

Variational Quantum Classifier의 소규모 데이터셋에서의 성능 비교: 고전적 머신러닝 모델 및 양자 커널 방법 대비 분석

[저자명]
전자공학부

대학교명

Email: [email]

[지도교수명]
전자공학부

대학교명

Email: [email]

Abstract—양자 머신러닝(Quantum Machine Learning, QML)은 양자 컴퓨팅의 고유한 특성을 활용하여 고전적 머신러닝의 한계를 극복할 수 있는 잠재력을 가진 분야로 주목받고 있다. 본 연구에서는 Variational Quantum Classifier(VQC)와 Quantum Support Vector Classifier(QSVC)를 구현하고, 세 가지 표준 벤치마크 데이터셋(Iris, Wine, Breast Cancer)에서 고전적 머신러닝 모델(SVM, MLP, Random Forest)과 체계적으로 성능을 비교하였다. 실험은 Qiskit 기반 시뮬레이터에서 수행되었으며, 데이터 인코딩 방식(Angle, ZZ), 변분 회로 구조(RealAmplitudes, EfficientSU2), 옵티마이저(COBYLA, SPSA), 회로 깊이(1-4 reps), 옵티마이저 수렴 분석, 양자 노이즈 영향 등 다양한 조건에서의 성능 변화를 분석하였다. 모든 실험은 10개의 랜덤 시드로 반복 수행하여 통계적 신뢰성을 확보하였으며, 모델 간 성능 차이는 Wilcoxon signed-rank test로 검정하였다. 실험 결과, VQC는 회로 깊이를 증가시켜도(reps=4) Iris 66.7%, Breast Cancer 87.8%에 머물러 고전적 모델(95% 이상) 대비 통계적으로 유의하게 낮았다($p < 0.005$). 반면 동일한 양자 특징 공간을 활용하는 QSVC는 Breast Cancer 데이터셋에서 $95.4 \pm 1.3\%$ 의 정확도를 달성하여 고전 SVM($95.2 \pm 1.7\%$)과 동등한 성능을 보였다. 수렴 분석에서는 COBYLA 옵티마이저가 maxiter 설정과 무관하게 약 70-85회 함수 평가 후 조기 수렴하여, VQC의 성능 병목이 변분 최적화 자체에 있음을 실증하였다. 본 연구는 NISQ 시대에서 양자 분류기의 접근법(변분 최적화 vs 양자 커널) 실질적 성능 차이를 분석한 기초 연구로서 의의를 갖는다.

Index Terms—양자 머신러닝, 변분 양자 분류기, 양자 커널, 벤치마크, NISQ, Qiskit

I. 서론

A. 연구 배경

양자 컴퓨팅은 큐비트(qubit)의 중첩(superposition)과 얽힘(entanglement) 특성을 활용하여 특정 문제에서 고전 컴퓨터 대비 지수적 속도 향상을 기대할 수 있는 새로운 계산 패러다임이다 [1]. 특히 양자 머신러닝(QML)은 양자 컴퓨팅과 머신러닝을 결합하여 데이터 분류, 최적화, 생성 모델 등의 문제에서 양자 이점(quantum advantage)을 탐색하는 연구 분야이다 [2].

NISQ(Noisy Intermediate-Scale Quantum) 시대의 양자 컴퓨터는 큐비트 수가 제한적이고 노이즈가 존재하여 완전한 양자 알고리즘 실행이 어렵다 [10]. 이러한 제약 하에서 변분 양자 알고리즘(Variational Quantum Algorithm, VQA)이 유력한 접근법으로 부상하였다 [3], [9]. VQA는 파라미터화된 양자 회로(Parameterized Quantum Circuit, PQC)와 고전 옵티마이저를 결합한 하이브리드 양자-고전 알고리즘으로 [11], 그 중 Variational Quantum Classifier(VQC)는 지도학습 분류 문제에 적용되는 대표적인 QML 모델이다.

B. 연구 목적

본 연구의 목적은 다음과 같다:

- 1) VQC와 QSVC를 구현하고 다양한 설정에서의 성능을 분석한다.
- 2) 동일한 데이터셋에서 고전적 ML 모델과 공정하게 비교한다.
- 3) 데이터셋 크기, 인코딩 방식, 회로 깊이, 옵티마이저 수렴, 양자 노이즈 등이 양자 분류기 성능에 미치는 영향을 분석한다.
- 4) 변분 최적화 방식(VQC)과 양자 커널 방식(QSVC)의 성능 차이를 비교한다.
- 5) NISQ 환경에서 양자 분류기의 실질적 가능성과 한계를 논의한다.

C. 연구 기여

본 연구의 주요 기여는 다음과 같다: (1) 동일한 양자 인코딩 하에서 변분 최적화 방식(VQC)과 양자 커널 방식(QSVC)의 성능 차이를 체계적으로 비교하여, 현재의 성능 병목이 양자 특징 공간이 아닌 최적화 방법론에 있음을 실증한 점, (2) VQC의 회로 깊이를 최대 4 reps까지 증가시켜도 고전 모델 대비 열위임을 통계적으로 검증한 점(Wilcoxon signed-rank test, $p < 0.005$), (3) 이진분류(Breast Cancer)와 다중분류(Iris, Wine)에서의 양자 분류기 성능 차이를 비교 분석한 점이다.

