

東京大学量子ソフトウェア寄付講座

第2回量子ソフトウェアワークショップ:

これからの量子シミュレーション ～古典から量子へ、量子から古典へ～

# 量子機械学習への期待

—量子カーネル学習を中心として—

2023年1月30日

友野孝夫

[takao.tomono@toppan.co.jp](mailto:takao.tomono@toppan.co.jp)

凸版印刷株式会社デジタルイノベーション本部

# PROFILE（講演者）

**友野孝夫** (とも の たかお), Ph.D. IEEE Senior Member

凸版印刷株式会社  
デジタルイノベーション本部  
プリンシパルサイエンティスト

---

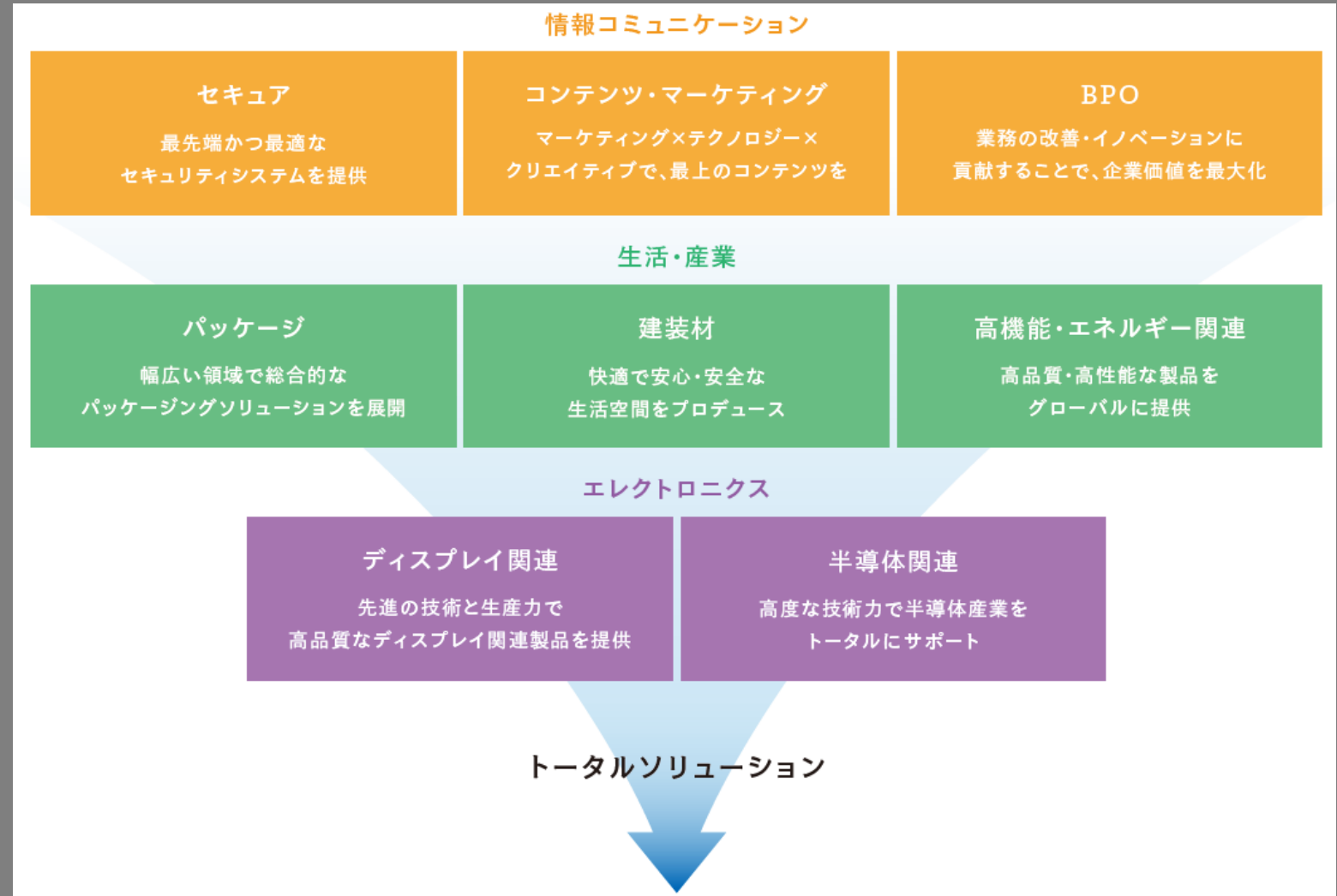
学生時代に応用物理を専攻、大学4年生の時に超電導量子干渉計の開発を経験

1984年からフォトニクス（量子光学、光計測、光学設計）、半導体・微細加工の研究開発に従事  
国内外の複数の会社で、研究開発、製品化を経験後、2003年、凸版印刷(株)入社。

（光による）微細加工、コンピュータービジョンの仕事を経て、  
現在、社内で量子を知るデータサイエンティストの育成と（光）量子アルゴリズムの研究に従事。

# PROFILE (会社)

凸版印刷は1900年の創業以来、  
ものづくりを中心に  
サービスに至るまで  
幅広い事業を展開中



# 動機・目的・目標・マイルストーン

動機・目的	量子コンピュータでの社会課題の解決
目標の設定	我々製造会社にとって、工場の出荷検査工程は非常に重要 量子機械学習の製造工程への適用（産業利用）
分類 (教師有学習)	2値：良品、不良品（主に検査工程） 3値：正常、故障予兆検知、異常検知（主に製造装置）
ニーズ	近年の特徴：少量多品種 ⇒少量のデータでの学習モデルの構築
今回の発表	特徴空間が広い量子カーネル学習を中心に研究開発を進めてきました。  今回、興味深い現象を見つけたので、紹介します。 さらには、今後の期待について述べます。

