma^2$



Boss's intuition

Overfitting?

Engineer's efforts

Alpaydin (2010)

- A powerful model against Boss's intuition, then… Fail…

In fact…

There always exists a BIG GAP between Prediction and Decision

工程師 老闆

Analytics(構析) Framework

POLab

Data Description

- Data Source & Goal
- Summary & Visualization
- Load Packages
- Problem Clarification

Data Preprocessing

- Redundant & Missing Value Imputation
- Transform & Rescale
- Feature Engineering
- Data Augmentation

Feature Selection

- Filter Methods
- Wrapper Methods
- Embedded Methods
- Dimension Reduction
- Physical Causality

Modeling & Validation

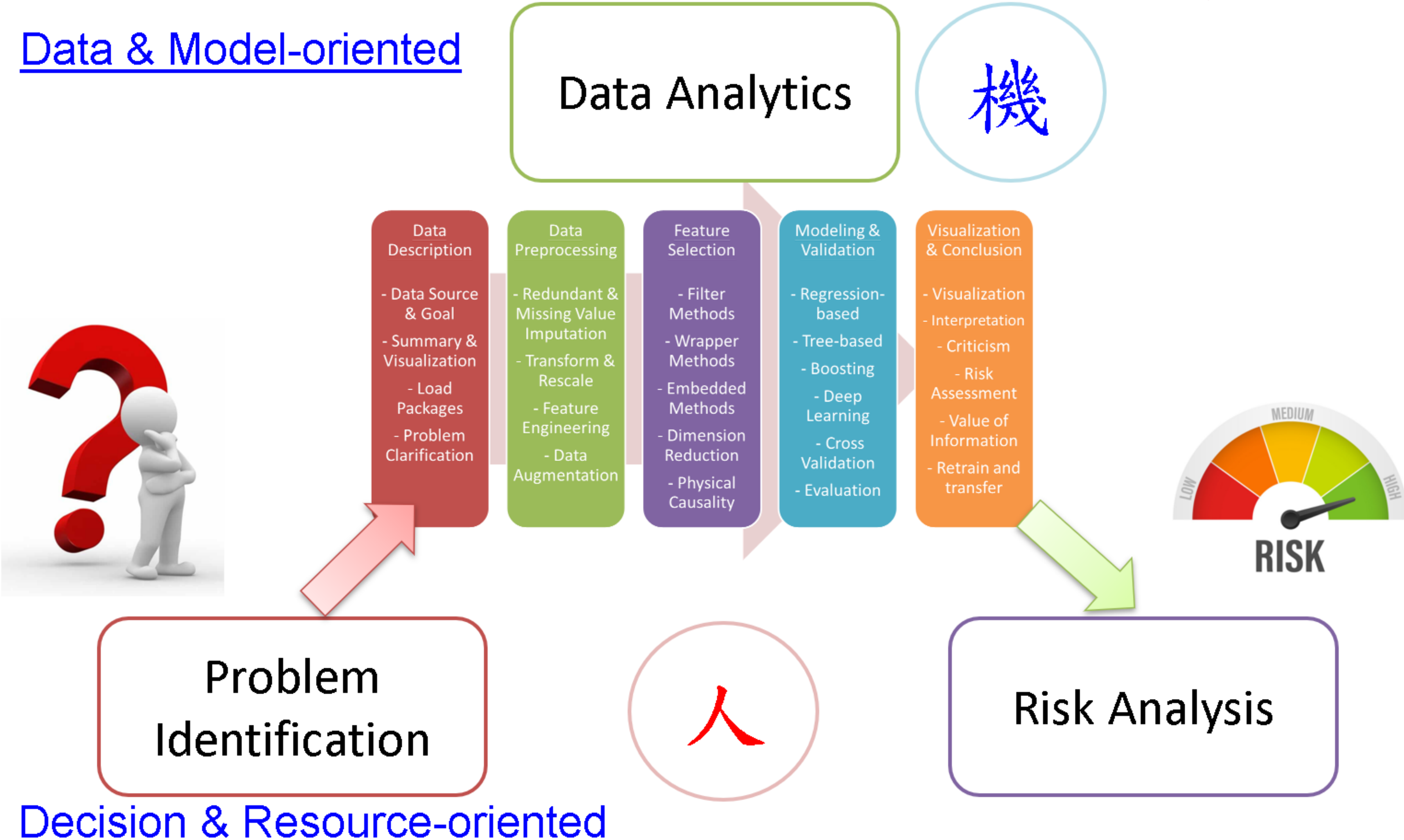
- Regression-based
- Tree-based
- Boosting
- Deep Learning
- Cross Validation
- Evaluation

Visualization & Conclusion

- Visualization
- Interpretation
- Criticism
- Risk Assessment
- Value of Information
- Retrain and transfer

以視覺化起，也以視覺化終

Analytics(構析) Framework



Beyond the Prediction…

問題與因果

1. Problem Identification
2. Causality

批判與資訊價值

3. Criticism
4. Value of Information

風險與決策

5. Risk Assessment
6. Prescriptive Analytics

適應與擴充

7. Concept Drift
8. Domain Adaptation

Beyond the Prediction…

問題與因果

1. Problem Identification
2. Causality

批判與資訊價值

3. Criticism
4. Value of Information

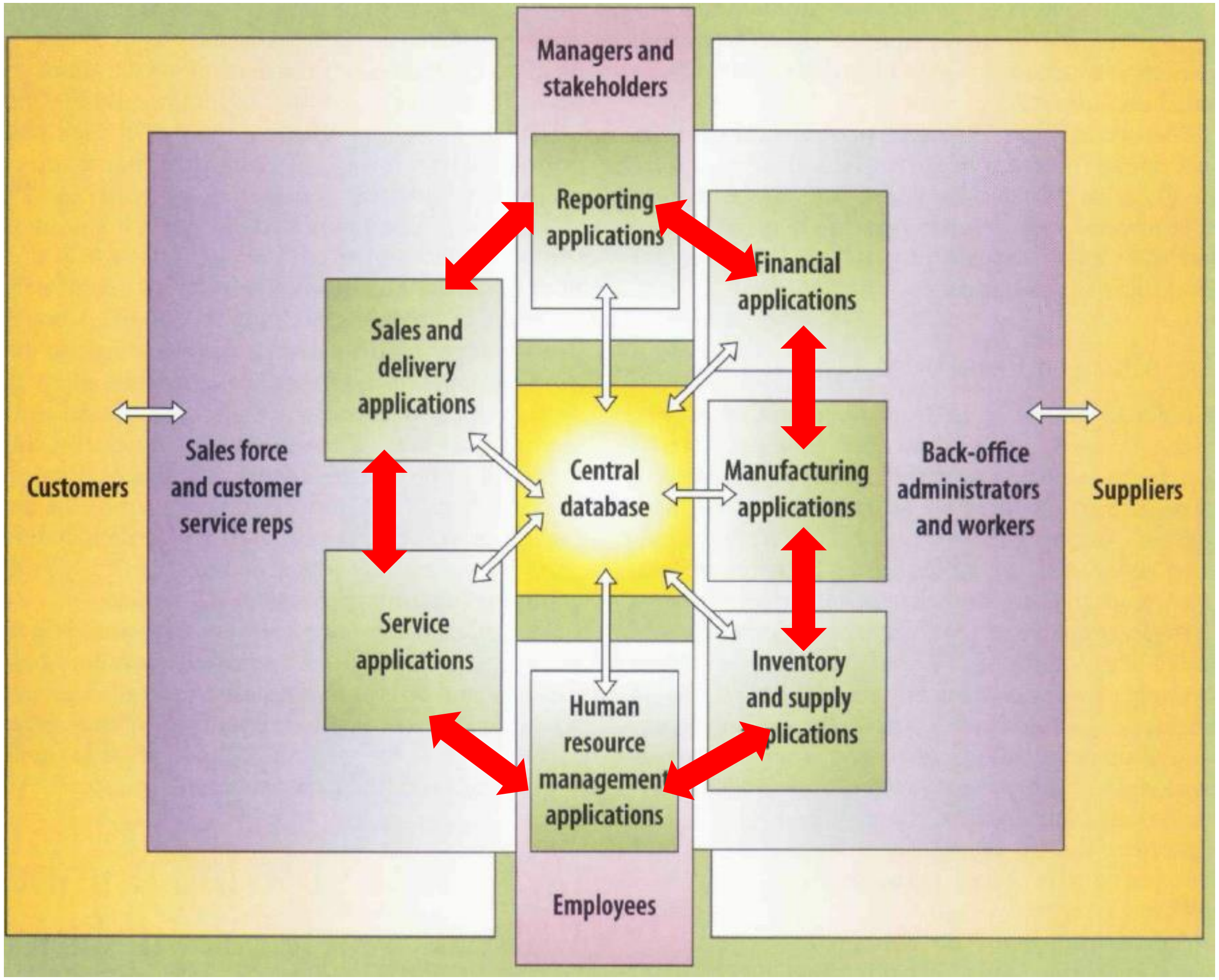
風險與決策

5. Risk Assessment
6. Prescriptive Analytics

適應與擴充

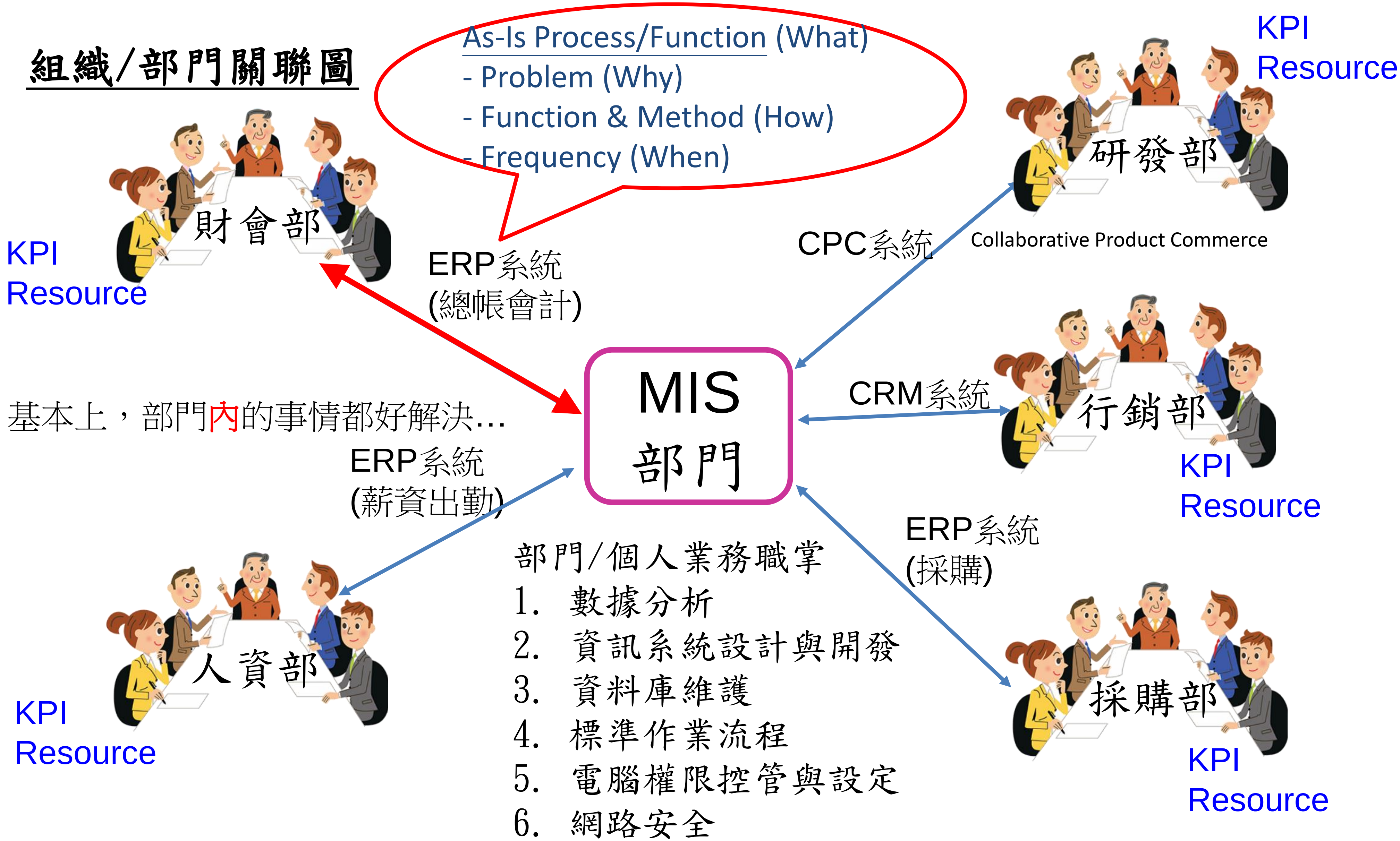
7. Concept Drift
8. Domain Adaptation

- 怎麼找問題？擊破點？如何從點影響到面？
  - 過去企業功能的水平展開，長期下來造成管理碎片化...