D. 연구 질문

소규모 데이터셋에서 양자 분류기(VQC, QSVC)가 고전적 ML 모델(SVM, MLP, Random Forest) 대비 어떤 조건에서 이점을 보이는가? 변분 최적화 방식과 양자 커널 방식의 성능 차이는 무엇에 기인하는가?

II. 이론적 배경

A. 양자 컴퓨팅 기초

양자 컴퓨팅의 기본 단위인 큐비트는 $|0\rangle$ 과 $|1\rangle$ 상태의 중첩으로 표현된다:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1 \quad (1)$$

양자 게이트는 큐비트의 상태를 유니타리 변환하며, 대표적으로 Pauli 게이트(X, Y, Z), Hadamard 게이트(H), 회전 게이트(R_X, R_Y, R_Z), CNOT 게이트 등이 있다.

B. Variational Quantum Classifier (VQC)

VQC는 세 단계로 구성된다 [4], [11]:

- 1) **데이터 인코딩 (Feature Map):** 고전 데이터 \mathbf{x} 를 양자 상태로 변환하는 유니타리 연산 $U(\mathbf{x})$
 - 2) **변분 회로 (Ansatz):** 학습 가능한 파라미터 θ 를 가진 양자 회로 $W(\theta)$
 - 3) **측정 및 분류:** 양자 상태를 측정하여 분류 결과를 얻음
- 전체 양자 상태는 다음과 같이 표현된다:

$$|\psi(\mathbf{x}, \theta)\rangle = W(\theta) \cdot U(\mathbf{x})|0\rangle^{\otimes n} \quad (2)$$

고전 옵티마이저가 손실 함수를 최소화하도록 파라미터 θ 를 반복적으로 업데이트한다.

C. 데이터 인코딩 방식

본 연구에서는 두 가지 인코딩 방식을 비교한다:

- **Angle Encoding:** 각 특징값을 R_Y 게이트의 회전 각도로 사용한다. 특징 수만큼의 큐비트가 필요하며, 구현이 단순하고 회로 깊이가 얕다.
- **ZZ Feature Map:** 특징 간 상호작용을 ZZ 게이트를 통해 인코딩한다 [5]. 비선형적 특징 공간 매핑이 가능하나 회로가 더 깊다.

D. 변분 회로 구조 (Ansatz)

- **RealAmplitudes:** R_Y 게이트와 CNOT 게이트로 구성된 하드웨어 효율적 ansatz [12]. 파라미터 수가 적어 학습이 빠르다.
- **EfficientSU2:** $SU(2)$ 회전 게이트(R_Y, R_Z)와 얽힘 게이트로 구성. 표현력이 더 높으나 파라미터 수가 많다.

E. Quantum Support Vector Classifier (QSVC)

QSVC는 양자 커널(quantum kernel)을 활용한 분류 방식으로, VQC와 근본적으로 다른 접근법을 취한다 [5]. 양자 커널 $K(x_i, x_j)$ 는 두 데이터 포인트의 양자 상태 간 충실도(fidelity)로 정의된다:

$$K(x_i, x_j) = |\langle 0|^{\otimes n} U(x_i)^\dagger U(x_j) |0\rangle^{\otimes n}|^2 \quad (3)$$

QSVC는 변분 파라미터 최적화가 불필요하므로 barren plateau 문제가 발생하지 않으며, 커널 행렬 계산 후 고전 SVM의 볼록 최적화를 활용하여 전역 최적해를 보장한다 [14].

F. 관련 연구

Schuld et al. [4]은 circuit-centric quantum classifier의 이론적 프레임워크를 제시하였다. Havlicek et al. [5]은 양자 강화 특징 공간에서의 지도학습을 제안하며, 양자 커널 방법의 이론적 기반을 마련하였다. Abbas et al. [6]은 양자 신경망의 표현력과 학습 가능성을 분석하여 barren plateau 문제를 논의하였다. Cerezo et al. [9]은 변분 양자 알고리즘의 포괄적 리뷰를 통해 비용 함수 지형, 최적화 전략, 표현력 간의 관계를 분석하였다. Huang et al. [13]은 양자 머신러닝에서 데이터의 역할을 분석하여, 고전 데이터에 대한 양자 이점의 조건을 이론적으로 제시하였다. Liu et al. [15]은 양자 지도학습에서 엄밀한 양자 속도 향상이 가능한 조건을 증명하였다.

III. 실험 방법

A. 데이터셋

세 가지 UCI 표준 벤치마크 데이터셋을 사용하였다 (Table I).

TABLE I
실험에 사용된 데이터셋

데이터셋	원본 특징 수	사용 특징 수	클래스 수	샘플 수
Iris	4	4	3	150
Wine	13	4 (PCA)	3	178
Breast Cancer	30	4 (PCA)	2	569

Wine과 Breast Cancer 데이터셋은 원본 특징 수가 큐비트 수를 초과하므로 PCA(주성분 분석)를 적용하여 4차원으로 축소하였다. PCA 적용 시 분산 기반 차원축소의 정확성을 위해 StandardScaler(평균 0, 표준편차 1)를 먼저 적용한 후 PCA를 수행하였다. 양자 인코딩에 필요한 $[0, \pi]$ 범위 변환은 PCA 이후 MinMaxScaler를 통해 수행하였다. 정규화 파라미터는 학습 데이터에서만 적합(fit)한 후 테스트 데이터에 적용(transform)하여 데이터 누출(data leakage)을 방지하였다. 데이터셋은 80:20 비율로 학습/테스트 분할하였다(층화 추출).