# AGENDA

- 1) 量子コンピュータの可能性
- 2) 機械学習の評価指標
- 3) 量子/古典カーネルの学習モデル構築過程
- 4) 量子機械学習への期待

# AGENDA

- 1) 量子コンピュータの可能性
- 2) 機械学習の評価指標
- 3) 量子/古典カーネルの学習モデル構築過程
- 4) 量子機械学習への期待

# 量子コンピュータに期待されている4つの応用分野

## 量子シミュレーション



2)量子世界の解明には量子システムで動く計算機が適している

- ・医薬品開発
- ・触媒探索
- ・素材探索 など

## 暗号



3)No-cloning定理より、コピーが不可なので、究極のセキュリティとして有望。

- ・安全保障の暗号通信
- ・金融分野の暗号通信
- ・暗号解読 など

## 数理最適化



4)計算スピードupで、NP困難問題を解決。グラフ理論のノードとエッジをqubitに適用可能。

- ・ルート最適化
- ・ポートフォリオ最適化
- ・スケジューリング など

## 注目

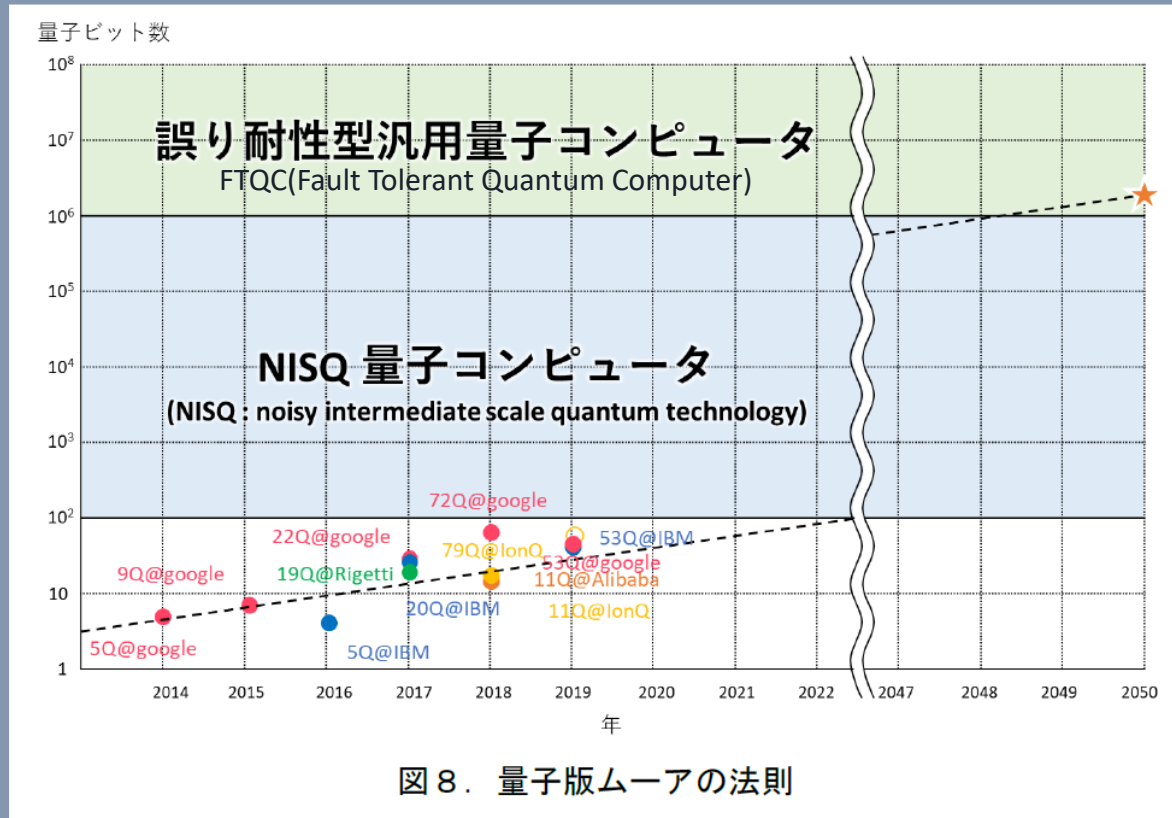
## 機械学習



4)計算スピードupな上、線形代数・行列を用いた計算が量子計算に親和性がある。

- ・工場の出荷検査
- ・不正検知と故障予知
- ・マーケティング分析 など

# 量子版ムーアの法則—NISQからFTQCへ—



量子コンピュータは2050年頃にFTQCが出現すると予測。

現在：  
クラウドで利用可能になった量子コンピュータと古典コンピュータ（従来型）とを併用して活用する時代（NISQ量子コンピュータ時代）。

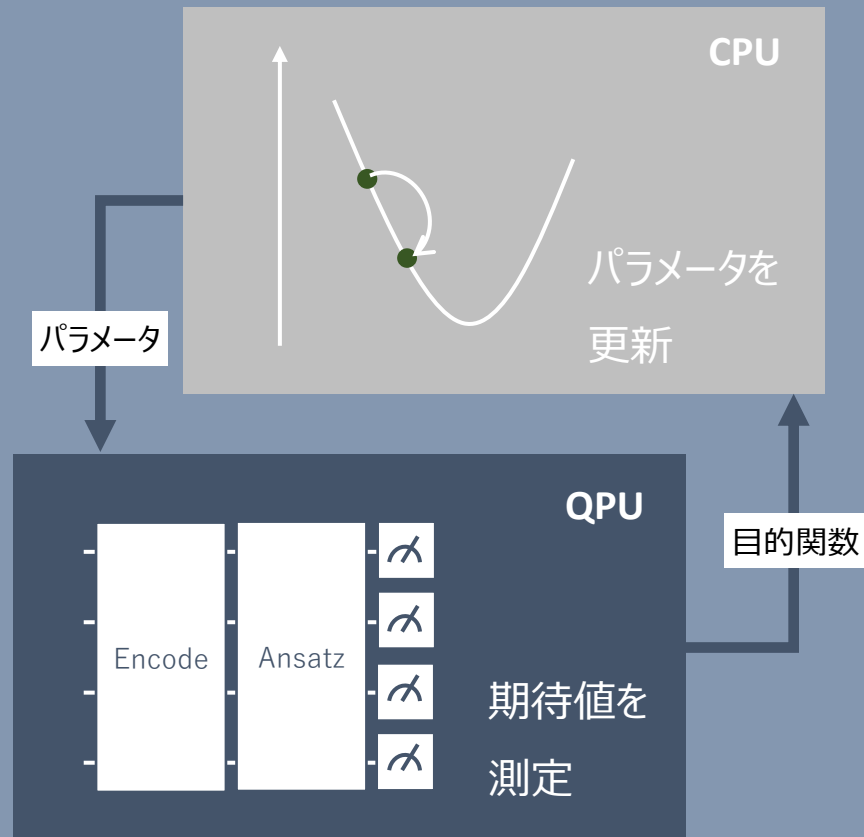
NISQ量子コンピュータで可能なことを考えることが特に民間企業では重要



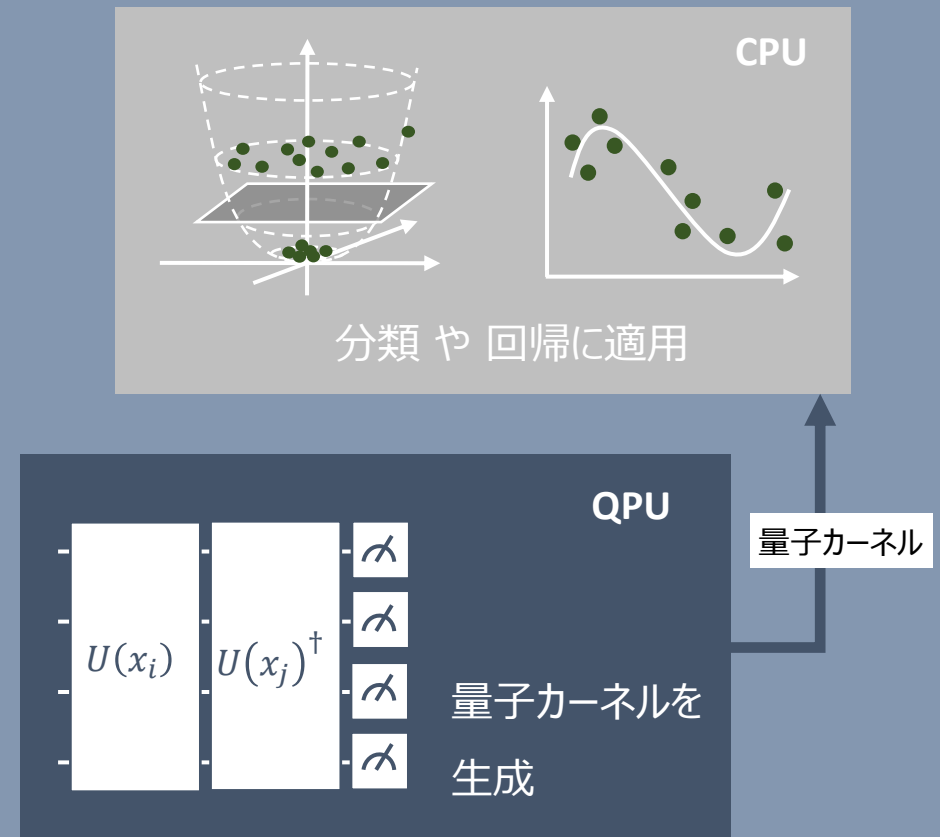
# NISQ時代の機械学習のアルゴリズム

量子古典ハイブリッドで活用する場合、量子機械学習の手法は大きく以下2つに分けられる。

## ■ 量子変分回路



## ■ 量子カーネル法



# AGENDA

- 1) 量子コンピュータの可能性
- 2) 機械学習の評価指標
- 3) 量子/古典カーネルの学習モデル構築過程
- 4) 量子機械学習への期待

# 評価指標の意味

- 分類などにおける評価は 1 つの評価指標だけではない  
予測に対する正解や誤りの中にも区別があり、**目的に応じて様々な評価指標を用いる**

- 例)
- ・正解率 (Accuracy)
  - ・適合率 (Precision)
  - ・偽陽性率 (FPR)
  - ・F値 (F1値)
  - ・再現率 (Recall)
  - ・偽陰性率 (FNR)

		予測	
		正	負
実際の結果	正	TP (真陽性)	FP (偽陽性)
	負	FN (偽陰性)	TN (真陰性)

# 評価指標: Accuracy

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

[概要] 全ての予測のうち、正解した割合

[用途] クラスごとに重要度の差が無い時に使用  
(犬と猫の分類、人と動物など)

		予測	
		正	負
実際の結果	正	TP (真陽性)	FP (偽陽性)
	負	FN (偽陰性)	TN (真陰性)

# 評価指標:F1-score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

[概要] PrecisionとRecallの調和平均

[用途] FPとFNをバランスよく注目して評価する指標

$$Precision = \frac{TP}{TP + FP}$$

適合率  
この値が高いほど、間違った分類が少ない

予測

		正	負
実際の結果	正	TP (真陽性)	FP (偽陽性)
	負	FN (偽陰性)	TN (真陰性)

$$Recall = \frac{TP}{TP + FN}$$

再現率  
この値が高いほど性能がよく、間違ったPositiveの判断が少ない

予測

		正	負
実際の結果	正	TP (真陽性)	FP (偽陽性)
	負	FN (偽陰性)	TN (真陰性)

# 評価指標の意味

例: ウイルスの感染者予測と実際の感染結果



実際の結果

		検査による予測	
		陽性	陰性
実際の結果	陽性		「陰性」と予測したのに、 実際は「陽性」だった
	陰性	「陽性」と予測したのに、 実際は「陰性」だった	

ここが少なければ  
陽性者のすり抜け  
が減る

ここが少なければ  
無駄に病床を利用  
せずに済む

# FPRとTPRについて

$$FPR = \frac{FP}{FP + TN}$$

偽陽性率  
Negativeとの予測で  
Positiveと判定された割合

		予測	
		正	負
実際の結果	正	TP (真陽性)	FP (偽陽性)
	負	FN (偽陰性)	TN (真陰性)

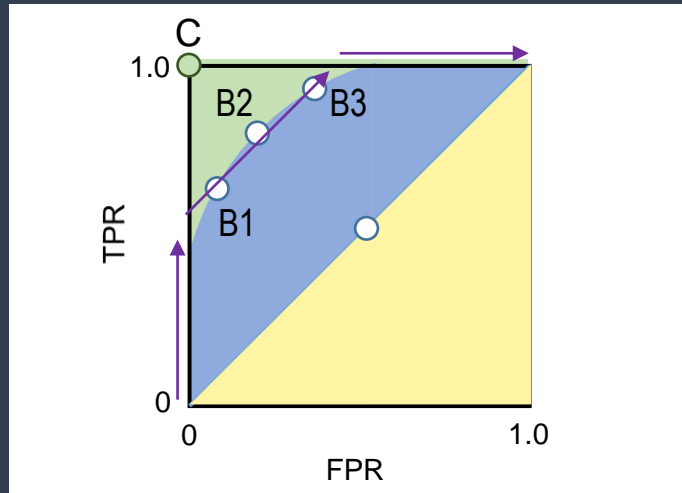
$$Recall, TPR = \frac{TP}{TP + FN}$$

再現率  
Positive と推測すべきデータのう  
ちどのくらい回収できたかの指標

		予測	
		正	負
実際の結果	正	TP (真陽性)	FP (偽陽性)
	負	FN (偽陰性)	TN (真陰性)