Davenport, T. H. (1998). Putting the enterprise into the enterprise system. Harvard Business Review, July-August, 121-131.

問題與因果



https://slidesplayer.com/slide/11473184/

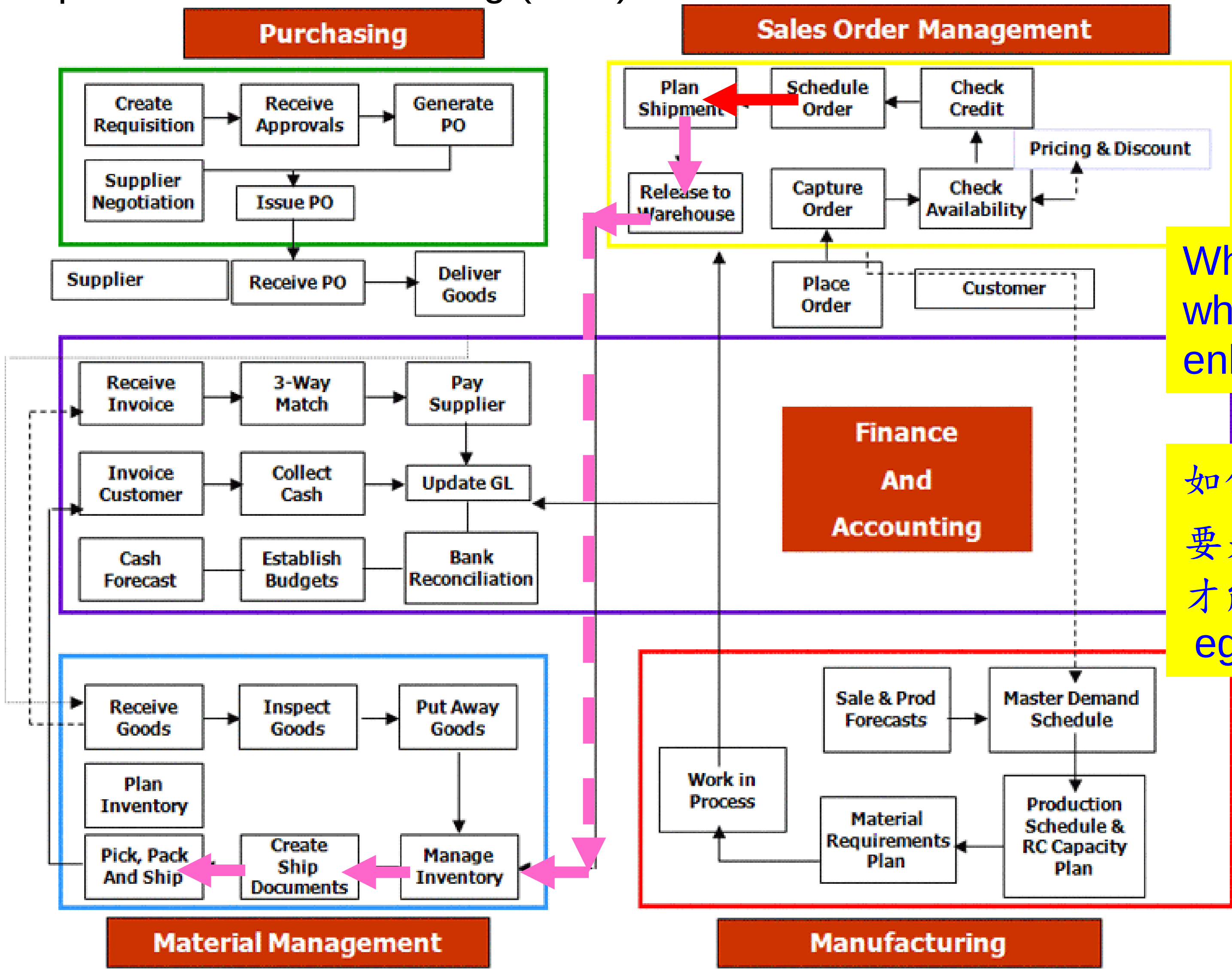
問題與因果

跨部門業務檢核表 (cross-dept. work checklist)

我所屬 <u>工程</u> 處 ______ 部門/課		
與他 <u>品保</u> 部門	As-Is (文字或作圖)	To-Be AI (文字或作圖)
流程業務 (Who/What) 欲改善的業務與流程為何？資源與決策關係人為何？	流程(針對某一業務/功能)： 人工瑕疵檢測 資源(會牽涉到什麼資源)： 技術員、電子顯微鏡…	AOI瑕疵檢測 機台、大數據影像資料庫…
頻率與重要性 (When) 該業務發生的頻率與重要性為何？	頻率(該業務發生頻率)： 高(一分鐘XXXX顆檢查) 重要性(該業務的重要性)： 高(關鍵XXX製程站點品檢)	高(一分鐘可達XXXX顆檢查) 高(關鍵XXX製程站點品檢)
問題 (Why) 目前該業務有什麼問題、困難與挑戰？	問題(目前遇到什麼問題)： 人為誤判率高 問題的問題(根本原因在哪)： 誤判率高的理由是品檢員判斷不一、長時間眼睛疲勞…	AOI可能overkill率高 由於影像數據還在收集中，AI模型訓練仍加強
功能與方法 (How) 目前該業務處理的方法或提供什麼功能？	功能/方法(怎麼做這業務)： 人工操作電子顯微鏡用肉眼觀察量測 投入/輸出(功能投入/輸出)： 投入：技術員、電子顯微鏡 輸出：良品/不良品(bin code)	用AI光學影像自動判斷 投入：光學檢測機 輸出：良品/不良品(bin code)
資料 (Where) 業務中會用到的分析數據從哪裡收集？	資料(欄位名稱與資料來源)： 缺陷影像、FDC數據根因分析 資料品質(資料可靠與否)： 根據機台或光學影像解析度而異，數據跟抽樣頻率有關	檢測員bin code判斷影像與結果、工程師複檢判斷結果差(尤其隨機defects的狀況人工不易歸類分bin)
績效KPI改善與衝突點(Conflict) 透過To-Be改善自己部門的KPI，是否會造成該業務相關部門他們KPI的惡化	我方：工程處：檢測員誤判率高(false alarm)、分bin錯誤率高、需要複判且時間長 對口： 品保處：「寧可錯殺一百，不可放過一人」	工程處：AOI可協助找出製程根因來源、縮短故障排除時間 品保處：AOI使人力成本降低

怎麼找問題？擊破點？如何從點影響到面？

Enterprise Resource Planning (ERP)



Where is the process which can be enhanced by AI?

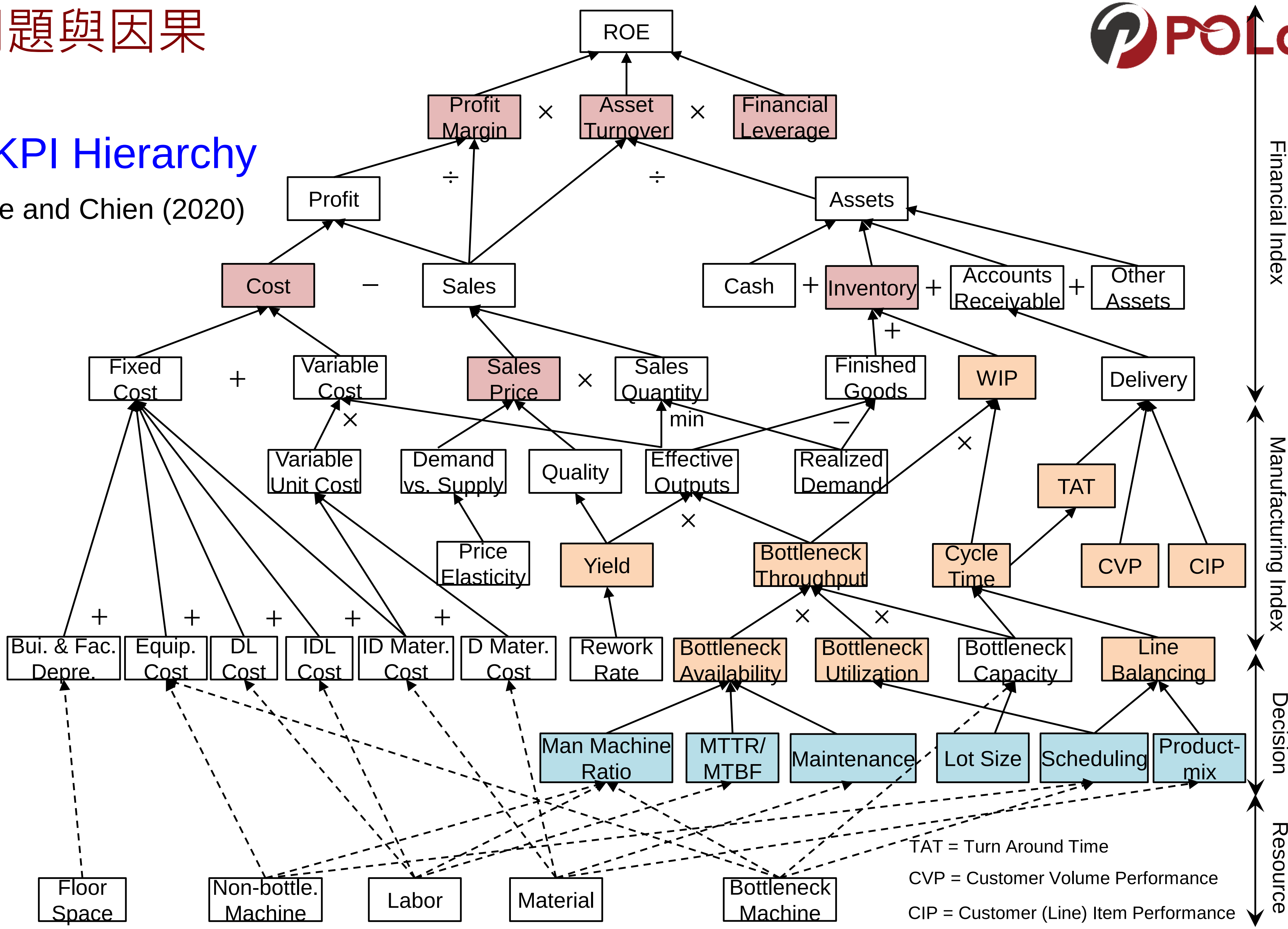
如何從點連到面的影響？
要先有**流程地圖**！
才能"定位"問題
eg. ERP map