B. 양자 모델 설정

- **큐비트 수:** 4 (특징 수와 동일)
- **인코딩 방식:** Angle Encoding, ZZ Feature Map
- **Ansatz:** RealAmplitudes, EfficientSU2
- **옵티마이저:** COBYLA (gradient-free), SPSA (stochastic gradient)
- **최대 반복 횟수:** 200
- **회로 깊이(reps):** 1, 2, 3, 4
- **시뮬레이터:** Qiskit StatevectorSampler (이상적) 및 Qiskit Aer (노이즈)

C. QSVC 모델 설정

- **양자 커널:** FidelityQuantumKernel (Qiskit Machine Learning)
- **인코딩 방식:** Angle Encoding, ZZ Feature Map
- **고전 분류기:** SVM (precomputed kernel)

D. 고전 모델 설정

- **SVM:** RBF 커널, 기본 하이퍼파라미터
- **MLP:** 은닉층 (64, 32), 최대 500 에폭, ReLU 활성화
- **Random Forest:** 100개 결정 트리, 기본 하이퍼파라미터

E. 실험 설계

일곱 가지 실험을 수행하였다:

- **실험 1:** 데이터셋별 전체 모델 성능 비교 (3 데이터셋 \times 8 모델)
- **실험 2:** 학습 데이터 크기에 따른 성능 변화 (Iris 및 Breast Cancer, 20%–100%)
- **실험 3:** 옵티마이저 비교 (COBYLA vs SPSA, Iris 및 Breast Cancer)
- **실험 4:** 회로 깊이에 따른 성능 변화 (reps 1–4, Iris 및 Breast Cancer)
- **실험 5:** 양자 노이즈(depolarizing error)가 VQC 성능에 미치는 영향 (Iris 및 Breast Cancer)
- **실험 6:** 옵티마이저 수렴 분석 (maxiter 200, 500, 1000, Iris 및 Breast Cancer)
- **실험 7:** QSVC와 VQC 및 고전 모델 비교

F. 통계적 신뢰성

모든 실험은 10개의 랜덤 시드(0, 1, 7, 13, 21, 42, 77, 99, 123, 456)로 반복 수행하였다. 각 시드에서 데이터 분할, 모델 초기화, 최적화가 독립적으로 수행되었으며, 결과는 평균 \pm 표준편차(mean \pm std)로 보고한다. 모델 간 성능 차이의 통계적 유의성은 Wilcoxon signed-rank test ($\alpha = 0.05$)로 검정하였다.

G. 평가 지표

- **Test Accuracy:** 테스트셋에 대한 분류 정확도
- **F1-Score (Macro):** 클래스 불균형을 고려한 F1 점수
- **Training Time:** 모델 학습에 소요된 시간 (초)

H. 다중 클래스 분류 처리

Qiskit Machine Learning의 VQC는 다중 클래스 분류를 위해 parity mapping을 사용한다. n 개 큐비트의 측정 결과를 2^n 개 비트 문자열로 분류하며, 각 비트 문자열의 패리티(짝수/홀수 1의 개수)를 기준으로 클래스를 할당한다. 3클래스 문제(Iris, Wine)에서는 내부적으로 one-vs-rest 방식으로 확장된다.

I. 구현 환경

Python 3.12, Qiskit 2.3.0, Qiskit Machine Learning 0.9.0, Qiskit Aer 0.17.2, scikit-learn 1.8.0. Windows 11 로컬 PC에서 시뮬레이션 수행.

IV. 실험 결과

A. 실험 1: 데이터셋별 모델 성능 비교

모든 데이터셋에서 고전적 ML 모델이 VQC를 통계적으로 유의하게 상회하였다 (Wilcoxon, $p < 0.005$). VQC는 Breast Cancer 데이터셋(이진분류)에서 가장 좋은 성능을 보였으며, VQC-Angle-RA가 $86.6\pm 2.5\%$ 의 정확도를 달성하였으나 고전 모델($95.2\pm 1.7\%$) 대비 여전히 유의한 차이를 보였다. 최적 회로 설정인 VQC-Angle-RA-reps3도 Iris

TABLE II
데이터셋별 테스트 정확도 비교 (MEAN \pm STD, 10 SEEDS)

모델	Iris	Wine	Breast Cancer
SVM	.967\pm.030	.969\pm.026	.952 \pm .017
MLP	.967\pm.037	.958 \pm .031	.957\pm.012
RF	.957 \pm .037	.953 \pm .025	.939 \pm .019
VQC-Angle-RA	.643 \pm .033	.647 \pm .031	.866 \pm .025
VQC-Angle-ESU2	.590 \pm .042	.614 \pm .050	.864 \pm .037
VQC-Angle-RA-reps3	.667 \pm .000	.672 \pm .037	.858 \pm .041
VQC-ZZ-RA	.473 \pm .090	.428 \pm .052	.658 \pm .045
VQC-ZZ-ESU2	.457 \pm .065	.506 \pm .093	.768 \pm .078

TABLE III
데이터셋별 테스트 F1-SCORE (MACRO) 비교 (MEAN \pm STD, 10 SEEDS)