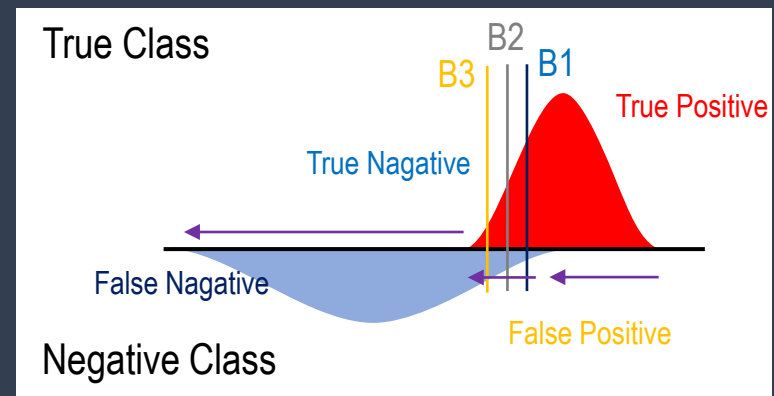
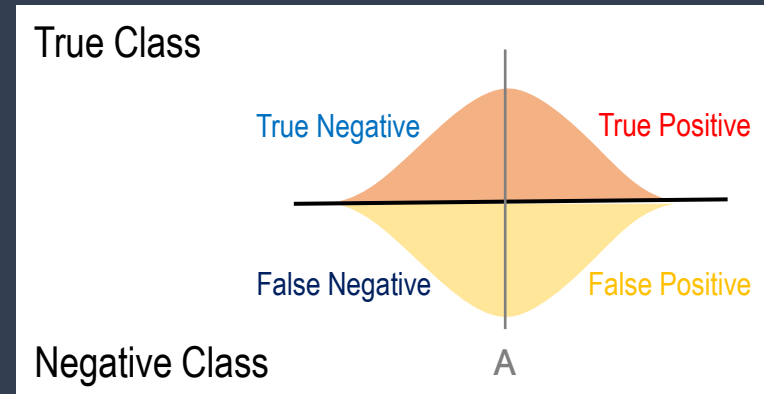
# ROC曲線について

当初：雑音の中から敵機を検出する受信者の特性を測るための方法として開発。  
現在、様々なところで活用されている。



ROC曲線の下面積：AUC (Area Under the Curve)  
これらは分類モデルの評価指標として活用。

ランダムで全く無効なモデル：	AUC=0.5	点A参照
ある程度構築された学習モデル：	AUC=0.8	点B参照
AUC が 1 のときが最良のモデル：	AUC=1	点C





# AGENDA

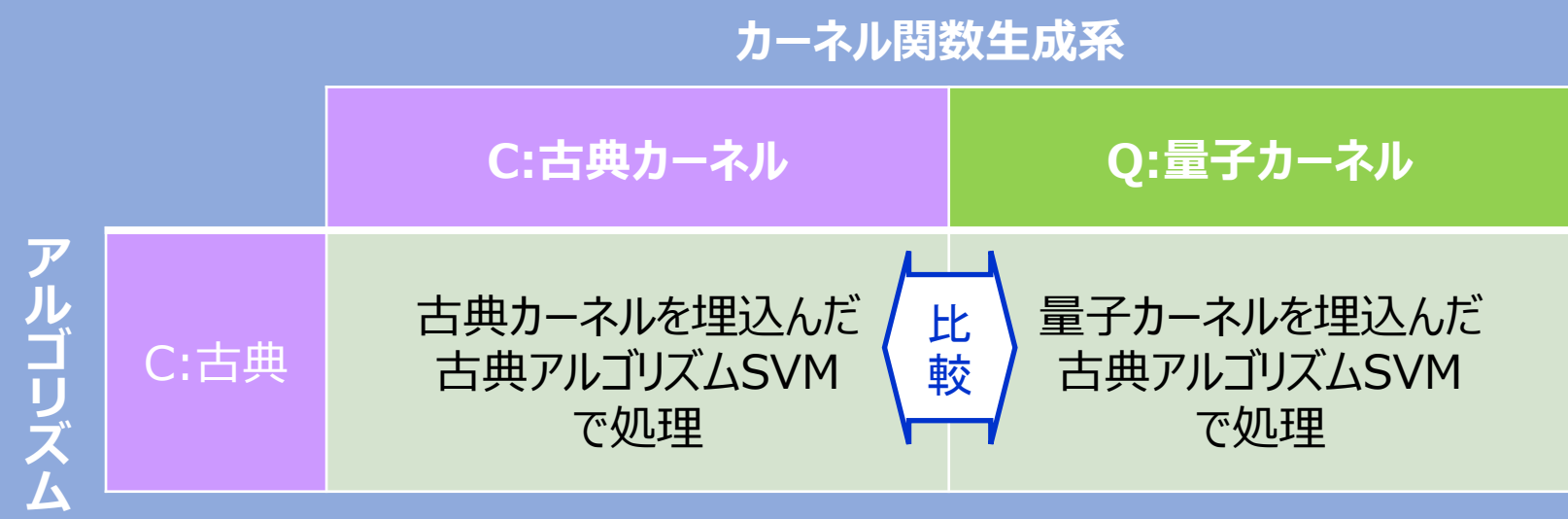
- 1) 量子コンピュータの可能性
- 2) 機械学習の評価指標
- 3) 量子/古典カーネルの学習モデル構築過程
- 4) 量子機械学習への期待

# 分類のアルゴリズムとその特徴

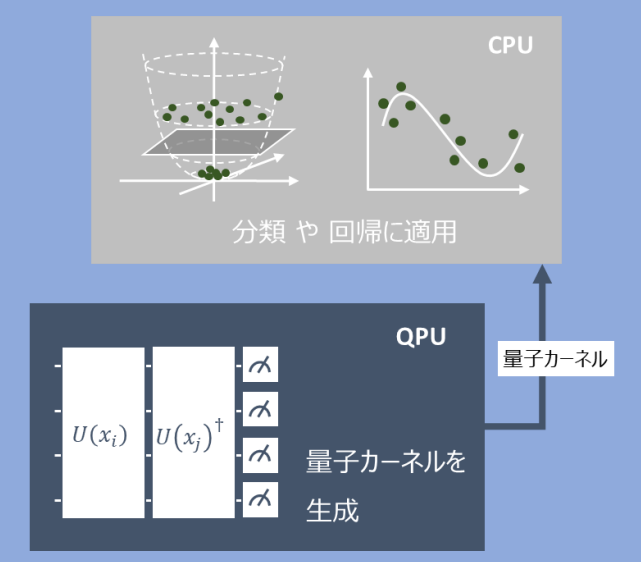
アルゴリズム	説明	特徴
決定木	樹木形のグラフで分類する手法	決定するとき明確
ロジスティック回帰	ある事象が起きる/起こらないの分類手法	最も単純なモデル
サポートベクターマシン(SVM)	データを分類する線で分ける手法(主として二値)	識別精度が高い
線形判別分析	$Z$ に線形写像し、 $z$ で2つのクラスに分離	ロバスト性がある
ナイーブベイズ	ある事象確率から原因確率を求める手法	高速だが精度が低い
深層学習	マルチタスク学習による汎化性能向上	多量の訓練データ必要
....	....	....

画像による良品不良品の検査工程では、SVMおよび深層学習が多く用いられる。  
 目的(少量データで学習モデル構築可、識別精度が高い)を考慮して、アルゴリズムとしてSVMを選択。

# カーネル法を用いた機械学習モデルの比較

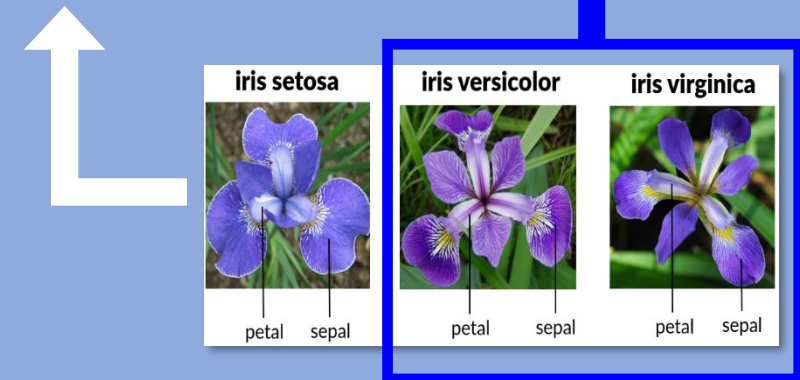


## ■ 量子カーネル法



# 利用するデータセット

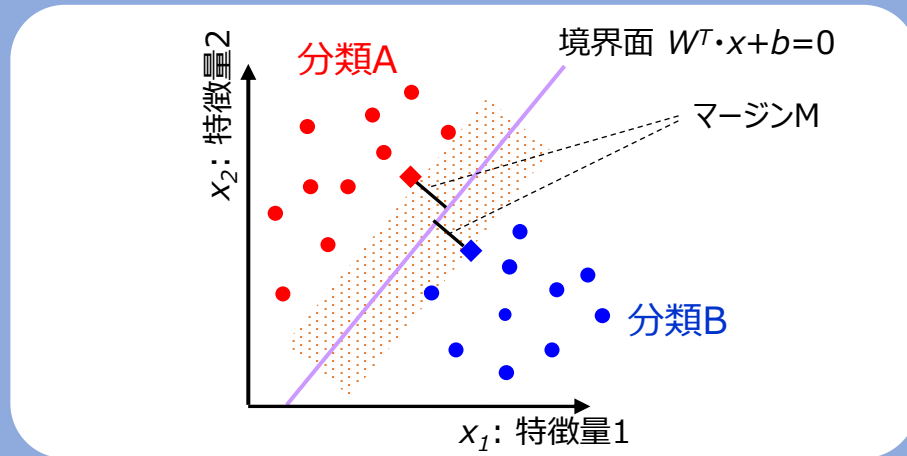
Datasets	Classes (Attributes)	Samples/class	Remark
Heart disease	2 (13)	138, 165	Positive, Negative
Wine	3 (13)	59, 71, 48	Three different cultivars
Iris_2	2 (4)	50, 50	Versicolor, Versinica
Iris_3	3 (4)	50, 50, 50	Versicolor, Versinica, Setosa