http://www.instep.co.th/innova/solution_erp.aspx

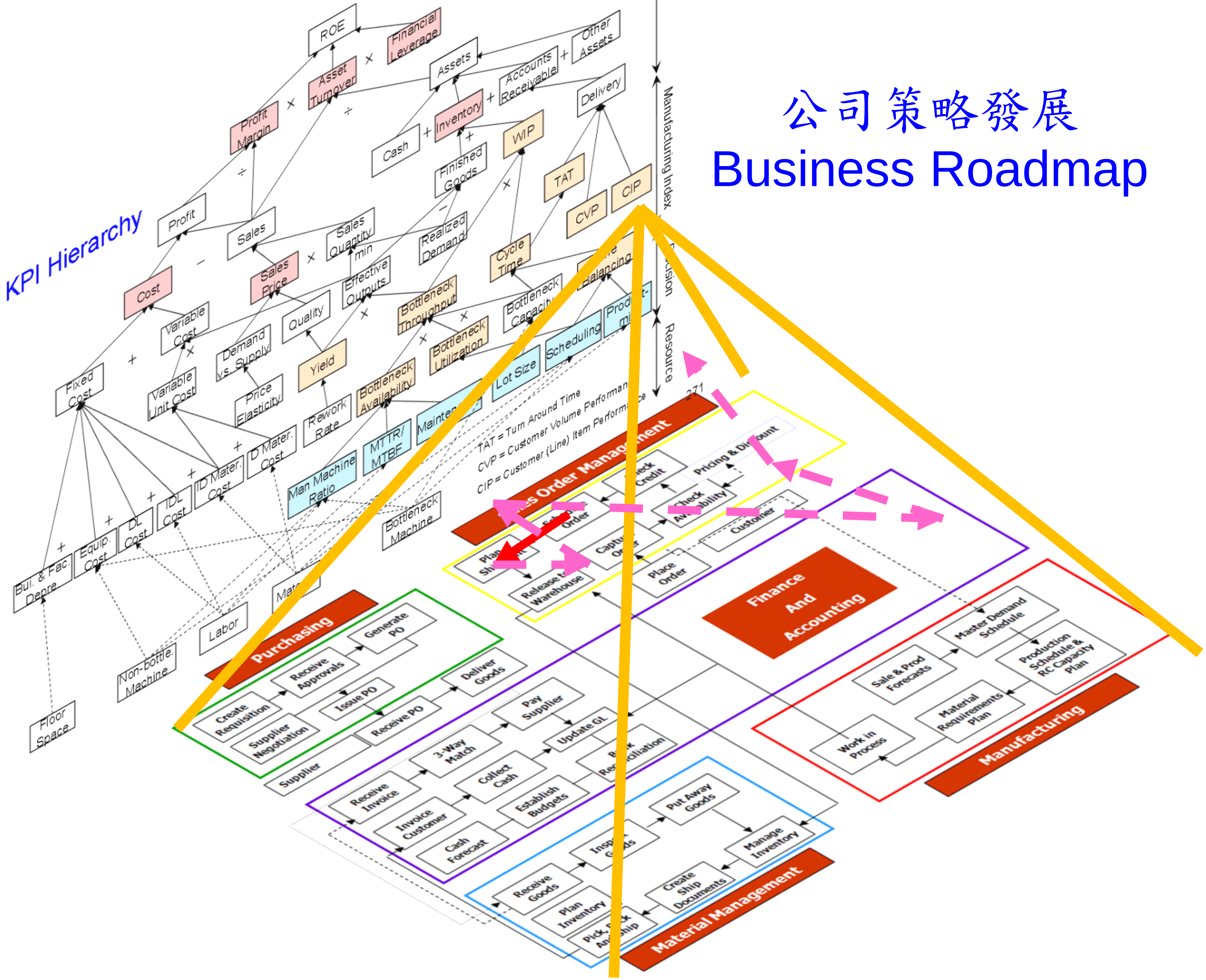
問題與因果

KPI Hierarchy

Lee and Chien (2020)



問題與因果

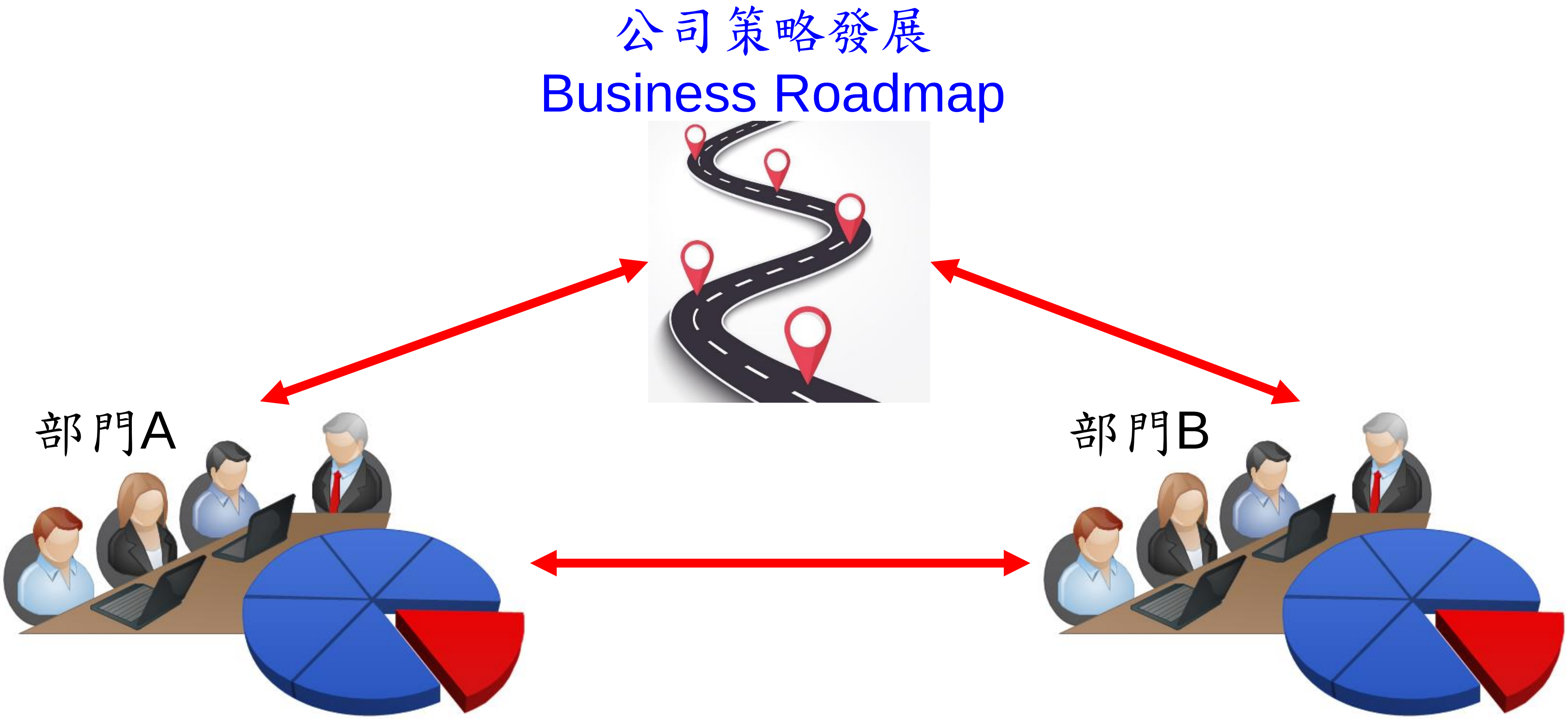


公司策略發展

Business Roadmap

問題與因果

- 怎麼找問題？擊破點？如何從點影響到面？(三要)
  - 必要：處理部門內的事 → 點
  - 需要：處理部門間的事 → 線
  - 想要：處理部門與公司策略發展的連結 → 面
    - 現在的專案執行，在地圖上所帶來的「連鎖效應」，最終如何達成公司策略目標



Troubleshooting...

什麼樣的因導致這個問題(果)呢？

其實...
ML/DS在這件事情上...苦手了...

本來希望用ML/DS來確認因果關係
但卻造成了更多問題...(因果釐清問題)

□ 48 Journal Editor-in-Chiefs…

PERSPECTIVE

SPECIAL SECTION

Control of Confounding and Reporting of Results in Causal Inference Studies

Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals

David J. Lederer[1,2*], Scott C. Bell[3*], Richard D. Branson[4*], James D. Chalmers[5*], Rachel Marshall[6*], David M. Maslove[7*], David E. Ost[8*], Naresh M. Punjabi[9*], Michael Schatz[10*], Alan R. Smyth[11*], Paul W. Stewart[12*], Samy Suissa[13*], Alex A. Adjei[14], Cezmi A. Akdis[15], Élie Azoulay[16], Jan Bakker[17,18,19], Zuhair K. Ballas[20], Philip G. Bardin[21], Esther Barreiro[22], Rinaldo Bellomo[23], Jonathan A. Bernstein[24], Vito Brusasco[25], Timothy G. Buchman[26,27,28], Sudhansu Chokroverty[29], Nancy A. Collop[30,31], James D. Crapo[32], Dominic A. Fitzgerald[33], Lauren Hale[34], Nicholas Hart[35], Felix J. Herth[36], Theodore J. Iwashyna[37], Gisli Jenkins[38], Martin Kolb[39], Guy B. Marks[40], Peter Mazzone[41], J. Randall Moorman[42,43,44], Thomas M. Murphy[45], Terry L. Noah[46], Paul Reynolds[47], Dieter Riemann[48], Richard E. Russell[49,50], Aziz Sheikh[51], Giovanni Sotgiu[52], Erik R. Swenson[53], Rhonda Szczesniak[54,55], Ronald Szymusiak[56,57], Jean-Louis Teboul[58], and Jean-Louis Vincent[59]

[1]Department of Medicine and [2]Department of Epidemiology, Columbia University Irving Medical Center, New York, New York; Editor-in-Chief, *Annals of the American Thoracic Society*; [3]Department of Thoracic Medicine, The Prince Charles Hospital, Brisbane, Queensland, Australia; Editor-in-Chief, *Journal of Cystic Fibrosis*; [4]Department of Surgery, University of Cincinnati, Cincinnati, Ohio; Editor-in-Chief, *Respiratory Care*; [5]University of Dundee, Dundee, Scotland; Deputy Chief Editor, *European Respiratory Journal*; [6]London, England;

Lederer, et al. 2019. Control of Confounding and Reporting of Results in Causal Inference: Studies Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. Annals of the American Thoracic Society, Volume 16, Number 1, 22-28.

- Key Principle #1: Causal inference requires careful consideration of confounding (干擾混淆)
  - Preferred variable selection methods
    - 1. Historical confounder definition with purposeful variable selection
    - 2. Causal models using directed acyclic graphs
  - Variable selection methods do not adequately control for confounding
    - 3. P value– or model-based methods
    - 4. Methods based on β-coefficient changes
    - 5. Selection of variables to identify "independent predictors"
  - Do not present all of the effect estimates from a model designed to test a single causal association
    - **Design of Experiments (DOE)**, **Taguchi Method**, Fused Lasso, Grouped K-fold Cross Validation, Tree-based Method, etc.

Lederer, et al. 2019. Control of Confounding and Reporting of Results in Causal Inference: Studies Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. Annals of the American Thoracic Society, Volume 16, Number 1, 22-28.

因果關係的必要條件

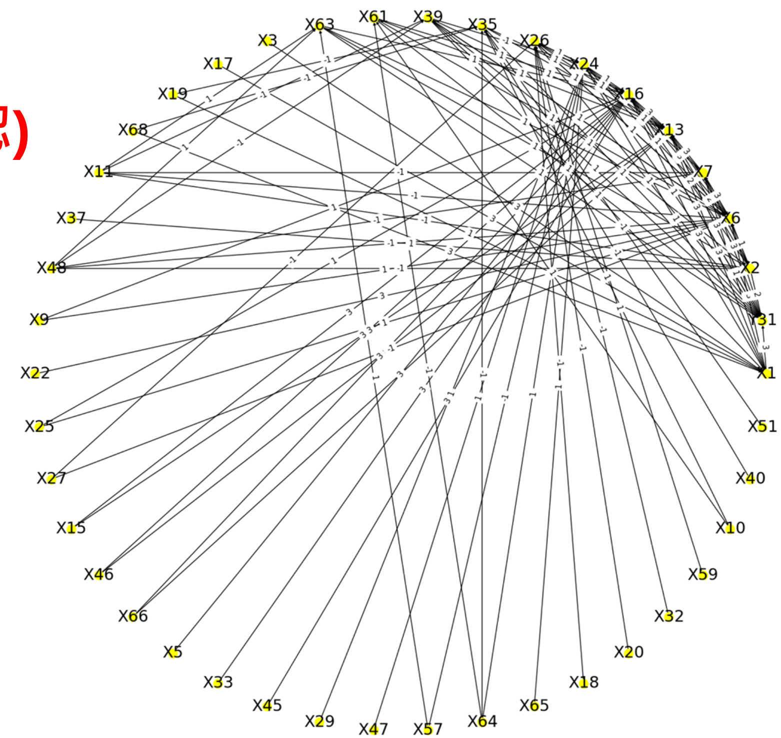
- 相關性 ← 大多時候ML/DS確認了這個...