모델	Iris	Wine	Breast Cancer
SVM	.967\pm.030	.971\pm.025	.948 \pm .019
MLP	.966 \pm .037	.960 \pm .029	.954\pm.013
RF	.957 \pm .037	.955 \pm .024	.935 \pm .020
VQC-Angle-RA	.534 \pm .028	.516 \pm .026	.845 \pm .030
VQC-Angle-ESU2	.490 \pm .037	.489 \pm .041	.842 \pm .045
VQC-Angle-RA-reps3	.556 \pm .000	.531 \pm .029	.832 \pm .053
VQC-ZZ-RA	.380 \pm .079	.328 \pm .043	.637 \pm .041
VQC-ZZ-ESU2	.365 \pm .052	.394 \pm .074	.748 \pm .081

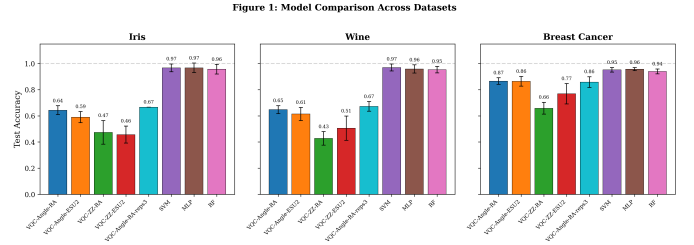


Fig. 1. 데이터셋별 모델 성능 비교

TABLE IV
학습 시간 비교 (초, MEAN \pm STD, 10 SEEDS)

모델	Iris	Wine	Breast Cancer
SVM	0.001 \pm 0.000	0.001 \pm 0.000	0.003 \pm 0.002
MLP	0.210 \pm 0.025	0.235 \pm 0.023	0.314 \pm 0.044
RF	0.108 \pm 0.004	0.111 \pm 0.002	0.123 \pm 0.004
VQC-Angle-RA	38.1 \pm 4.8	51.4 \pm 8.4	143.5 \pm 14.5
VQC-Angle-ESU2	66.7 \pm 5.4	97.3 \pm 15.8	300.3 \pm 24.1
VQC-Angle-RA-reps3	93.3 \pm 7.4	109.5 \pm 12.4	335.3 \pm 46.3
VQC-ZZ-RA	41.6 \pm 5.2	49.9 \pm 3.8	168.9 \pm 12.9
VQC-ZZ-ESU2	86.1 \pm 8.9	109.1 \pm 7.0	405.6 \pm 50.2

66.7%, Wine 67.2%, BC 85.8%에 머물렀다. F1-Score에서도 동일한 경향이 관찰되었으며, VQC의 F1-Score는 정확도보다 더 낮은 클래스별 분류 성능의 불균형을 시사한다. 인코딩 방식별로는 Angle Encoding이 ZZ Feature Map보다 전반적으로 우수하였다.

VQC의 학습 시간은 고전 모델 대비 수백~수만 배 느렸으며, EfficientSU2 ansatz는 파라미터 수가 많아 RealAmplitudes보다 약 2–2.5배의 학습 시간이 소요되었다. 전체 모델

Figure 2: Training Time Comparison

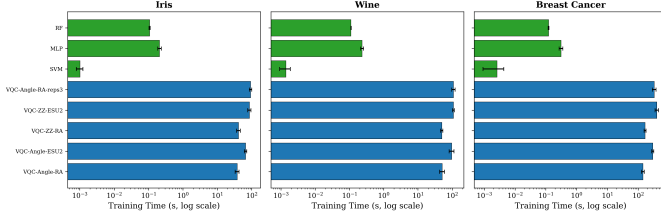


Fig. 2. 학습 시간 비교 (로그 스케일)

과 데이터셋에 대한 정확도 히트맵은 Fig. 6에 제시하였다.

B. 실험 2: 학습 데이터 크기에 따른 성능 변화

TABLE V
학습 데이터 비율별 테스트 정확도 (MEAN \pm STD, 10 SEEDS)

데이터셋	비율	VQC	SVM	MLP	RF
Iris	20%	.593 \pm .051	.930 \pm .038	.947 \pm .031	.940 \pm .033
	40%	.627 \pm .039	.960 \pm .029	.967 \pm .021	.960 \pm .020
	60%	.633 \pm .033	.980 \pm .016	.970 \pm .031	.967 \pm .021
	80%	.650 \pm .017	.963 \pm .031	.970 \pm .031	.960 \pm .029
	100%	.620 \pm .050	.967 \pm .030	.967 \pm .037	.957 \pm .037
BC	20%	.861 \pm .043	.941 \pm .023	.946 \pm .017	.931 \pm .014
	40%	.869 \pm .024	.944 \pm .018	.955 \pm .010	.930 \pm .022
	60%	.868 \pm .017	.950 \pm .021	.959 \pm .010	.935 \pm .023
	80%	.873 \pm .019	.954 \pm .020	.958 \pm .012	.940 \pm .017
	100%	.875 \pm .012	.952 \pm .017	.957 \pm .012	.939 \pm .019

Figure 3: Effect of Training Data Size on Performance

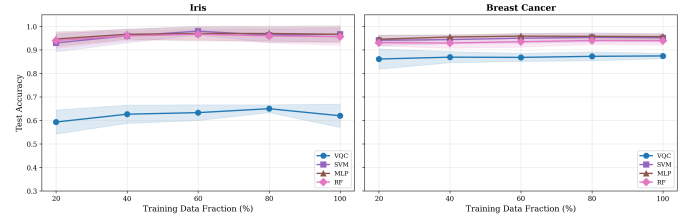


Fig. 3. 학습 데이터 크기에 따른 성능 변화 (Iris 및 Breast Cancer)

고전 모델은 20%의 데이터만으로도 93% 이상의 정확도를 달성하였다. VQC는 Iris에서 데이터 양 증가에 따른 명확한 성능 향상 경향을 보이지 않았으며, 59.3%–65.0% 범위에서 정체되었다. Breast Cancer(이진분류)에서는 VQC가 86.1%–87.5%로 상대적으로 높으나 역시 데이터 증가에 의한 개선은 미미하였다. 이는 VQC의 성능 병목이 데이터가 아닌 최적화에 있음을 시사한다.