データセットは下記のサイトより入手 <https://archive.ics.uci.edu/ml/index.php>  
<https://www.kaggle.com/>

Irisの2値分類のデータセットは既存のデータセットより作成

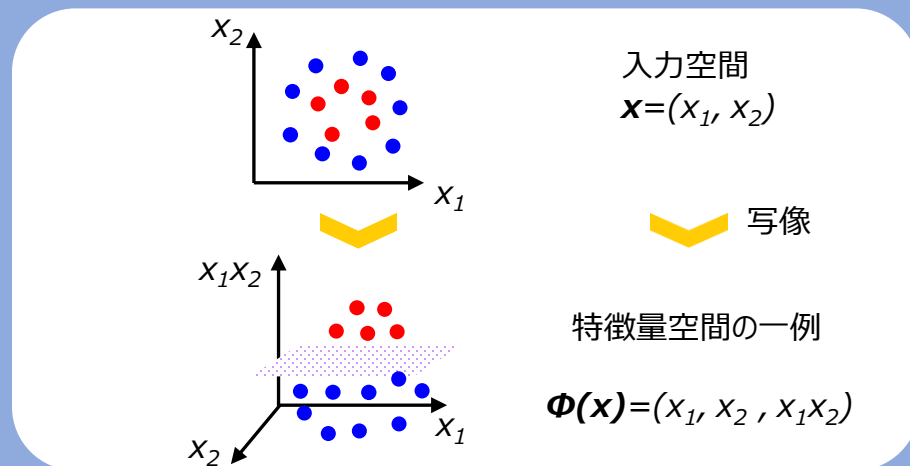
# 古典カーネル法を用いたSVM



## SVM

分離直線（面）からの距離が最大になるように設定する手法。

◆◆がサポートベクターであり、このマージンMが最大になる直線（面）を探す手法



## カーネル法（カーネルトリック）

次元を拡張して、特徴量を埋め込むことで非線形分離が可能になる手法。

多項式カーネル:  $\kappa(x_i, x_k) = (x_i^T \cdot x_k + c)^d$

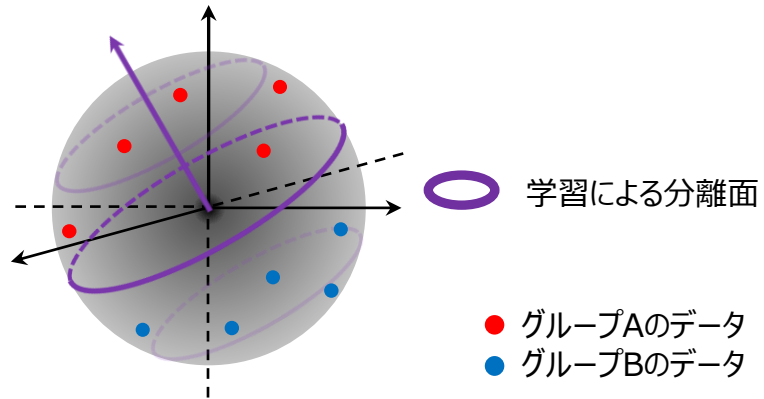
ガウスカーネル:  $\kappa(x_i, x_k) = e^{-\frac{|x_i - x_k|^2}{2\sigma^2}}$

tanhカーネル:  $\kappa(x_i, x_k) = \tanh(x_i^T \cdot x_k)$

など。

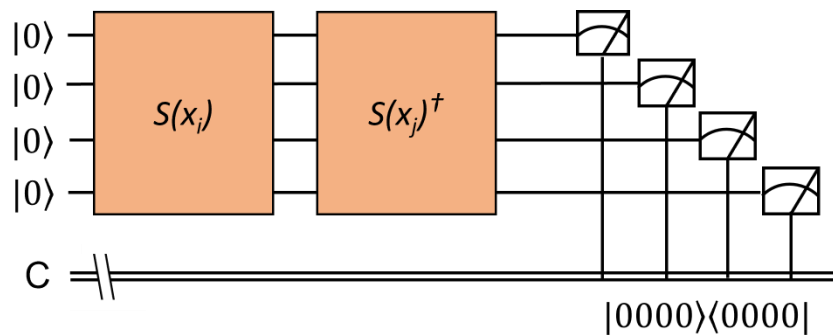
古典カーネルの課題：カーネルトリックで特徴を埋め込む関数 $\Phi$ が限られる。

# 量子カーネル法を用いたSVM



古典データにより量子データに変換され、ブロッホ球（特徴量空間）上にマッピングされる。

量子回路でマッピングされたデータで学習モデル（分離面）が構築され、分類Aと分類Bに分け



量子カーネル関数  $\kappa(x, x') = |\langle \phi(x') | \phi(x) \rangle|^2$

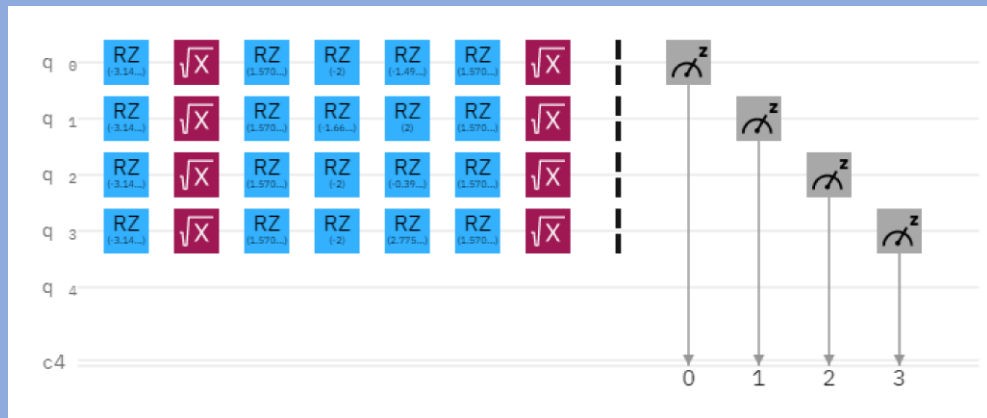
データエンコーディング回路  $|\phi(x)\rangle = |S(x)\rangle|0\rangle$

量子回路図  $\langle 0 | S^\dagger(x) S(x') | 0 \rangle \langle 0 | S^\dagger(x') S(x) | 0 \rangle$   
 $= \langle \phi(x) | \phi(x') \rangle \langle \phi(x') | \phi(x) \rangle$

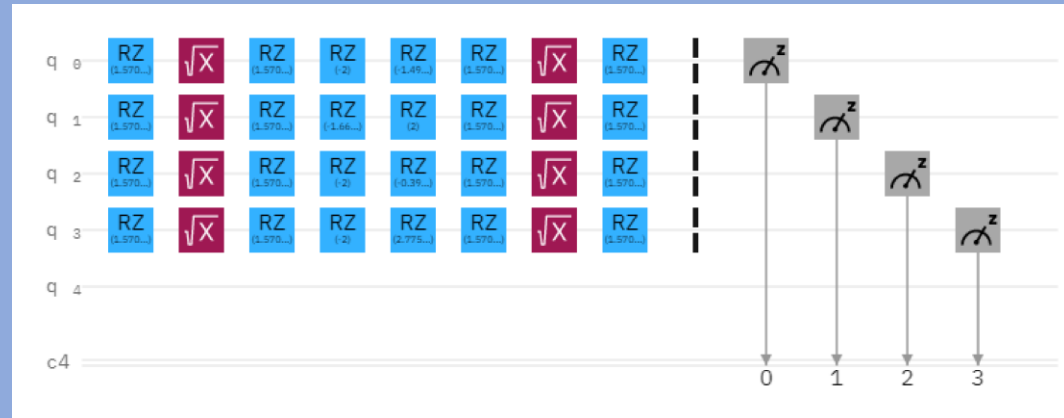
グラム行列

数式に示す通り、全てのデータの組合せについて計算する。  
 シミュレータ：Qiskit, blueqat

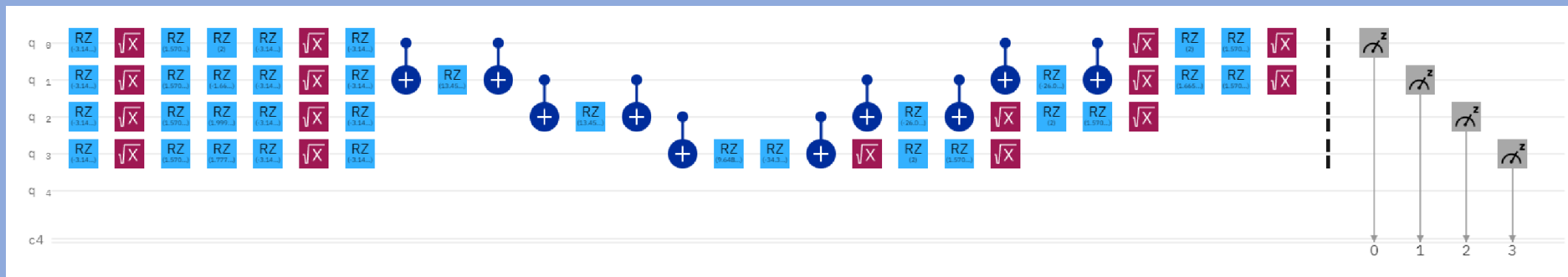
# 量子カーネルの量子回路 (Qiskitの例)