- 前因後果(時間順序)

- 連貫(動)性

問題與因果

- 物理上重要因子沒有選進來？
  - **Granger causality tests (時間前後關係確認)**
    - directed acyclic graphs
  - 改變資料區間長短
    - 取長時間資料 → generalized因子選入
    - 取短時間資料 → specific因子選入
  - CRISP-DM 標準流程
    - Business understanding (商業理解)
    - **Data understanding (數據理解)**
    - Data preparation (數據預備)
    - Modeling (塑模)
    - Evaluation (評估)
    - Deployment (佈署)

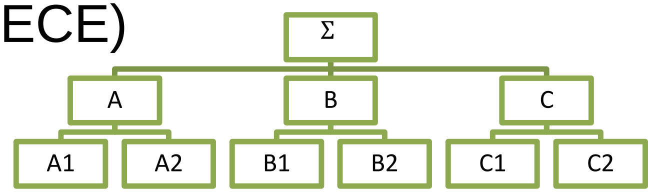


數據好理解了，則數據分析後的結果就好理解了。

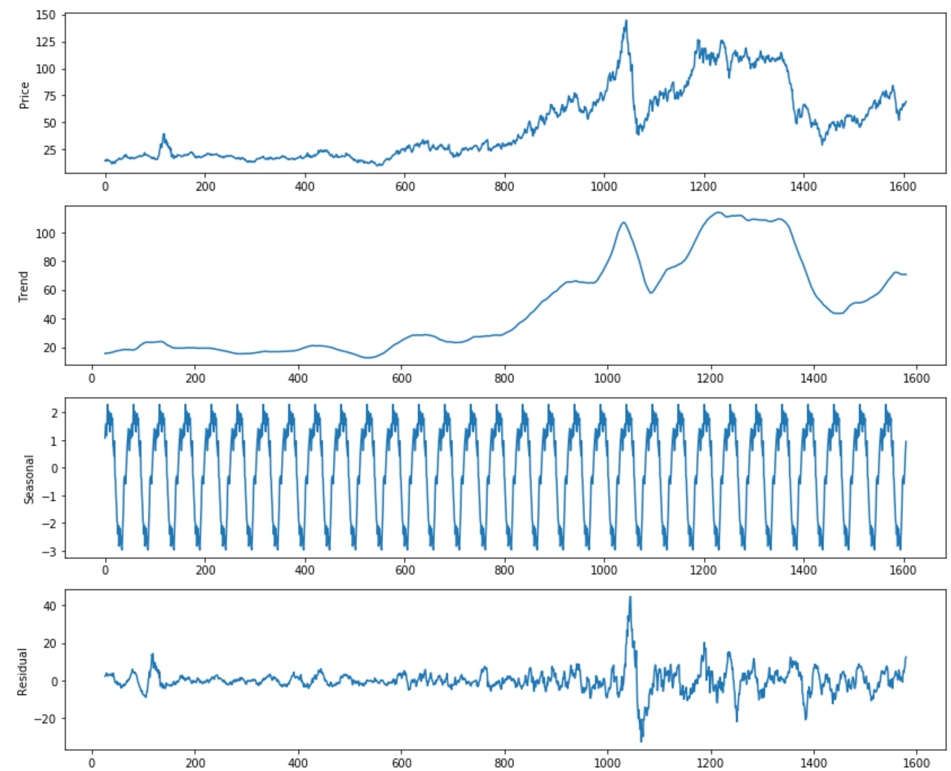
Wu, T.-Y., Lee, C.-Y., Chang, C.-H., Jiang, L.-D. 2020. Data science for Brent oil price prediction. 2020 Chinese Institute of Industrial Engineers (CIIE) Conference & Annual Meeting, Yunlin, Taiwan.

問題與因果

- 數據理解 (原則：分而治之Divide-and-Conquer strategy)
  - 橫向展開：將因子分類後各別類別篩選因子 (留意MECE)
    - 供需(量)時間序列：上下游產品、互補品、替代品等
    - 價格時間序列：上下游產品、互補品、替代品
    - 總體經濟：GDP、M1B、PMI、CPI(或年增率)、Interest Rate、Exchange Rate等
    - 地域(空間)相關：政治、經濟、文化、經緯度、空間距離、地形地貌、time lag等
    - 新聞輿情相關：社群網路、報章雜誌…

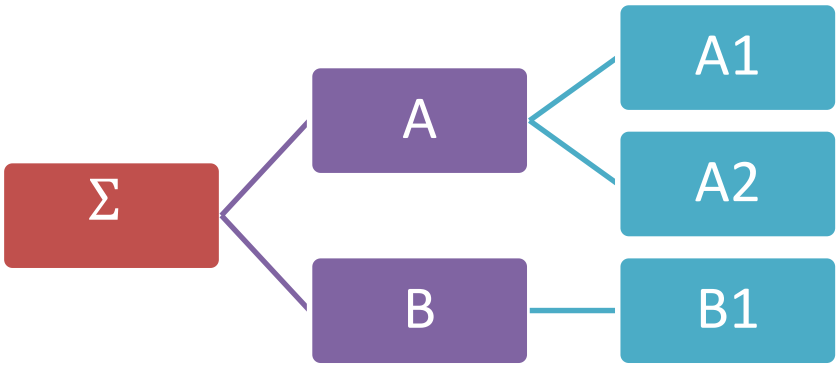


  - 縱向分解：時間序列分解、信號分解
    - Time Series Decomposition
      - Trend (趨勢)
      - Cyclicality (long cycle) (週期)
      - Seasonality (short cycle) (季節)
      - Randomness (隨機)
    - Hilbert-Huang Transform
      - Empirical mode decomposition
      - Intrinsic mode functions (IMF)



人 + 機

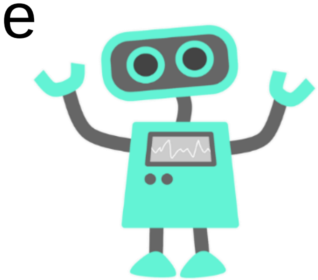
Rule-based



Model-based

$$Y = f(x)$$

Model-free



機 + 機

平台整合 (eg. Google map, 自駕車)

Beyond the Prediction…

PQLab

問題與因果

1. Problem Identification
2. Causality

批判與資訊價值

3. Criticism
4. Value of Information

風險與決策

5. Risk Assessment
6. Prescriptive Analytics

適應與擴充

7. Concept Drift
8. Domain Adaptation

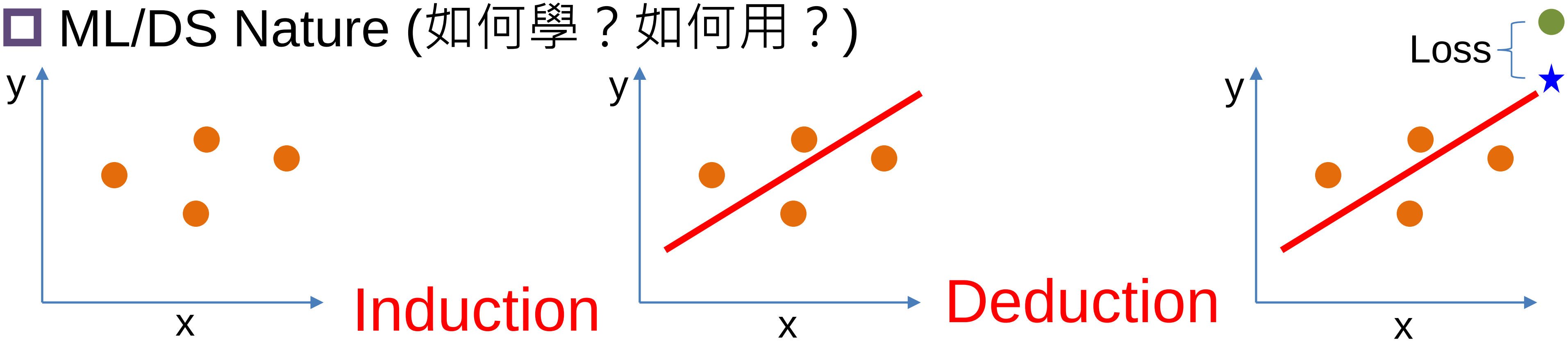
ML/DS的預測結果都是可信的嗎？

如果不可信...(那我們為啥AI?)

理由為何？WHY？

批判與資訊價值

- ML/DS Nature (如何學？如何用？)



Induction
(歸納法)

Deduction
(演繹法)

Specification
特殊化
(樣本)

0 → 1

Generalization
一般化
(母體)

1 → N

Specification
特殊化
(樣本)

Data Focus (學)
Model Focus

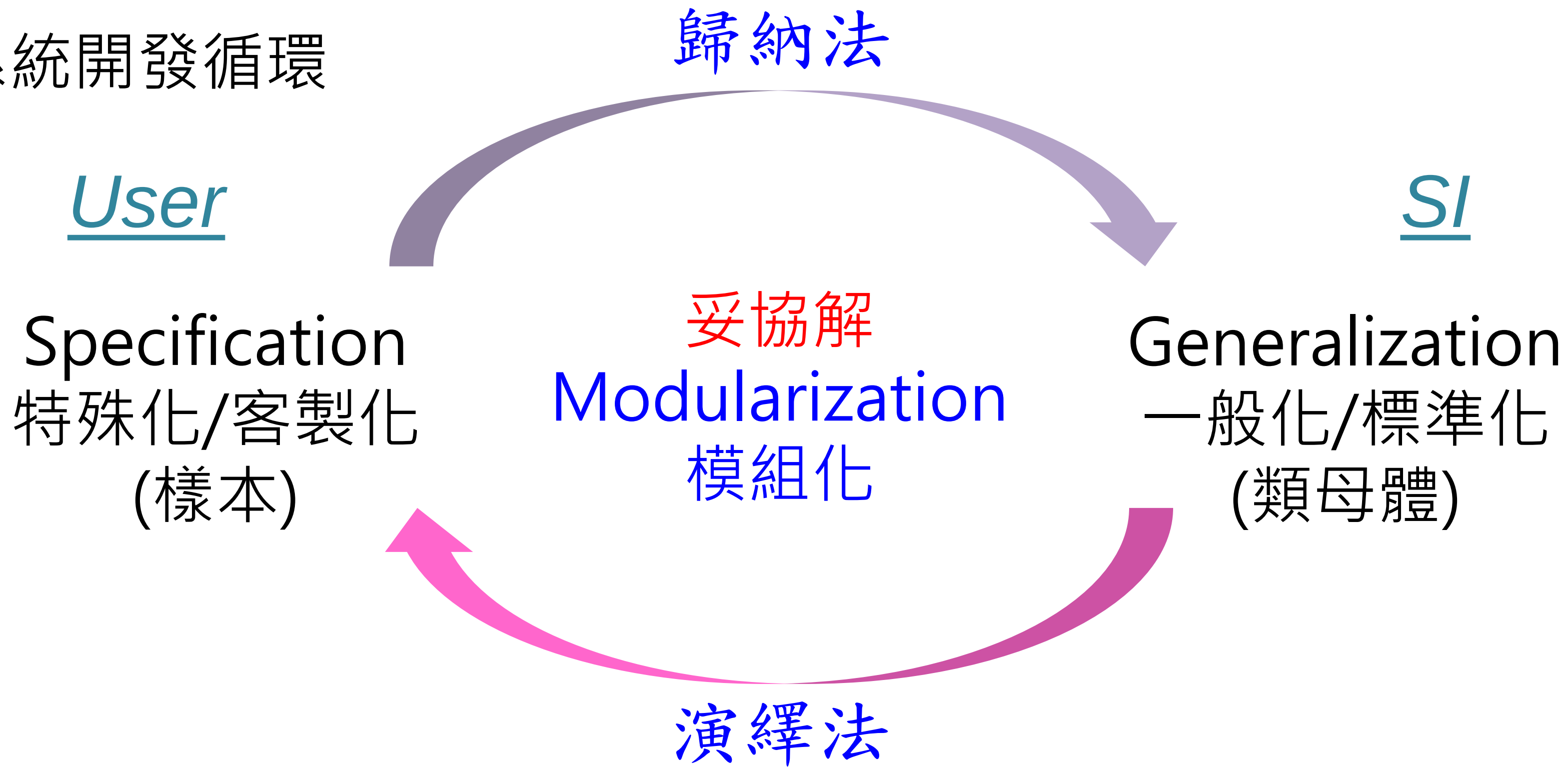
Application Focus (用)
Adaptation Focus

批判從訓練的 數據/個案/樣本 下手	批判從模型的 假設/前提 下手
- 找反例 - 批判樣本的 代表性/小樣本/收集偏誤	- 模型的假設不滿足 (i.e. 使用條件不對) - 另有他因 可能導致一樣的結果 - 因果關係 (相關性、時間順序、連貫性)

學以致用！學用落差？

批判與資訊價值

- 系統開發循環



- 在A公司做得好，不見得可以直接地導入在B公司
- 特殊化→一般化→特殊化→一般化→特殊化→一般化→...
- Mindset轉變
  - As-Is: 採購軟硬體、安裝設備、撰寫程式、教育訓練
  - To-Be: 結合協作體系(跨領域專業)及客戶/業務導向，共創共贏

PO Lab

ML/DS的預測模型提供了資訊...