C. 실험 3: 옵티마이저 비교

TABLE VI
COBYLA vs SPSSA 비교 (MAXITER=200, MEAN \pm STD, 10 SEEDS)

데이터셋	옵티마이저	Test Accuracy	Time (s)
Iris	COBYLA	0.630 \pm 0.038	36.4
	SPSA	0.617 \pm 0.037	212.9
BC	COBYLA	0.844 \pm 0.086	158.0
	SPSA	0.868 \pm 0.024	769.9

Iris에서는 COBYLA가 SPSSA보다 약 1.3%p 높은 정확도를 달성하였으나, Breast Cancer에서는 SPSSA(86.8%)가 COBYLA(84.4%)보다 높았다. 그러나 학습 시간은 SPSSA가 COBYLA 대비 약 5–6배 소요되었다.

Figure 4: Optimizer Comparison

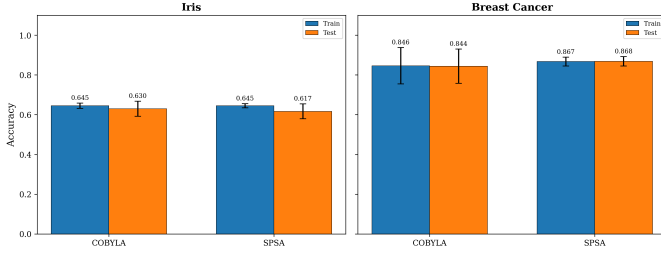


Fig. 4. 옵티마이저 비교: 정확도 및 학습 시간 (Iris 및 Breast Cancer)

D. 실험 4: 회로 깊이에 따른 성능 변화

TABLE VII

회로 깊이 (REPS) 별 성능 (COBYLA, MAXITER=200, MEAN \pm STD, 10 SEEDS)

데이터셋	Reps	Test Accuracy	Time (s)
Iris	1	0.627 \pm 0.039	35.2
	2	0.653 \pm 0.022	65.0
	3	0.663 \pm 0.010	85.8
	4	0.667\pm0.000	108.3
BC	1	0.869 \pm 0.026	140.3
	2	0.853 \pm 0.031	229.6
	3	0.861 \pm 0.036	321.8
	4	0.878\pm0.020	410.0

Figure 5: Effect of Circuit Depth

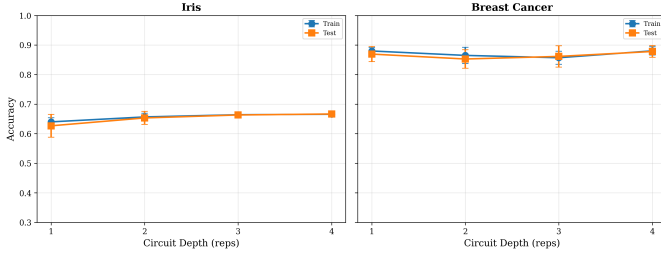
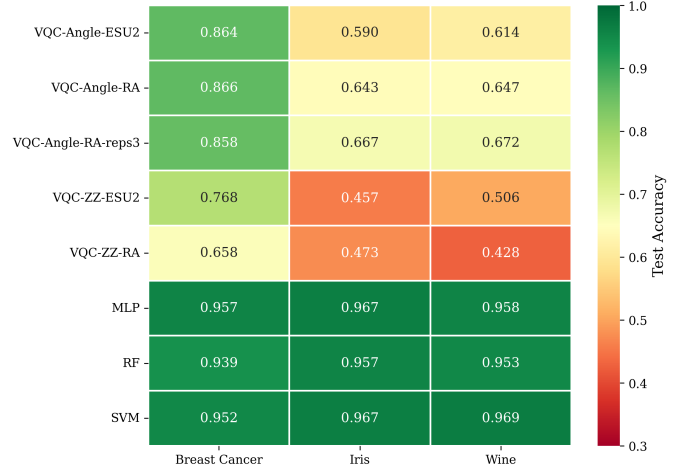


Fig. 5. 회로 깊이에 따른 성능 변화 (Iris 및 Breast Cancer)

Iris에서 reps=4에서 최고 성능(66.7 \pm 0.0%)을 달성하였으며, 10개 시드 모두 동일한 결과를 보여 가장 안정적인 수렴을 나타냈다. Breast Cancer에서도 reps=4에서 87.8 \pm 2.0%로 최고 성능을 보였다. 4큐비트 소규모 회로에서는 barren plateau 현상이 관찰되지 않았으며 [7], 이는 Sim et al. [20]이 분석한 저큐비트 회로의 높은 표현력과 일치한다.