Y-特徴量マップを用いた量子カーネルの量子回路図

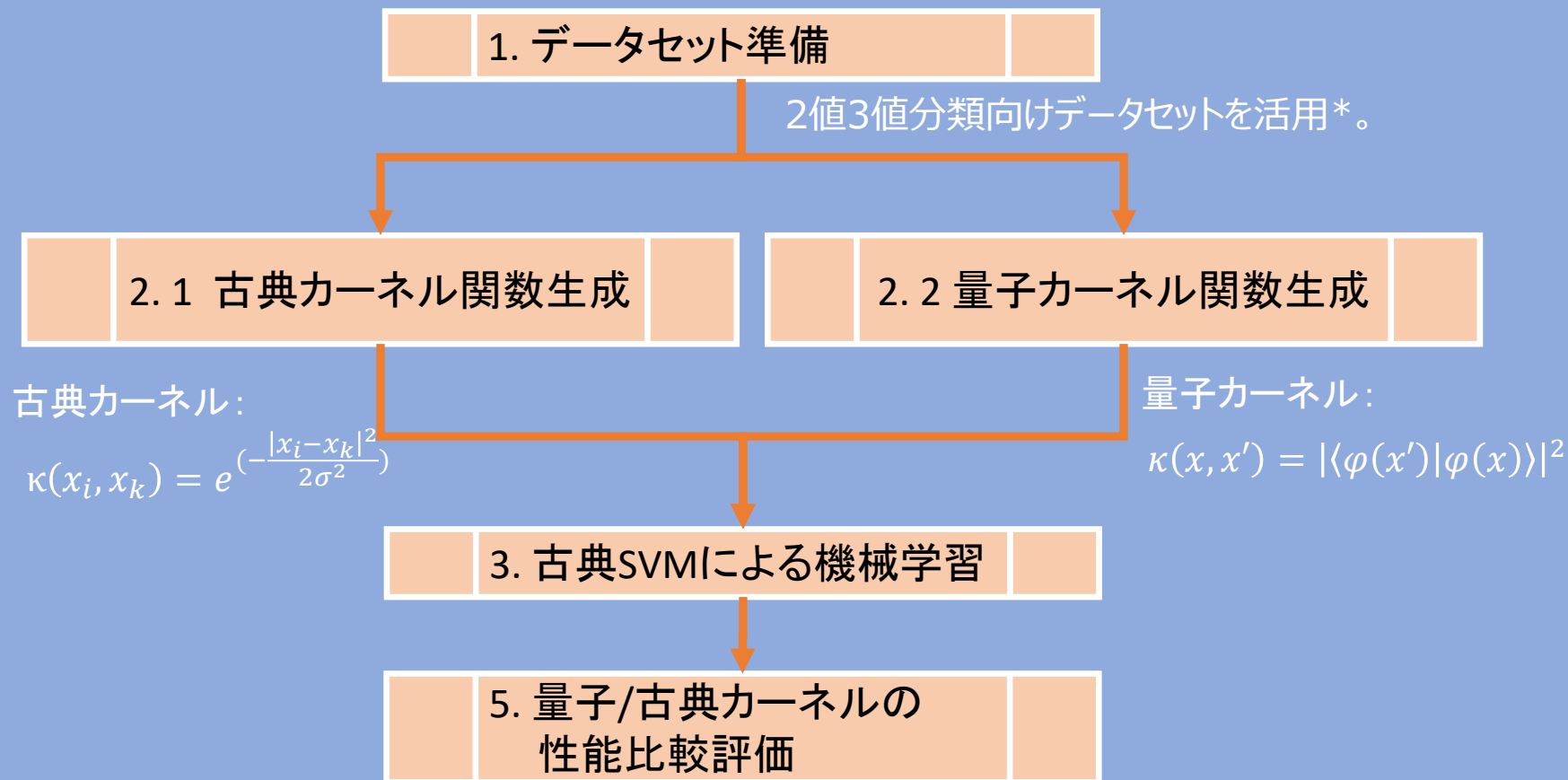


Z-特徴量マップを用いた量子カーネルの量子回路図

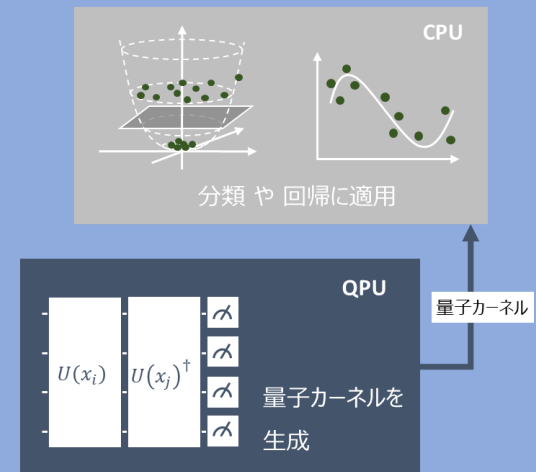


Y-ZZ特徴量マップ (量子もつれを持つ特徴マップ) を用いた量子カーネルの量子回路図

# データセットの性能評価の流れ



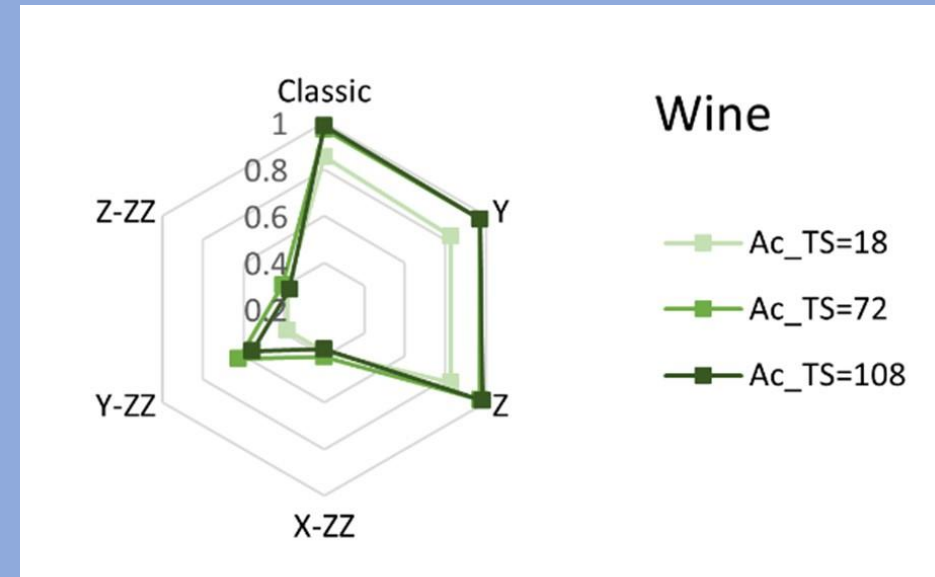
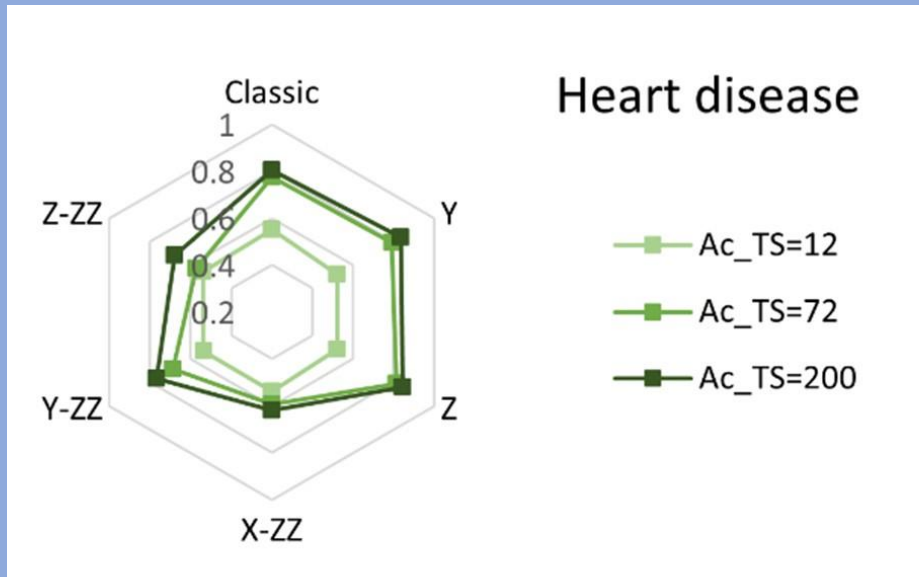
## ■ 量子カーネル法