其價值(value)如何？

可以衡量(measure)嗎？

若跟決策與資源連結，就有機會衡量報酬(reward)或悔惜(regret)

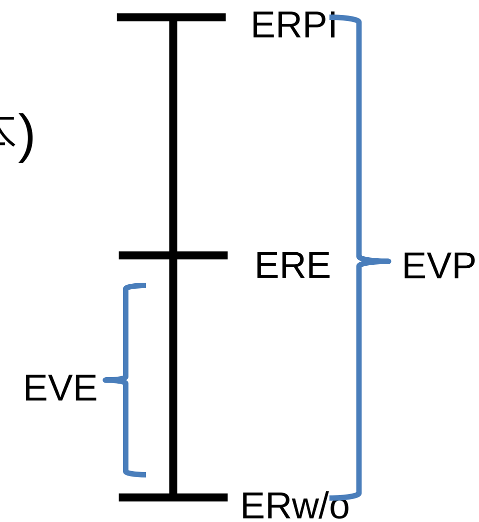
批判與資訊價值

完全資訊期望價值 (Expect Value of Perfect Information, EVPI)

- 強假設：模型提供資訊100%預測準確能消除不確定性(uncertainty)，i.e. 完全資訊 (perfect information).
- EVPI = 完全資訊期望報酬(ERPI) - 無資訊期望報酬(ERw/o)
- EVPI衡量決策者對於未來不確定性願意支付的金額用以換來完全資訊

實驗期望價值 (Expected Value of Experimentation, EVE)

- 因為很難獲得完全資訊，所以透過作實驗來產生資訊(樣本)
  - 不完全資訊(incomplete information)
- EVE = 實驗期望報酬(ERE) - 無資訊期望報酬(ERw/o)
- EVE計算若大於實驗成本則代表實驗值得進行
  - 實驗成本就是ML/DS建模與模擬所花的成本



(Raiffa and Schlaifer, 1961)

貝氏分析(Bayesian Analysis) for EVPI and EVE

- 假設：預測結果的混亂矩陣視作概似函數(likelihood function)

POLab

外科手術加護病房拔管預測Extubation Prediction in Surgical Intensive Care Unit (SICU)

- Data Source: ICIP of case hospital in southern Taiwan (2015-2016)
- 23 variables: biochemistry, arterial blood gas (ABG), blood cell, Glasgow coma scale (GCS), APACHE, extubation, etc.
  - Imbalanced dataset between 1565 and 626,894
- 359 observations including 49 failure cases (i.e., reintubation)
- Summary and feature selection

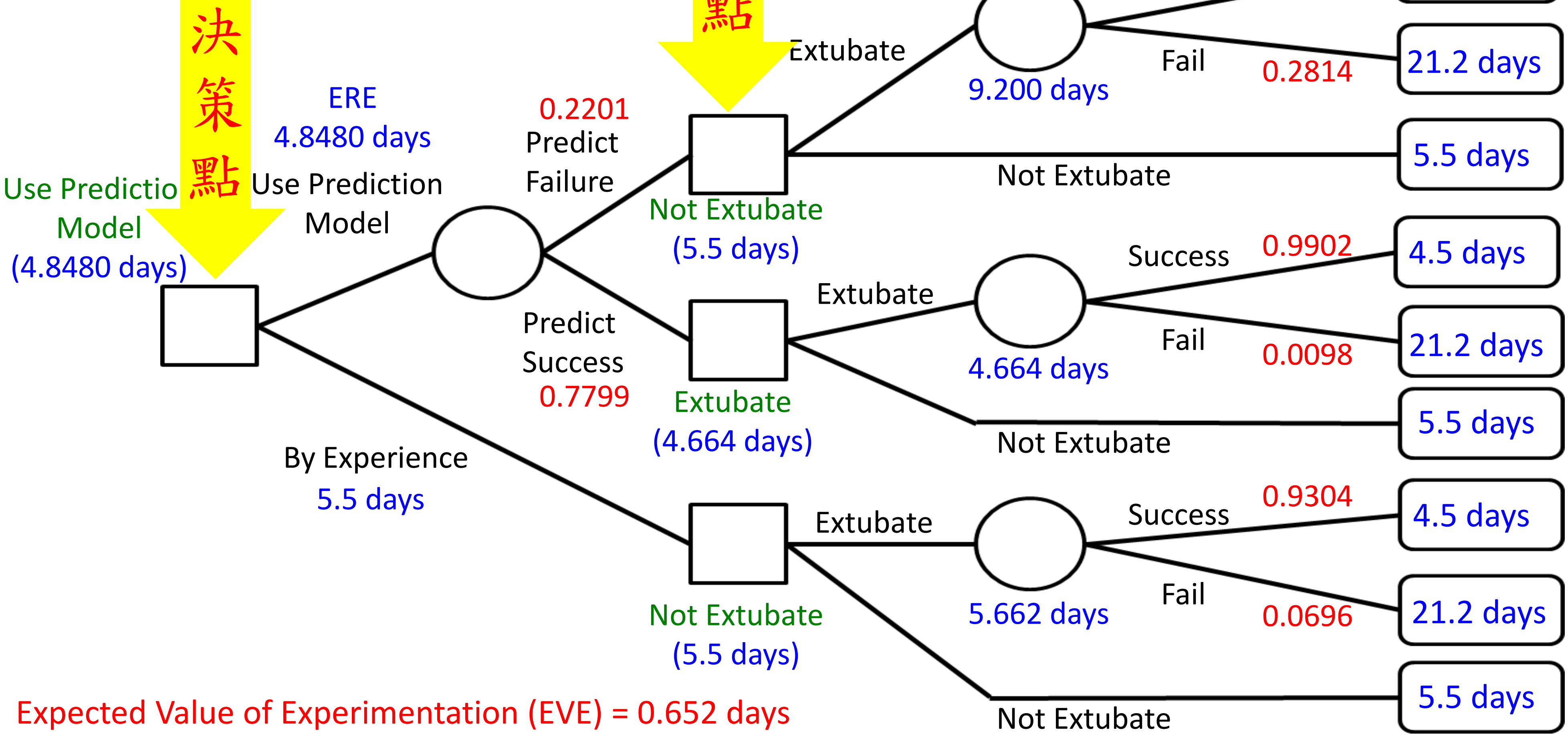
	Success (Mean/Std)	Failure (Mean/Std)
APACHEII	12.11/5.21	17.82/5.80
RSBI	48.04/27.91 breath/(min × L)	67.88/34.41 breath/(min × L)
Heart Rate	91.49/17.06 bpm	94.69/14.91 bpm
White Blood Cells	12.05/4.85 $10^3/\mu L$	12.87/4.37 $10^3/\mu L$
Na^+	138.47/4.29 mmol/L	140.98/6.69 mmol/L
Glu	175.47/62.66 mg/dL	182.14/55.54 mg/dL
PaO_2/FiO_2	367.78/96.62 mmHg	324.26/75.72 mmHg
Hct (ABG)	34.16/5.35%	31.64/4.04%
Age	58.83/15.40	64.58/16.89
Weight	64.15/14.77 kg	62.03/13.12 kg

Variables	Freq.
ApacheII	292
WBC	155
Eye_Opening	114
Heart_Rate	111
Glu	108
Na	103
Hct (ABG)	100
RSBI	90
Platelets	64
Weight	62
Verbal_Response	61
PT_INR	59
ARTmean_BP	54
$pO_2\_FiO_2$	53
PIMAX	44
Gender_men	36
ICU_Emergency	19

Tsai et al. (2019)

批判與資訊價值

- 外科手術加護病房拔管預測
  - 決策分析(decision analysis)的 策樹
  - B sian analysis for posteri estimation

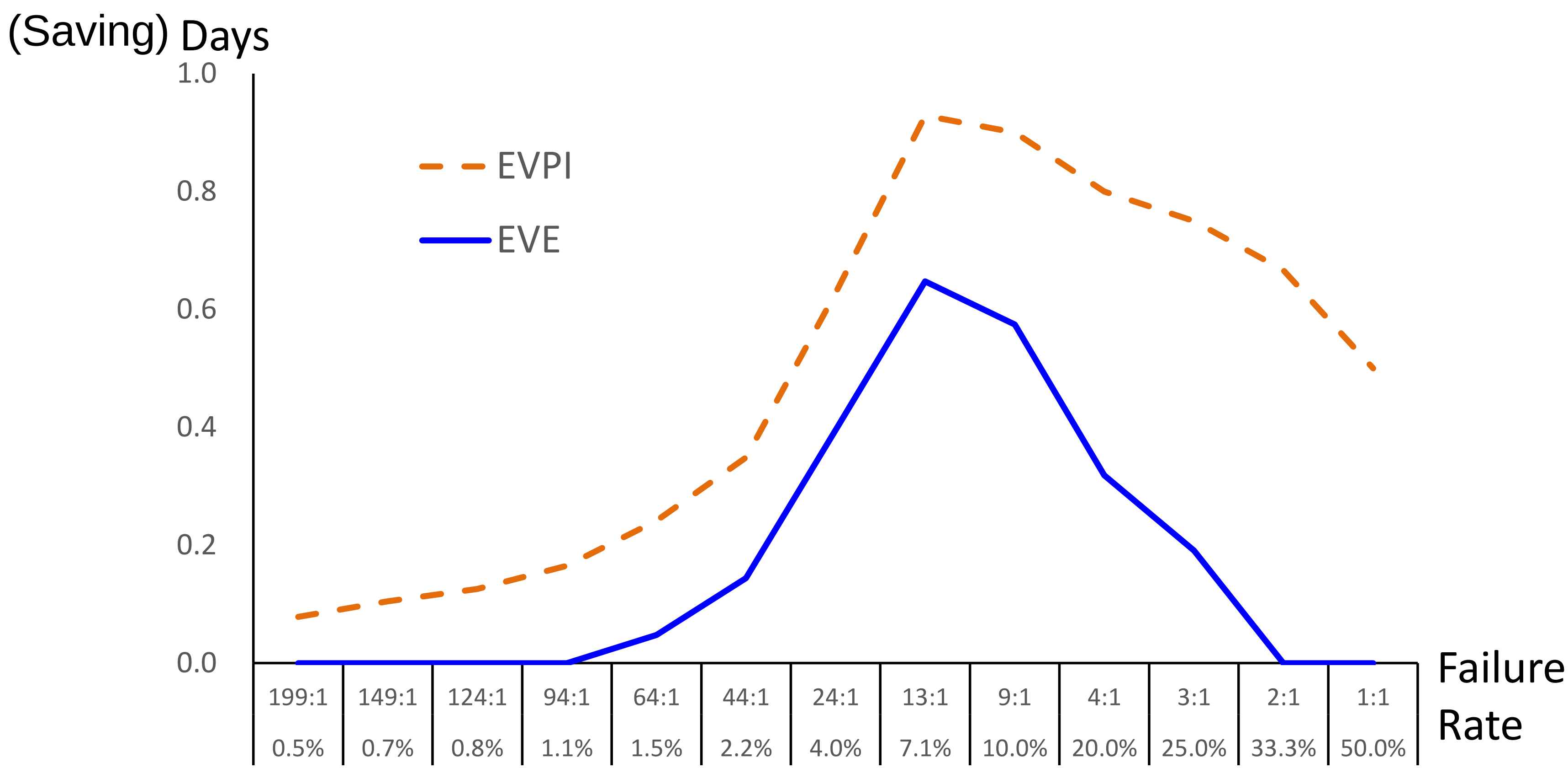


Expected Value of Experimentation (EVE) = 0.652 days
(Saving)