Figure 6: Test Accuracy Heatmap (All Models x Datasets)

Fig. 6. 테스트 정확도 히트맵 (전체 모델 \times 데이터셋)

E. 실험 5: 양자 노이즈 영향

TABLE VIII

DEPOLARIZING 노이즈 수준별 VQC 성능 (MEAN \pm STD, 10 SEEDS)

데이터셋	Error Rate	Test Accuracy	Test F1
Iris	Ideal	.640\pm.033	.531\pm.029
	$p = 0.001$.623 \pm .047	.517 \pm .042
	$p = 0.005$.607 \pm .044	.503 \pm .040
	$p = 0.01$.600 \pm .033	.497 \pm .029
	$p = 0.05$.560 \pm .044	.464 \pm .040
BC	Ideal	.871\pm.026	.851\pm.032
	$p = 0.001$.868 \pm .019	.847 \pm .024
	$p = 0.005$.874 \pm .028	.854 \pm .034
	$p = 0.01$.873 \pm .030	.852 \pm .039
	$p = 0.05$.822 \pm .044	.780 \pm .067

Figure 7: Impact of Quantum Noise on VQC Performance

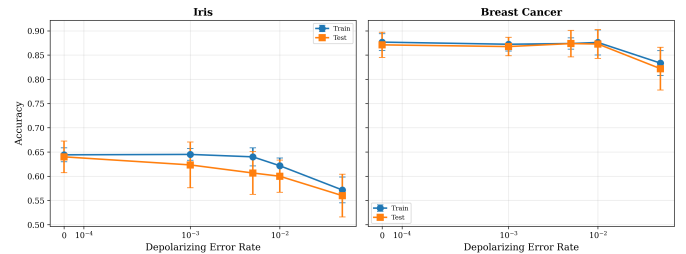


Fig. 7. 양자 노이즈가 VQC 성능에 미치는 영향 (Iris 및 Breast Cancer)

Depolarizing 노이즈가 증가함에 따라 VQC 성능이 전반적으로 감소하였다. Iris에서는 이상적 조건 64.0%가 $p = 0.05$ 에서 56.0%로 약 8%p 단조 하락하였다. Breast Cancer에서는 $p = 0.05$ 에서 82.2%로 이상적 조건(87.1%) 대비 약 5%p 하락하였으나, 중간 노이즈 수준($p = 0.005$, $p = 0.01$)에서는 표준편차 범위 내의 변동으로 비단조적 양상을 보였다. 높은 노이즈 수준에서의 성능 저하는 NISQ 환

경에서 양자 오류 완화(error mitigation) 기술의 중요성을 확인한다 [8]. 본 실험의 이상적 조건은 AerSimulator(shot noise 포함)를 사용하였으며, 이는 실험 1의 Statevector-Sampler(shot noise 없음)와 다른 시뮬레이터이므로 실험 1과 직접 비교할 수 없다.

F. 실험 6: 옵티마이저 수렴 분석

TABLE IX
MAXITER별 VQC 성능 및 수렴 특성 (COBYLA, MEAN \pm STD, 10 SEEDS)

데이터셋	maxiter	Test Accuracy	Time (s)	평가 횟수
Iris	200	0.643 \pm 0.026	53.6	79.2
	500	0.647 \pm 0.027	51.4	78.3
	1000	0.620 \pm 0.040	48.4	74.2
BC	200	0.879 \pm 0.017	217.0	85.9
	500	0.873 \pm 0.025	209.3	80.9
	1000	0.868 \pm 0.034	217.5	82.7

Figure 8: Convergence Analysis

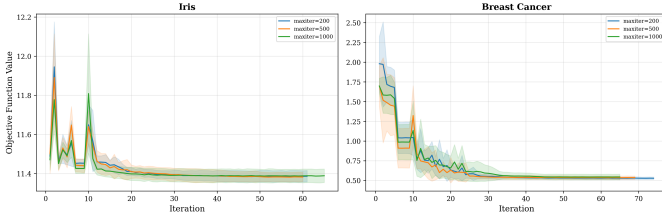


Fig. 8. 옵티마이저 수렴 분석: (좌) 학습 손실 수렴 곡선, (우) maxiter 별 최종 정확도

COBYLA 옵티마이저는 maxiter 설정과 무관하게 약 74–86회의 함수 평가 후 자체 수렴 기준에 도달하여 조기 종료하였다. Iris와 Breast Cancer 모두에서 학습 시간이 세 조건에서 거의 동일한 점에서도 이를 확인할 수 있다. maxiter를 5배(1000)로 증가시켜도 정확도가 개선되지 않고 오히려 소폭 하락하였으며(Iris: 64.3% \rightarrow 62.0%, BC: 87.9% \rightarrow 86.8%), 이는 VQC의 성능 병목이 반복 횟수 부족이 아닌 비블록 손실 지형의 구조적 한계에 기인함을 시사한다.