\* <https://archive.ics.uci.edu/ml/index.php>  
<https://www.kaggle.com/>



# 古典/量子カーネル法の性能評価：もつれの効果



量子もつれを用いたパウリ-ZZ特徴マップの効果は  
今回用いた2つのデータセットでは効果がないことが分かった。



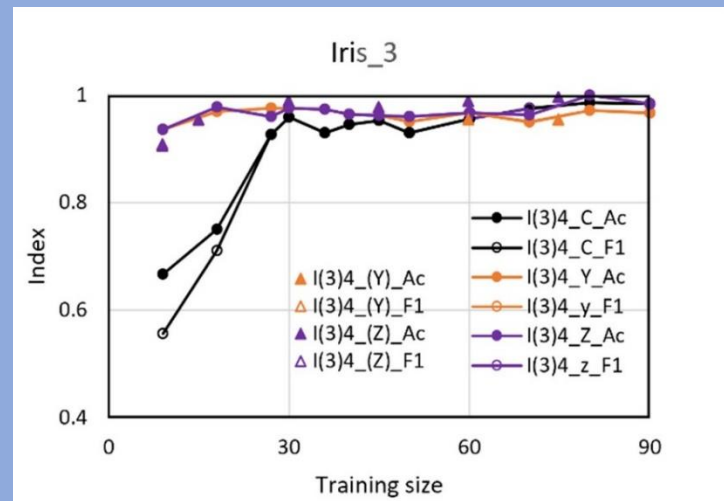
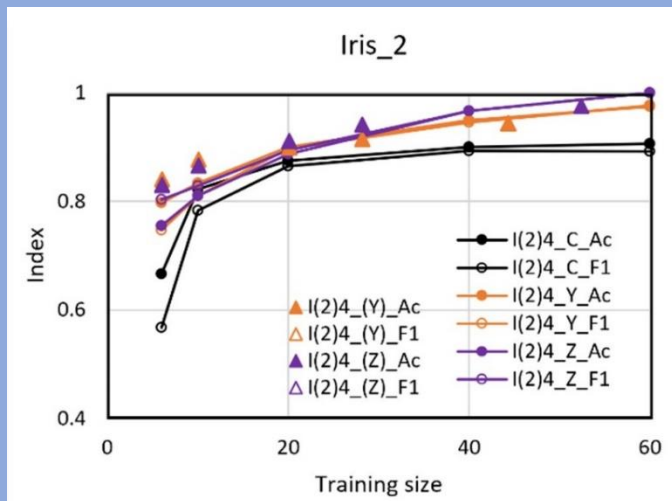
量子カーネルについてはパウリ特徴マップを用いることにした。

# パウリ特徴マップの効果について

Kernel	Heart disease		Iris_2		Wine		Iris_3	
Classic	Ac: 0.825	F1: 0.825	Ac: 0.750	F1: 0.667	Ac: 0.971	F1: 0.971	Ac: 1.000	F1: 1.000
	ts=240		ts=80		ts=144		ts=110	
Y	Ac: 0.857	F1: 0.857	Ac: 1.000	F1: 1.000	Ac: 1.000	F1: 1.000	Ac: 1.000	F1: 1.000
	ts=240		ts=75		ts=144		ts=110	
Z	Ac: 0.857	F1: 0.825	Ac: 1.000	F1: 1.000	Ac: 1.000	F1: 1.000	Ac: 1.000	F1: 1.000
	ts=240		ts=60		ts=144		ts=80	

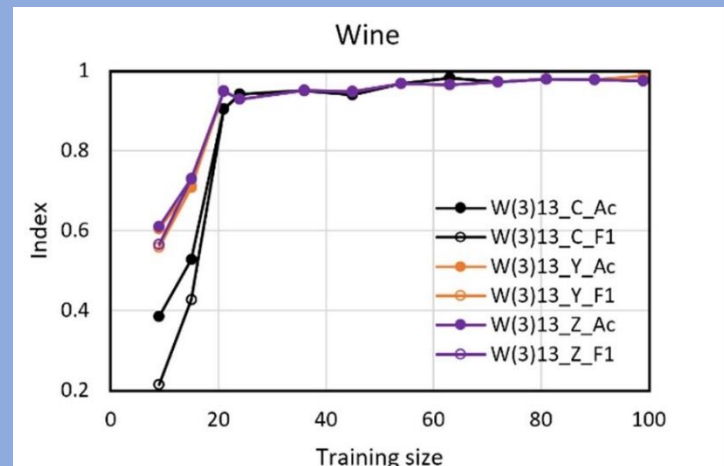
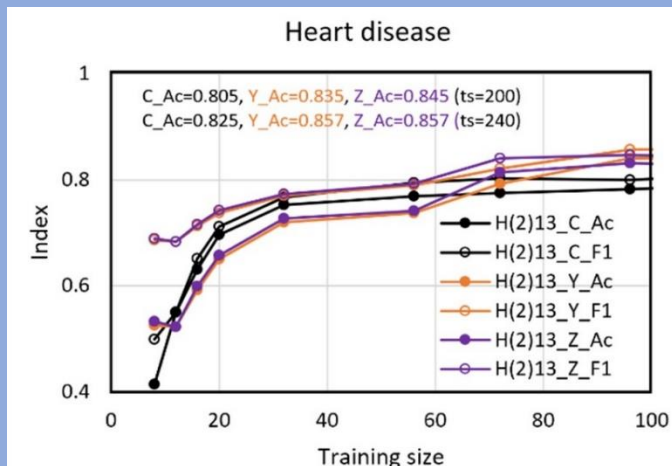
古典カーネルを用いた機械学習よりも、量子カーネルを用いた機械学習のほうが、速く、精度およびF1値が1に到達、もしくは**一定値に到達**を確認。

# 古典と量子カーネルによる性能評価：パウリゲートの効果



▲△ ◆◇ (はIBMの実機 (Bogota)。  
○● ○● ○● (はシミュレータ (Qiskit))