Tsai et al. (2019)

PoLab

- 外科手術加護病房拔管預測
  - Is always prediction model useful? "It Depends" on the prior…
    - → Applicable condition for building prediction model…



Tsai et al. (2019)

Beyond the Prediction…

問題與因果

1. Problem Identification
2. Causality

批判與資訊價值

3. Criticism
4. Value of Information

風險與決策

5. Risk Assessment
6. Prescriptive Analytics

適應與擴充

7. Concept Drift
8. Domain Adaptation

- For decision support…let us think about one simple case
  - If prediction model A with accuracy 95%, but it causes a "big loss" when the forecast is inaccurate.
  - If prediction model B with accuracy 90%, but it causes a "little loss" when the forecast is inaccurate.
  - If you are BOSS, which model do you prefer? A or B?

Risk Assessment (AI模型的風險評估)

- 風險辨識
  - 包括辨認會發生什麼？如何、為何、何處、何時發生？
  - 資源8M1I與組織關聯圖(跨部門的風險)
    - 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)
  - 模型預測風險、資料收集/品質風險、上線風險、使用者的風險…
- 風險分析
  - 包括找出事件發生的機率、風險之等級及其影響；須先確認既有偵測或控制機制，是既有機制下之機率等級及影響
  - 風險 = 影響(S) x 機率(P) x 偵測(D)
- 風險評量
  - 與風險基準比較，設定優先順序
  - 單因子敏感度分析(可控/不可控的因子所帶來的不確定性大小)
- 風險處理
  - 風險規避(risk avoidance)、風險降低(risk reduction)、風險轉移(risk transfer)、風險保有(risk retention)

行政院研究發展考核委員會(2009)，風險管理及危機處理作業手冊。https://cqa.nsysu.edu.tw/var/file/64/1064/img/2505/468727368.pdf

風險辨識檢核表(搭配組織關聯圖一起看)

AI品質預測/生產週期預測模型			
所屬：製造部 KPIs: move, WIP	(A)CIM/IT KPIs: 效率	(B)IE/PC KPIs: cost, 交期, cycle time	(C)工程/R&D KPIs: 良率
1人(Man)	(A1)配合執行人力	(B1)使用者的風險	
2機(Machine)			(C2)跨部門協調/依賴性
3料(Material)			
4方法(Method)		(B4)流程分析	
5測量(Measure)	(A5)數據取得		
6時間(Minutes)	(A6)模型Retrain		
7資金(Money)			
8環境(Mother nature/environment)			
9資訊(information)	(A9)一般化特殊化		(C9)online預測準度
…			

行政院研究發展考核委員會(2009)，風險管理及危機處理作業手冊。https://cqa.nsysu.edu.tw/var/file/64/1064/img/2505/468727368.pdf

風險分析 (風險 = 影響 x 機率 x 偵測)

- 影響(severity)之敘述分類表

等級	衝擊後果	形象/財產損失	KPI影響
3	非常嚴重	公司形象受損	KPI嚴重惡化
2	嚴重	跨部門形象受損	KPI中度惡化
1	輕微	部門/個人形象受損	KPI輕度惡化

- 可能性(possibility)之敘述分類表

等級	可能性分類	發生機率百分比
3	幾乎確定	51-100%
2	可能	21-50%
1	幾乎不可能	0-20%

- 風險本質圖像分析

影響(衝擊或後果)	風險分布		
非常嚴重(3)	3 (high risk) 高度危險的風險，管理階層需督導所屬研擬計畫並提供資源	6 (high risk) 高度危險的風險，管理階層需督導所屬研擬計畫並提供資源	9 (extreme risk) 極度危險的風險，需立即採取行動
嚴重(2)	2 (moderate risk) 中度危險的風險，必須明定管理階層的責任範圍	4 (high risk) 高度危險的風險，管理階層需督導所屬研擬計畫並提供資源	6 (high risk) 高度危險的風險，管理階層需督導所屬研擬計畫並提供資源
輕微(1)	1 (low risk) 低度危險的風險，以一般步驟處理	2 (moderate risk) 中度危險的風險，必須明定管理階層的責任範圍	3 (high risk) 高度危險的風險，管理階層需督導所屬研擬計畫並提供資源
	幾乎不可能(1)	可能(2)	幾乎確定(3)
		機率	

行政院研究發展考核委員會(2009)，風險管理及危機處理作業手冊。https://cqa.nsysu.edu.tw/var/file/64/1064/img/2505/468727368.pdf

風險與決策

等級	偵測分類	準確度/穩定性
3	無機制	低
2	有部分機制	中
1	有準確偵測機制	高

風險評量

- 據風險分析結果以判定需優先處理的風險
- 設計風險控制機制：可/不可控制因子

風險項目或可能發生情境	風險本質分析			風險等級	可能控制機制/作法	殘餘風險等級
	影響(S)	機率(P)	偵測(D)	(R)=(S)x(P)x(D)		(R)=(S)x(P)x(D)
B1	2	3	3	18	Action	2x2x2=8
C9	2	2	3	12	Action	1x2x2=4
…						

- 風險處理
  - 風險規避(risk avoidance)、風險降低(risk reduction)、風險轉移(risk transfer)、風險保有(risk retention)
  - 如何權衡你的風險？Eg. 誤放 vs. 誤宰

行政院研究發展考核委員會(2009)，風險管理及危機處理作業手冊。https://cqa.nsysu.edu.tw/var/file/64/1064/img/2505/468727368.pdf

PASS/FAIL Quality Prediction (128 lots in production line)

- Classification → tradeoff between confusion matrix → Cost-Sensitive
- Cost/penalty of two risks (Type I vs. Type II) → Prescriptive Analytics

Model A		Prediction FAIL	Prediction PASS
True	FAIL	61	7
True	PASS	29	31

Model B		Prediction FAIL	Prediction 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

Wu, T.-Y. (吳宗諭) 2019. Cost-Sensitive Method. https://github.com/wutsungyu/Cost-Sensitive
Lee, C.-Y., and Chien, C.-F., 2020. Pitfalls and Protocols of Data Science in Manufacturing Practice. Journal of Intelligent Manufacturing.

Beyond the Prediction…

POLab

問題與因果

1. Problem Identification
2. Causality

批判與資訊價值

3. Criticism
4. Value of Information

風險與決策

5. Risk Assessment
6. Prescriptive Analytics

適應與擴充

7. Concept Drift
8. Domain Adaptation

Retraining and Transfer

Concept Drift & Domain Adaptation

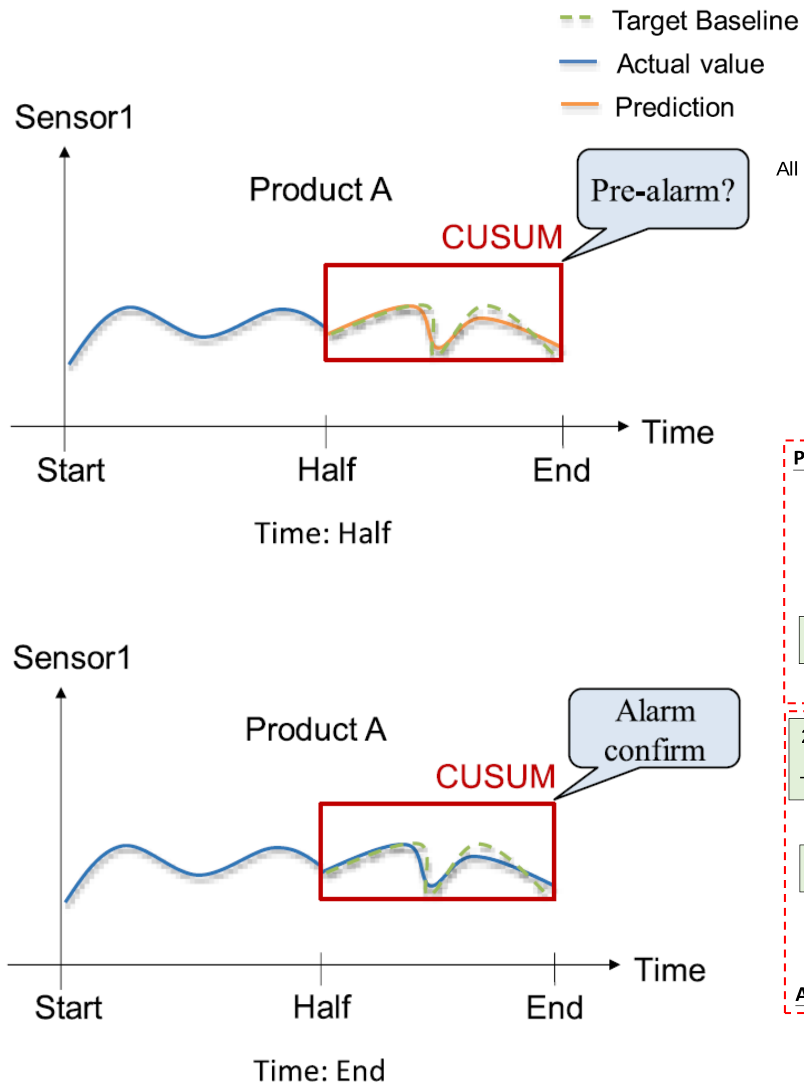
- Concept drift detection (何時模型不準了需要retraining)
  - Hypothesis test and control chart
- Concept drift understanding (釐清模型不準的根本原因)
  - Time of concept drift occurs (When)
  - The severity of concept drift (How)
  - The drift regions of concept drift (Where)
- Drift adaptation (如何retrain模型)
  - Training new models for global drift
  - Model ensemble for recurring drift
  - Adjusting existing models for regional drift