G. 실험 7: QSVC (양자 커널 SVM) 비교

TABLE X
QSVC vs VQC vs CLASSICAL SVM 비교 (MEAN \pm STD, 10 SEEDS)

모델	Iris	Wine	Breast Cancer
QSVC-Angle	.967\pm.037	.969\pm.026	.954\pm.013
QSVC-ZZ	.927 \pm .039	.861 \pm .071	.903 \pm .032
VQC-Angle-RA	.643 \pm .033	.647 \pm .031	.866 \pm .025
SVM (Classical)	.967 \pm .030	.969 \pm .026	.952 \pm .017

QSVC-Angle은 모든 데이터셋에서 VQC 대비 9–32%p 높은 정확도를 달성하였으며, 고전 SVM과 거의 동등한 성능을 보였다. Breast Cancer에서는 QSVC-Angle(95.4 \pm 1.3%)이 고전 SVM(95.2 \pm 1.7%)과 동등한 정확도를 달성하였다. 이 결과는 VQC의 낮은 성능이 양자 특징 공간의 한계가 아닌 변분 최적화 과정의 한계에 기인함을 강력히 시사한다.

Figure 9: VQC vs QSVC vs Classical Models

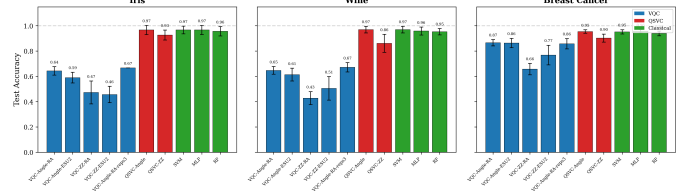


Fig. 9. VQC vs QSVC vs 고전 모델 비교

TABLE XI

통계적 유의성 검정 (최적 VQC vs 고전 모델, WILCOXON SIGNED-RANK TEST)

데이터셋	비교	VQC Acc	Classical Acc	p-value
Iris	reps3 vs SVM	.667	.967	.002**
Iris	reps3 vs MLP	.667	.967	.002**
Iris	reps3 vs RF	.667	.957	.002**
Wine	reps3 vs SVM	.672	.969	.002**
Wine	reps3 vs MLP	.672	.958	.002**
Wine	reps3 vs RF	.672	.953	.002**
BC	RA vs SVM	.866	.952	.002**
BC	RA vs MLP	.866	.957	.002**
BC	RA vs RF	.866	.939	.002**

V. 논의

A. VQC 성능 분석 및 최적화 병목

VQC의 성능 병목 원인을 다각도로 분석한 결과:

1) **비블록 최적화와 조기 수렴**: 실험 6에서 COBYLA가 maxiter와 무관하게 약 74–86회 평가 후 조기 수렴하였으며, 실험 2에서 데이터 크기 증가도 VQC 성능을 개선하지 못하였다. 이는 변분 양자 알고리즘에서 널리 보고된 최적화 지형의 문제와 일치한다 [9].

2) **QSVC와의 비교로 확인**: 실험 7에서 동일한 Angle Encoding을 사용하면서 양자 커널 방식을 적용한 QSVC가 VQC 대비 9–32%p 높은 정확도를 달성하여, 양자 특징 공간 자체의 표현력은 충분하나 변분 최적화가 이를 효과적으로 활용하지 못함을 확인하였다 [17].

3) **인코딩 방식의 영향**: Angle Encoding이 ZZ Feature Map보다 VQC와 QSVC 모두에서 우수하여, 이 차이가 최적화 방식에 독립적임을 확인하였다.

B. 양자 노이즈의 영향

실험 5에서 높은 수준의 depolarizing 노이즈 ($p = 0.05$)가 VQC 성능을 유의하게 저하시켰음을 확인하였다. Iris에서 64.0% \rightarrow 56.0%, Breast Cancer에서 87.1% \rightarrow 82.2%로 하락하여 [19], NISQ 환경에서 양자 오류 완화 기술의 필요성을 확인하였다.

C. VQC vs QSVC: 접근법별 분석

최적화 방식: VQC는 비블록 최적화(COBYLA)로 지역 최솟값에 빠질 수 있으나, QSVC는 블록 최적화(SVM dual)로 전역 최적해가 보장된다 [14].

확장성: VQC는 데이터 크기에 선형 비례하나, QSVC는 커널 행렬 계산에 $O(n^2)$ 소요. 대규모 데이터에서는 VQC가 유리할 수 있다 [18].

D. 계산 비용 분석

QSVC가 정확도 측면에서 고전 SVM과 동등한 성능을 보였으나, 학습 시간에서는 현저한 차이가 존재한다. Breast Cancer 데이터셋에서 QSVC-Angle의 학습 시간은 약 502초로, 고전 SVM(0.003초) 대비 약 17만 배 느렸다. 이는 양자 커널 행렬 계산의 $O(n^2)$ 복잡도에 기인하며, 현재의 고전 시뮬레이터 기반 실험에서는 양자 하드웨어의 병렬 상태 탐색이 발현되지 않는다. 따라서 QSVC의 실용적 가치는 고전적으로 시뮬레이션이 어려운 커널을 사용하거나, 실제 양자 하드웨어에서 실행할 때 평가되어야 한다.

E. 양자 이점에 대한 논의

분류 정확도 측면의 양자 이점은 관찰되지 않았으며, 이는 Bowles et al. [16]이 지적한 양자 ML 벤치마크의 일반적인 경향과 일치한다. 그러나 QSVC가 고전 SVM과 동등한 성능을 달성한 것은 양자 특징 공간의 유효성을 보여준다. Huang et al. [13]이 지적한 바와 같이, 고전 데이터에 대한 양자 이점은 데이터의 구조와 양자 모델의 귀납적 편향이 일치할 때 발현될 수 있다. VQC와 QSVC의 성능 차이는 양자 컴퓨팅의 잠재력이 아닌 최적화 방법론의 선택이 현재의 병목임을 시사한다. Liu et al. [15]이 증명한 양자 속도 향상의 조건은 본 실험의 소규모 정형 데이터셋에는 해당되지 않으며, 양자 이점이 발현되는 문제 구조에 대한 추가 연구가 필요하다.