Irisでシミュレータと実機で計算  
ほぼ、両方の結果は一致を確認。



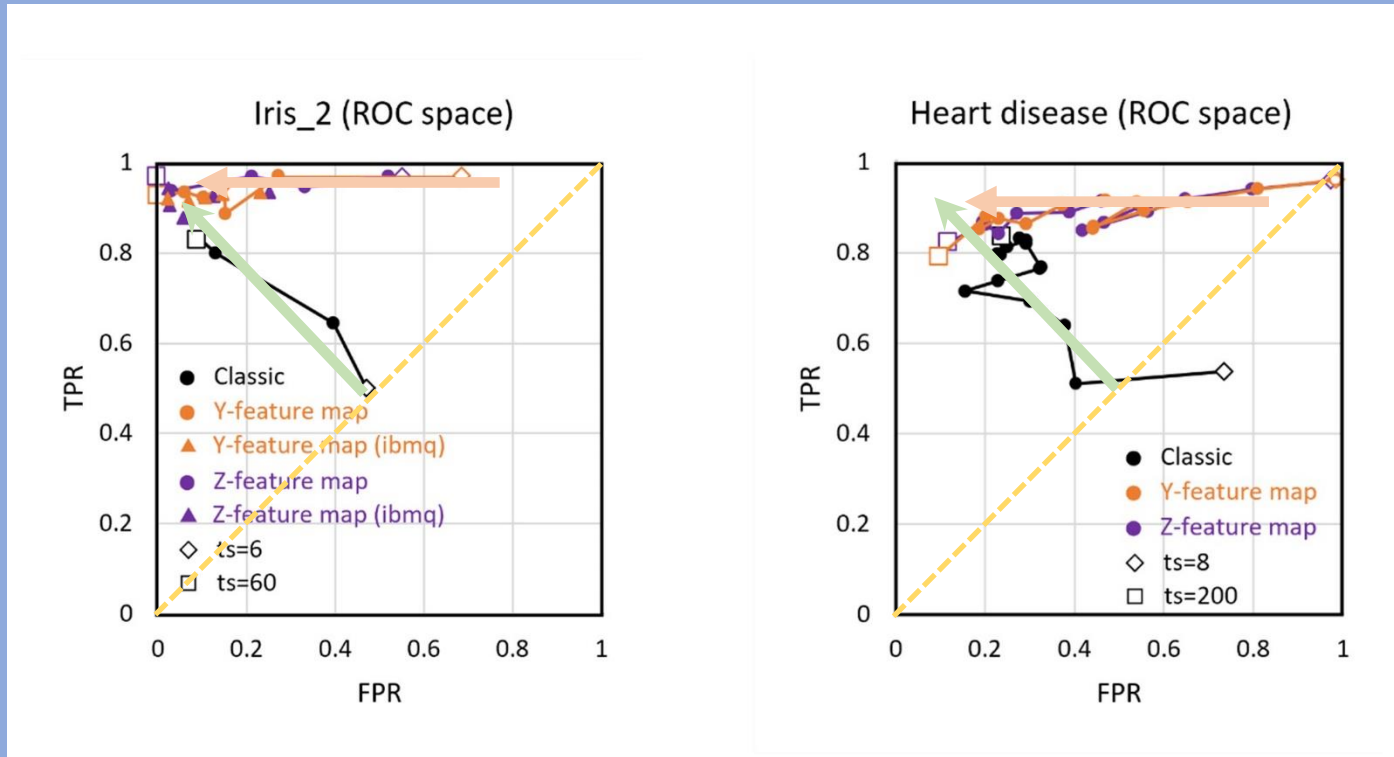
推測：

- 初期の学習  
訓練データ数が20で学習モデルの骨格形成
- 後半のゆっくりとした学習  
ゆっくりとロバスト性が増加

疑問：

精度とF1-scoreが初期状態で大きく異なる。

# 二値分類に注目して、ROC空間へのFPRとTPRのプロット



ランダムな学習モデルはこの位置にある。

古典カーネル法は、 $(FPR, TPR) = (0.5, 0.5) \Rightarrow (FPR, TPR) = (0, 1)$  に向かう。

量子カーネル法は、 $(FPR, TPR) = (1, 1) \Rightarrow (FPR, TPR) = (0, 1)$  に向かう。

# 訓練精度とテスト精度

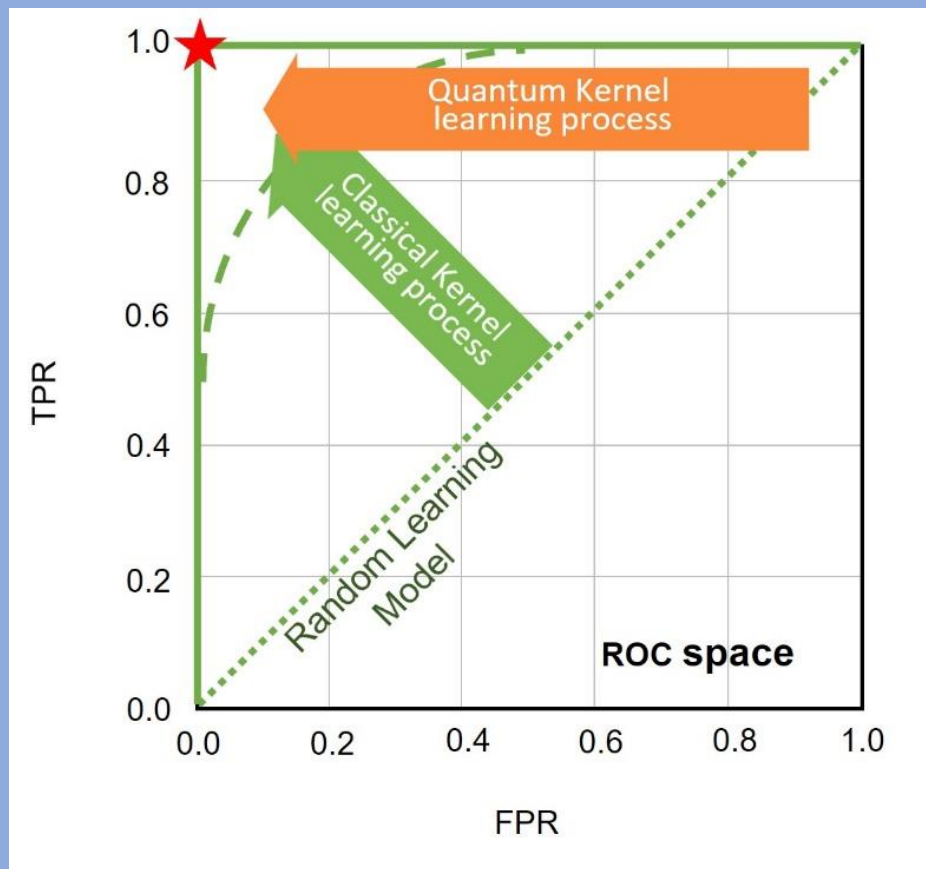
Training size	Accuracy: Classic		Accuracy: Y future map		Accuracy: Z future map	
	Training	Testing	Training	Testing	Training	Testing
12	0.667	0.550	1.000	0.522	1.000	0.523
72	0.889	0.775	0.986	0.792	0.986	0.814
200	0.905	0.806	0.980	0.835	0.980	0.845

古典カーネル法：訓練精度はテスト精度と同様に訓練サイズに応じて増加。  
一方、

量子カーネル法：訓練精度は1.00前後の状態を維持。

⇒探索空間領域が広い状態で学習が進むと考えられる。

# 仮説：量子カーネルと古典カーネルの学習過程の相違



古典カーネル法：

ランダムな学習モデルとして、緑の破線は一般的に周知。  
FPR, TPR = (0.5, 0.5) から (0, 1) に進むのは妥当な結果。

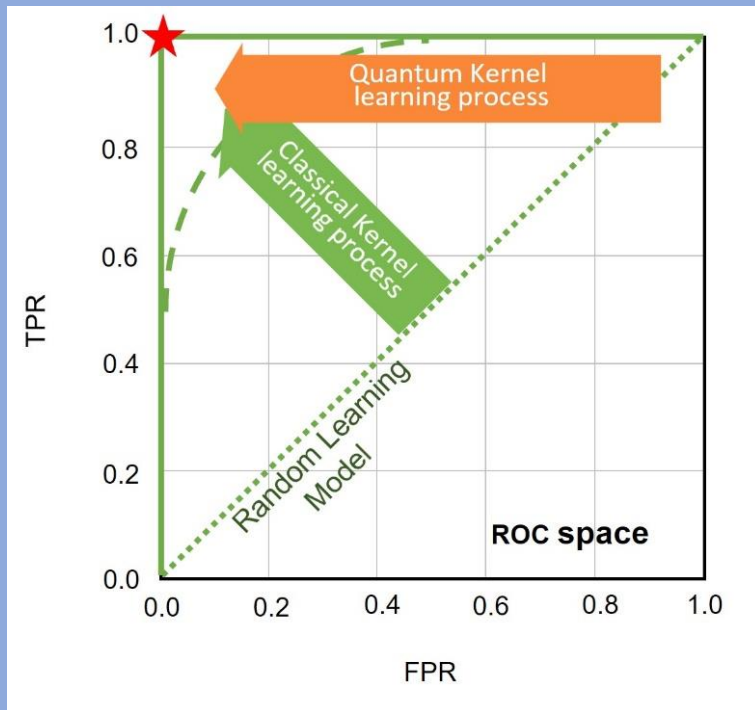
理想的な学習モデル：

★印であり、(FPR, TPR) = (0, 1)

量子カーネル法：

訓練精度の結果が示す通り、  
「探索空間領域(許容範囲)が広い状態から学習が進む  
そのため、(FPR, TPR) = (1, 1) から (0, 1) に学習が進み、  
TPRを高く維持しながら、FPRが減少する。」  
と考えています。

# 実課題適用上の考察



量子カーネル学習はTPRが高いところから始まる。

例：工場の不良品検知予測と実際の検査結果

## 検査による予測

	不良品	良品
不良品		「良品」と予測したのに、 実際は「不良品」だった
良品	「不良品」と予測したが、 実際は「良品」だった	

実際の結果

リコール率（TPR）は不良品（Positive）の疑いがある場合を漏れなく分類。

疑いのある良品は再度検査工程に回せばよい。

# 研究成果のまとめ

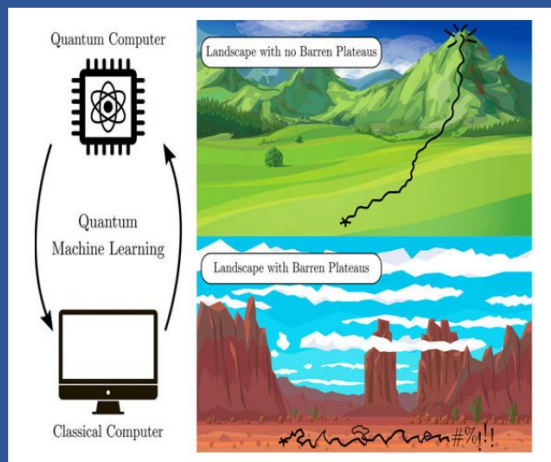
製造データを含む 2 値 3 値の複数のデータセットを用いて、  
古典と量子の機械学習モデルを各々構築した結果、  
ROC空間へのプロットより、初期状態での古典的カーネルと量子カーネルの学  
習プロセスの有意差があることを見出した。



# AGENDA

- 1) 量子コンピュータの可能性
- 2) 機械学習の評価指標
- 3) 量子/古典の学習モデル構築過程
- 4) 量子機械学習への期待

損失関数のランドスケープ内に台地のような平坦な形状が出てきて、最適化計算が進まない問題をBarren Plateauと呼ぶ。

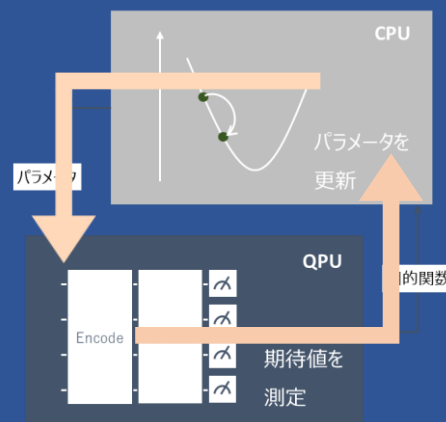


出典: phy.org<sup>4-1</sup>

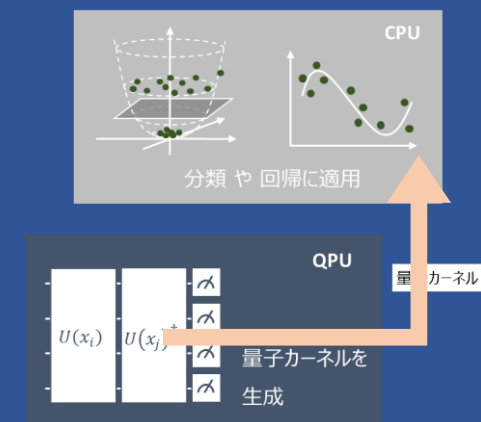
回避策:

- 浅い回路の活用<sup>4-2, 4-3</sup>
- 独自のarchitecturの活用<sup>4-4</sup>
- ローカルな観測で計算<sup>4-5</sup>
- 空間的、時間的に回路を工夫<sup>4-6</sup>
- 訓練の方法を工夫<sup>4-7</sup>

#### ■ 量子変分回路



#### ■ 量子カーネル法



量子変分回路では損失関数を評価しているのでBarren Plateauが起こりやすいが  
量子カーネルでは量子カーネルを古典アルゴリズムに埋め込んでいるのでBarren Plateausが起こりにくい。

4-1 <https://phys.org/news/2021-03-barren-plateaus-key-quantum-machine.html>

4-2 McClean, J.R., Boixo, S., Smelyanskiy, V.N. et al. Barren plateaus in quantum neural network training landscapes. Nat Commun 9, 4812 (2018).