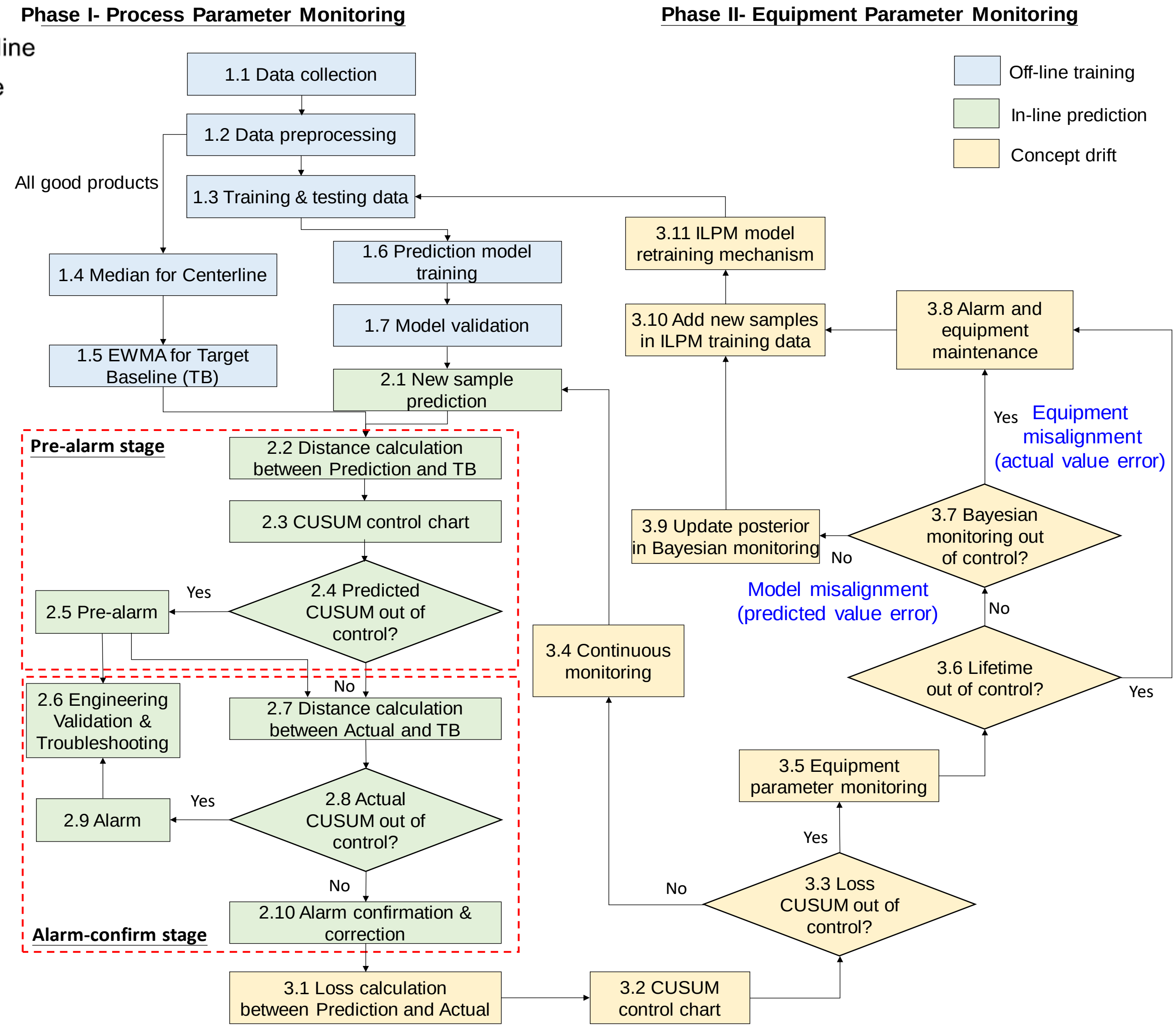
Transfer Learning

- Instance based transfer learning
- Feature based transfer learning
- Parameter based transfer learning
- Relational knowledge-based transfer learning

Lu et al. (2020). Learning under concept drift: a review. https://arxiv.org/abs/2004.05785v1
Ran, et al. (2019). A survey of predictive maintenance: systems, pruposes and approaches. IEEE Communications Surveys & Tutorials.

In-line Predictive Monitoring (ILPM)

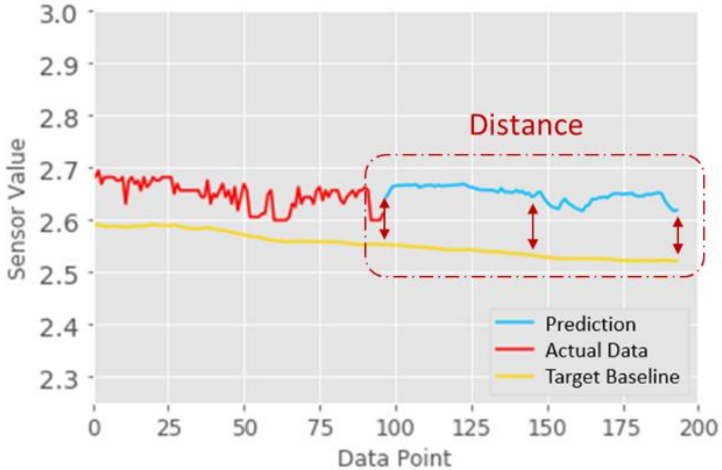


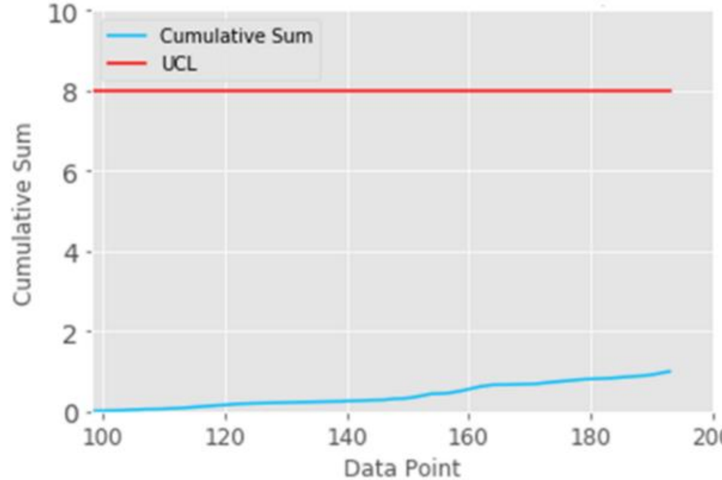


Lee, C.-Y., Wu, C.-S., and Hung, Y.-H., 2020. In-line Predictive Monitoring Framework. *IEEE Transactions on Automation Science and Engineering*.

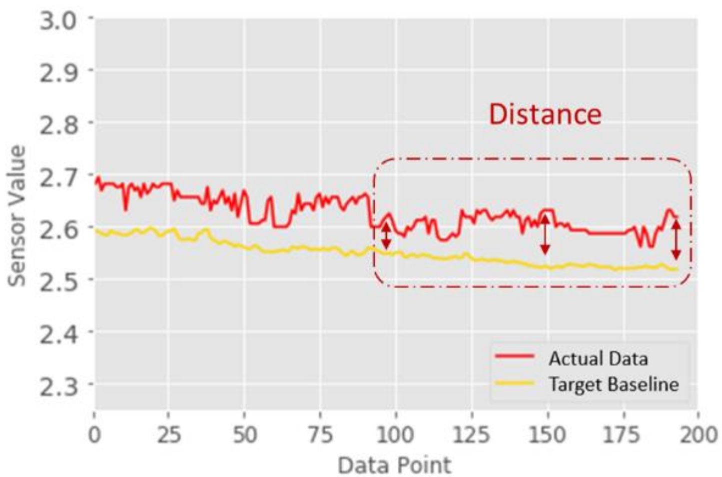
In-line Predictive Monitoring (ILPM)

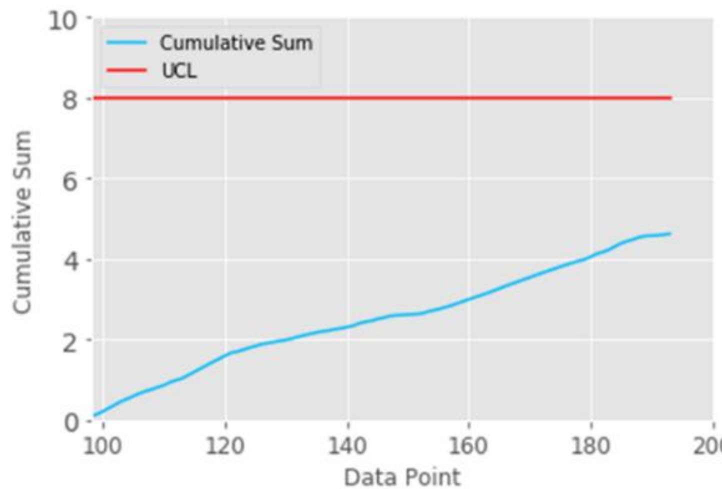
- Good product plots
  - Pre-alarm stage: (a) distance between predicted values and TB; (b) predicted CUSUM





  - Alarm-confirm stage: (a) distance between actual values and TB; (b) actual CUSUM



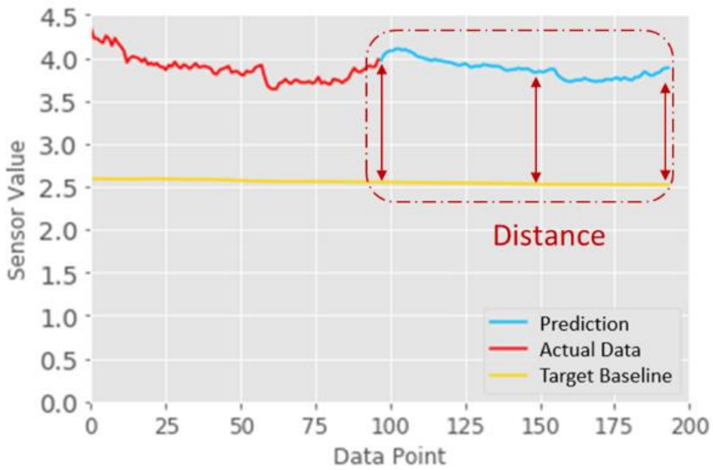


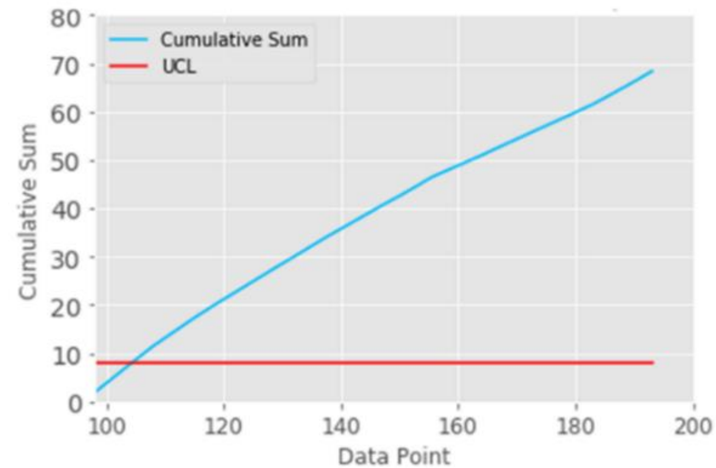
Lee, C.-Y., Wu, C.-S., and Hung, Y.-H., 2020. In-line Predictive Monitoring Framework. *IEEE Transactions on Automation Science and Engineering*.