F. 연구의 한계

- 1) Depolarizing 노이즈만 사용하여, 실제 하드웨어의 복합적 노이즈를 완전히 반영하지 못한다.
- 2) 단일 gradient-free 최적화(COBYLA/SPSA)만 사용. Gradient-based optimizer에서는 다른 결과가 나올 수 있다.
- 3) PCA 차원축소로 인한 정보 손실이 모든 모델에 영향을 미쳤을 수 있다.
- 4) 소규모 데이터셋만 사용하였다.

VI. 결론

본 연구에서는 VQC와 QSVC를 세 가지 표준 벤치마크 데이터셋에서 구현하고, 고전적 ML 모델과 체계적으로 비교하였다. 주요 결론:

- 1) VQC는 회로 깊이를 최대 4 reps까지 증가시켜도 고전적 ML 모델 대비 통계적으로 유의하게 낮은 성능을 보였다 (Wilcoxon, $p < 0.005$). 반면 QSVC-Angle은 Breast Cancer에서 $95.4 \pm 1.3\%$ 로 고전 SVM($95.2 \pm 1.7\%$)과 동등하였다.
- 2) VQC의 성능 병목은 변분 최적화의 구조적 한계이다. COBYLA가 maxiter와 무관하게 ~ 74 – 86 회 평가 후 조기 수렴하였다.
- 3) Angle Encoding이 ZZ Feature Map보다 VQC와 QSVC 모두에서 안정적이다.
- 4) 높은 수준의 양자 노이즈($p = 0.05$)는 VQC 성능을 유의하게 저하시킨다. Iris: $64.0\% \rightarrow 56.0\%$, BC: $87.1\% \rightarrow 82.2\%$.
- 5) 이진분류(Breast Cancer)에서 VQC 성능이 다중분류(Iris, Wine) 대비 현저히 높다 (86.6% vs 64.3 – 67.2%).

향후 연구에서는 (1) 실제 IBM Quantum 하드웨어 실험, (2) gradient-based optimizer를 활용한 VQC 최적화 개선,

(3) 양자 오류 완화 적용, (4) QSVC 확장성 분석, (5) 더 많은 큐비트에서의 barren plateau 분석 등을 통해 양자 분류기의 잠재력을 더 깊이 탐색할 계획이다.

REFERENCES

- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2010.
- [2] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” *Nature*, vol. 549, pp. 195–202, 2017.
- [3] A. Peruzzo et al., “A variational eigenvalue solver on a photonic quantum processor,” *Nature Communications*, vol. 5, p. 4213, 2014.
- [4] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, “Circuit-centric quantum classifiers,” *Physical Review A*, vol. 101, no. 3, p. 032308, 2020.
- [5] V. Havlicek et al., “Supervised learning with quantum-enhanced feature spaces,” *Nature*, vol. 567, pp. 209–212, 2019.
- [6] A. Abbas et al., “The power of quantum neural networks,” *Nature Computational Science*, vol. 1, pp. 403–409, 2021.
- [7] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, “Barren plateaus in quantum neural network training landscapes,” *Nature Communications*, vol. 9, p. 4812, 2018.
- [8] Y. Li and S. C. Benjamin, “Efficient variational quantum simulator incorporating active error minimisation,” *Physical Review X*, vol. 7, no. 2, p. 021050, 2017.
- [9] M. Cerezo et al., “Variational quantum algorithms,” *Nature Reviews Physics*, vol. 3, pp. 625–644, 2021.
- [10] J. Preskill, “Quantum computing in the NISQ era and beyond,” *Quantum*, vol. 2, p. 79, 2018.
- [11] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, “Parameterized quantum circuits as machine learning models,” *Quantum Science and Technology*, vol. 4, no. 4, p. 043001, 2019.
- [12] A. Kandala et al., “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets,” *Nature*, vol. 549, pp. 242–246, 2017.
- [13] H.-Y. Huang et al., “Power of data in quantum machine learning,” *Nature Communications*, vol. 12, p. 2631, 2021.
- [14] J. M. Kübler, S. Buchholz, and B. Schölkopf, “The inductive bias of quantum kernels,” in *Advances in Neural Information Processing Systems*, vol. 34, 2021.
- [15] Y. Liu et al., “A rigorous and robust quantum speed-up in supervised machine learning,” *Nature Physics*, vol. 17, pp. 1013–1017, 2021.
- [16] J. Bowles et al., “Better than classical? The subtle art of benchmarking quantum machine learning models,” arXiv:2403.07059, 2024.
- [17] M. Schuld, “Supervised quantum machine learning models are kernel methods,” *Journal of Machine Learning Research*, vol. 23, pp. 1–32, 2022.
- [18] S. Thanasilp, S. Wang, M. Cerezo, and Z. Holmes, “Exponential concentration in quantum kernel methods,” *Nature Communications*, vol. 15, p. 5200, 2024.
- [19] IBM Quantum, “IBM Quantum System Performance,” Technical Report, 2023.
- [20] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, “Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms,” *Advanced Quantum Technologies*, vol. 2, no. 12, p. 1900070, 2019.