4-3 Wang, S., Fontana, E., Cerezo, M. et al. Noise-induced barren plateaus in variational quantum algorithms. Nat Commun 12, 6961 (2021).

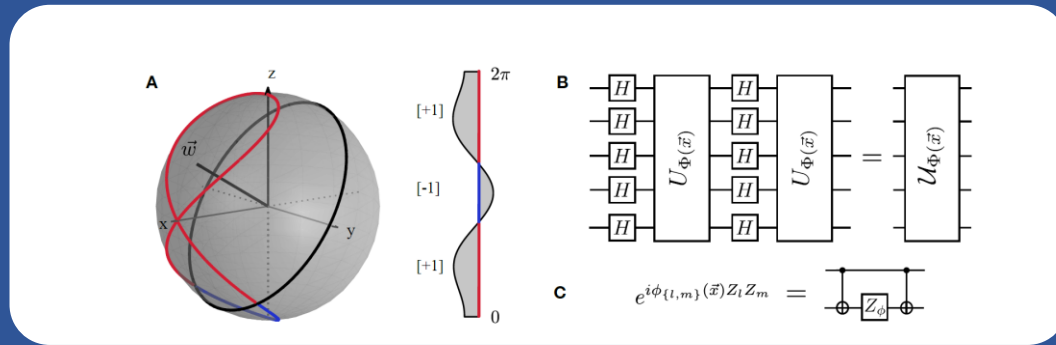
4-4 Arthur Pesah, M. Cerezo, et al. Absence of Barren Plateaus in Quantum Convolutional Neural Networks. Phys. Rev. X 11, 041011 (2021)

4-5 Cerezo, M., Sone, A., Volkoff, T. et al. Cost function dependent barren plateaus in shallow parametrized quantum circuits. Nat Commun 12, 1791 (2021).

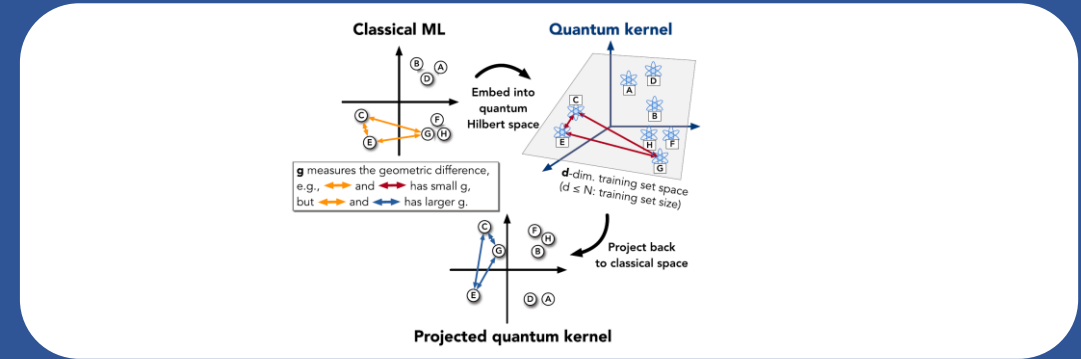
4-6 Tyler Volkoff and Patrick J Coles. Large gradients via correlation in random parameterized quantum circuits. Quantum Sci. Technol. 6 025008 (2021)

4-7 Skolik, A., McClean, J.R., Mohseni, M. et al. Layerwise learning for quantum neural networks. Quantum Mach. Intell. 3, 5 (2021).

# 量子カーネルとその課題



2019年にIBMがヒルベルト空間へ量子カーネル関数の埋込みを提案<sup>4-8</sup>。カーネル+SVMで、高い精度が得られることを実証。

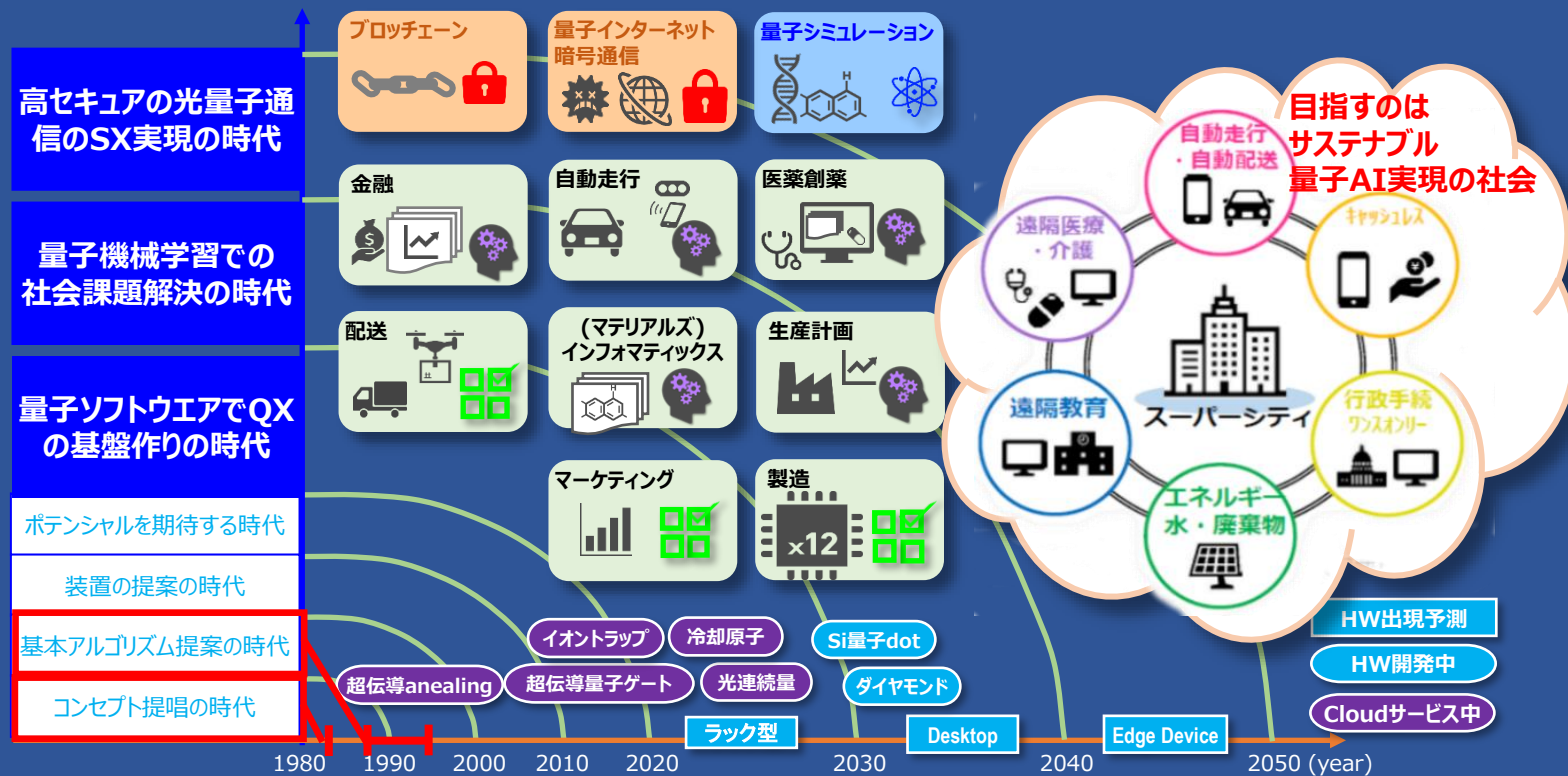


2021年にGoogleがカーネル関数が埋込まれた射影量子モデルを提案<sup>4-9</sup>。カーネル+NNで、古典よりも量子優位性を提示。

- モデル構築はデータの種類に依存する。
- 独自のカーネル関数を定義する必要がある。

カーネル法は、古典データを量子データには変換して、さらに、カーネル関数を埋込むことにより、分離面を学習により作る手法であり、更なる発展が今後期待される。

# 社会実装に向けたロードマップ (個人的見解)



現在は量子ソフトウェアでQXの基盤を作る時代と考えている。  
 実験データの蓄積により、NISQ時代に量子優位性を確立していく時代と考えている。

ロードマップは各種資料を基に、友野が2019年作成、2021revised  
 スーパーシティ構想は内閣府国家戦略特区スーパーシティ構想より引用  
<https://www5.cao.go.jp/keizai-shimon/kaigi/special/reform/wg6/190418/pdf/shiryous3-3.pdf>

# さいごに

近い将来、

- 量子コンピューターと古典コンピューターが共存
- 量子コンピューターが得意とする計算を識別することが重要
- 工場への量子機械学習の実装を検討する場合、データセットの選択が重要

そのためには、現在量子計算によるデータの蓄積が必要

今後、試行錯誤しながら有用な量子分類器などを構築し、社会実装に向けたトライアルを繰り返し実施たいと考えています。

謝辞

- 共同研究者である夏堀智子氏(量子/古典カーネルの学習モデル構築過程)に感謝します。
- 一部資料作成をしてくださった山中寛太氏(p8)、北澤優也氏(p10-13)に感謝します。

**TOPPAN**