In-line Predictive Monitoring (ILPM)

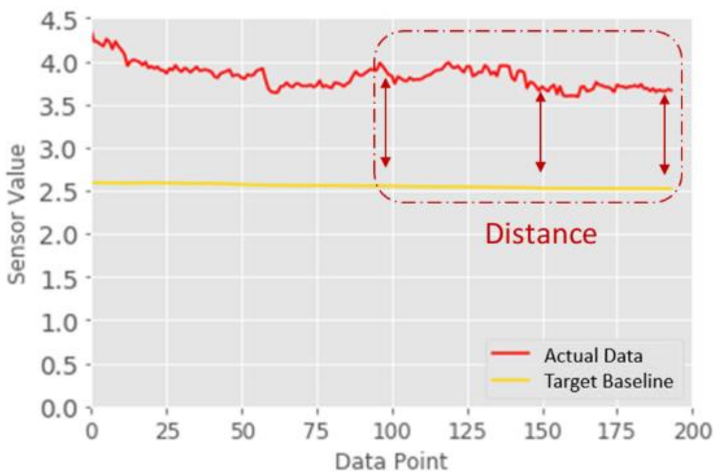
Defective product plots

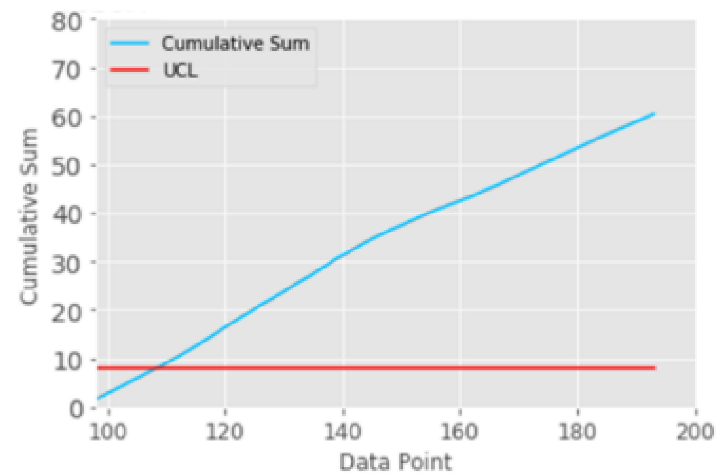
- Pre-alarm stage: (a) distance between predicted values and TB; (b) predicted CUSUM





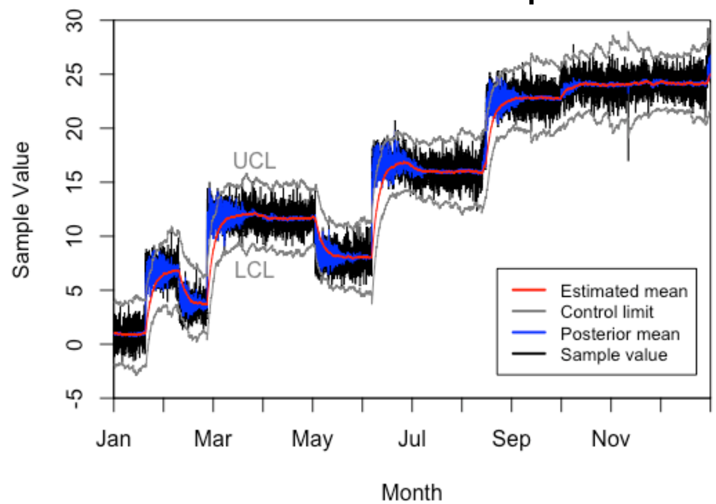
- Alarm-confirm stage: (a) distance between actual values and TB; (b) actual CUSUM

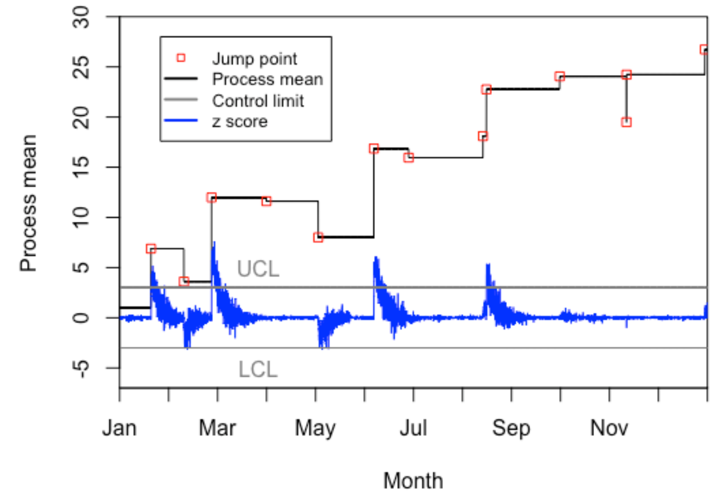




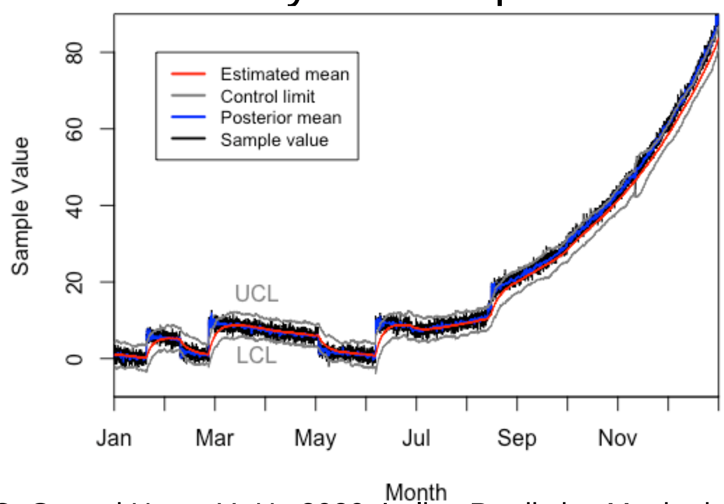
Lee, C.-Y., Wu, C.-S., and Hung, Y.-H., 2020. In-line Predictive Monitoring Framework. *IEEE Transactions on Automation Science and Engineering*.

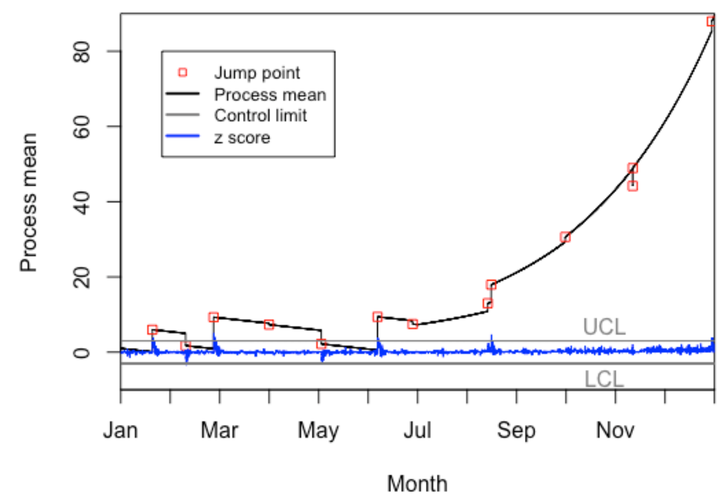
- Concept Drift of Model/Equipment Parameter
  - Bayesian Monitoring
  - Scenario 1: Random Jump





  - Scenario 2: Hybrid Jump with Trend





Lee, C.-Y., Wu, C.-S., and Hung, Y.-H., 2020. In-line Predictive Monitoring Framework. *IEEE Transactions on Automation Science and Engineering.*

結語

Conclusion

構析(Analytics)五階段

POLab

What	Why	When	How/Who	Where
What happened? (定義問題)	Why did it happen? (解釋根因)	When/What will happen next? (掌握趨勢)	How can we make it happen? Who takes it? (決策格局)	Where should build routines? (持續改善)
Descriptive 敘述性	Diagnostic 診斷性	Predictive 預測性	Prescriptive 處方性	Automating 自動性
數據收集 視覺化 界定利基	根本原因 物理解釋 因果關係	異質資訊 因子變異 系統趨勢	決策優化 資源槓桿 風險權衡	即時監控 適應性概念飄移 戰情財報
見	識	謀	斷	行

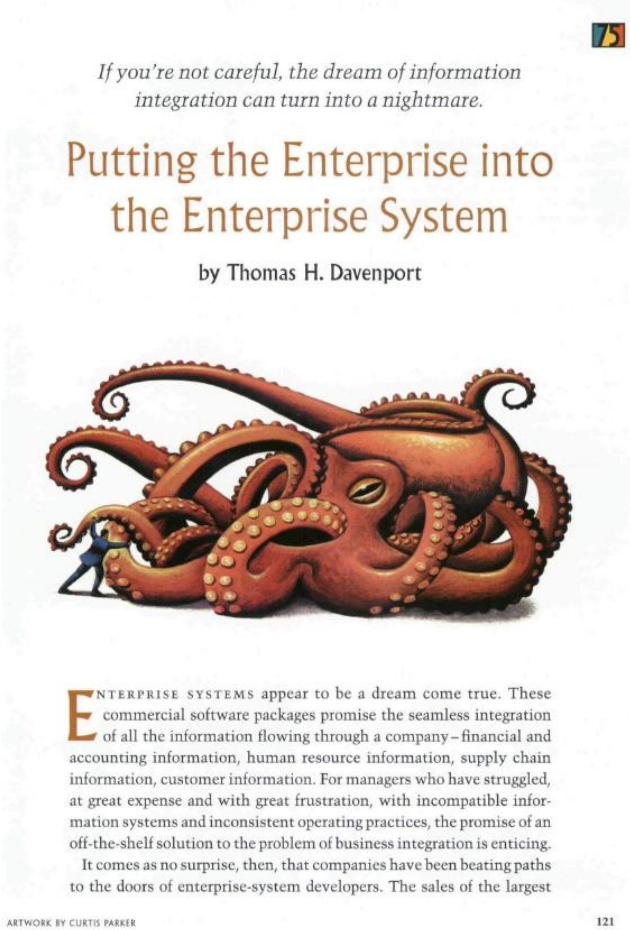
U N I S O N

Lee, C.-Y., and Chien, C.-F., 2020. Pitfalls and Protocols of Data Science in Manufacturing Practice. Journal of Intelligent Manufacturing.

AI導入是...誰的責任？

AI

“如果把安裝~~企業資源規畫系統~~的責任，全權交由資訊部門，是非常危險的。管理階層應該協調技術要務和業務要務，謹慎地控制企業系統的發展，否則很快就會反被系統控制。畢竟只要一不小心，資訊整合的美夢就會變成噩夢。”


75

If you're not careful, the dream of information integration can turn into a nightmare.

Putting the Enterprise into the Enterprise System

by Thomas H. Davenport

ENTERPRISE SYSTEMS appear to be a dream come true. These commercial software packages promise the seamless integration of all the information flowing through a company – financial and accounting information, human resource information, supply chain information, customer information. For managers who have struggled, at great expense and with great frustration, with incompatible information systems and inconsistent operating practices, the promise of an off-the-shelf solution to the problem of business integration is enticing.

It comes as no surprise, then, that companies have been beating paths to the doors of enterprise-system developers. The sales of the largest

ARTWORK BY CURTIS PARKER

121

HBR (1998)

- “但無論技術上的挑戰有多大，都不是企業系統失敗的主因。業務(business)問題才是最大的問題。失敗的原因在於，公司沒有調和企業系統的技術要務，以搭配企業本身的業務需求。”
- “企業系統的特質，就是將系統本身的邏輯，加在公司的策略、組織和文化上……量身打造(specific, customized)的流程，可能可以創造競爭優勢，企業系統仍會促使公司採用通用(generalized)的流程。

https://www.hbrtaiwan.com/article_content_AR0001429.html?utm_campaign=2010fanpage&utm_medium=GH_post&utm_source=Facebook&utm_content=1006_2200&utm_term=

但做AI大多是先花錢(投資)，而非先省錢..

這跟養小孩的道理類似...

單是教育訓練，就需要一筆花費...

除了花錢投資外...有時還要忍一下..

ML/DS的挑戰

- Unsupervised Learning
- Auto Labelling
  - 除正面表列外、可能也需要一些負面表列
- Metadata, Small Data, Big Data
- Data Generation/ Data Imbalance
- Transfer Learning & Domain Adaptation
  - 人 + 機 → 機 + 機
- Risk Assessment (AI預測不準時…)
- 稀少事件：每站都ok，那為啥還會有不良?
  - Control:單變量都in control但多變量out掉, scheduling/AMHS (每一站都delay一下…)
  - Environment: 載具破損, 溫濕度變化, 堆高機經過, 很多巧合造成的結果(環境異常)
- AI有時目的達到了，但過程有點歪掉…
  - Symbolic regression: 變數挑選ok; 但方程式不正確

誤把光頭當足球！球隊採用AI轉播　鏡頭瘋狂聚焦線審頭頂 (https://www.ettoday.net/news/20201101/1844532.htm)

蘇格蘭足球隊 Inverness Caledonian Thistle F.C.（恩華尼斯足球俱樂部，ICTFC）日前導入了名為 Pixellot 的 AI 攝影機…

智慧製造的挑戰

[Andrew Ng]

AI或許可以達到某些功能/目的，但仍需留意過程/物理特性
這部分目前還需要人的參與

感謝大家的支持跟參與
還請多多指教

Contact Information:
name: 李家岩 (Chia-Yen Lee)
phone: 02-33661206
email: chiayenlee@ntu.edu.tw
web: https://polab.im.ntu.edu.tw/

2020台灣人工智慧年會
Beyond the Prediction (預測之外)
http://polab.im.ntu.edu.tw/Talk/Beyond_the_Prediction.pdf