ISSN 2464-4617 (print)
ISSN 2464-4625 (online)
Quantum Informatics, Computing & Technology 2025
QC-Horizon 2025
https://www.qc-horizon.eu/

Quantum-inspired classifiers

Roberto Leporini
University of Bergamo
via dei Caniana, 2
Italy, 24127 Bergamo, BG
roberto.leporini@unibg.it

Cesarino Bertini
University of Bergamo
via dei Caniana, 2
Italy, 24127 Bergamo, BG
cesarino.bertini@unibg.it

ABSTRACT

Quantum-inspired machine learning is a new branch of machine learning based on the application of the mathematical formalism of quantum mechanics to devise novel algorithms for classical computers. We implement some quantum-inspired classification algorithms, based on quantum state discrimination, within a local approach in the feature space by taking into account elements close to the element to be classified. This local approach improves the accuracy in classification and motivates the integration with the classifiers. The quantum-inspired classifiers require the encoding of the feature vectors into density operators and methods for estimating the distinguishability of quantum states like the Helstrom state discrimination and the Pretty-Good measurement. We present a comparison of the performances of the local quantum-inspired classifiers against well-known classical algorithms in order to show that the local approach can be a valuable tool for increasing the performances of this kind of classifiers.

Keywords

quantum-inspired classifiers; machine learning; local approach.

1 INTRODUCTION

Quantum-inspired machine learning represents a novel area within machine learning that leverages the mathematical framework of quantum mechanics to develop new algorithms for classical computers. In this work, we implement several quantum-inspired classification algorithms rooted in quantum state discrimination, employing a local strategy within the feature space. Specifically, we implement quantum-inspired algorithms based on Helstrom discrimination. The method involves classifying an unlabeled data instance by identifying its k nearest training elements before applying the algorithm to these k neighbors. This local approach enhances classification accuracy.

Quantum-inspired classifiers necessitate the encoding of feature vectors into density operators and methods for assessing the distinguishability of quantum states, such as Helstrom state discrimination and Pretty-Good Measurement (PGM). In the experimental section, we present a performance comparison between our local quantum-inspired classifiers and established classical algorithms. The aim is to demonstrate the potential of the local approach as a valuable technique for improving the performance of these types of classifiers.

2 QUANTUM-INSPIRED CLASSIFICATION

The initial stage in quantum-inspired classification involves quantum encoding, which encompasses any method for mapping classical information into quantum states. Specifically, we consider encoding data vectors into density matrices within a Hilbert space H whose dimensionality is determined by the input space's dimension. Density matrices are positive semidefinite operators ρ with a trace of 1 and they serve as the mathematical tools for describing the physical states of quantum systems.

Pure states are a subset of density matrices. These are rank-1 projectors that can be directly associated with unit vectors up to a phase factor. A density operator ρ on a d-dimensional Hilbert space $\mathbb{C}^d$ can be expressed as:

$$\rho = \frac{1}{d}\Big(\mathtt{I}_d + \sqrt{\frac{d(d-1)}{2}}\sum_{j=1}^{d^2-1} b_j^{(\rho)}\sigma_j\Big), \qquad (1)$$

where $\{\sigma_j\}_{j=1,\ldots,d^2-1}$ are the standard generators of the special unitary group $SU(d)$, also known as generalized Pauli matrices, and $\mathtt{I}_d$ is the $d\times d$ identity matrix. The vector $\mathbf{b}^{(\rho)} = (b_1^{(\rho)},\ldots,b_{d^2-1}^{(\rho)})$, with $b_j^{(\rho)} = \sqrt{\frac{d}{2(d-1)}}\mathrm{tr}(\rho\,\sigma_j) \in \mathbb{R}$, is the Bloch vector associated to ρ which lies within the hypersphere of radius 1 in $\mathbb{R}^{d^2-1}$. For $d=2$, the qubit case, the density matrices

are in bijective correspondence to the points of the unit ball in $\mathbb{R}^3$, the so-called Bloch sphere, where the pure states are in one-to-one correspondence with the points of the spherical surface.

Complex vectors of dimension n can be encoded into density matrices of a $(n+1)$-dimensional Hilbert space H in the following way:

$$\mathbb{C}^n \ni \mathbf{x} \mapsto |\mathbf{x}\rangle = \frac{1}{\sqrt{\| \mathbf{x} \|^2 + 1}} \left(\sum_{\alpha=0}^{n-1} x_\alpha |\alpha\rangle + |n\rangle \right) \in \mathsf{H}, \tag{2}$$

where $\{|\alpha\rangle\}_{\alpha=0,...,n}$ is the computational basis of H, identified as the standard basis of $\mathbb{C}^{n+1}$. The map defined in (2), called *amplitude encoding*, encodes $\mathbf{x}$ into the pure state $\rho_{\mathbf{x}} = |\mathbf{x}\rangle\langle\mathbf{x}|$ where the additional component of $|\mathbf{x}\rangle$ stores the norm of $\mathbf{x}$. Nevertheless the quantum encoding $\mathbf{x} \mapsto \rho_{\mathbf{x}}$ can be realized in terms of the Bloch vectors $\mathbf{x} \mapsto \mathbf{b}^{(\rho_{\mathbf{x}})}$ saving space resources. The improvement of memory occupation within the Bloch representation is evident when we take multiple tensor products $\rho \otimes \cdots \otimes \rho$ of a density matrix ρ constructing a feature map to enlarge the dimension of the representation space [1].

Quantum-inspired classification methods rely on three main steps: encoding data vectors into quantum density matrices, calculating centroids within this quantum representation, and applying various quantum state distinguishability criteria such as Helstrom discrimination, the Pretty-Good measurement [2], and the geometric minimum-error measurement [3] to differentiate between classes.

Let us briefly recall the notion of quantum state discrimination. Given a set of arbitrary quantum states with respective a priori probabilities $R = \{(\rho_1, p_1), ..., (\rho_N, p_N)\}$, in general there is no a measurement process that discriminates the states without errors, i.e. a collection $E = \{E_i\}_{i=1,...,N}$ of positive semidefinite operators such that $\sum_{i=1}^N E_i = \mathtt{I}$, satisfying the following property: $\mathrm{tr}(E_i \rho_j) = 0$ when $i \neq j$ for all $i, j = 1, ..., N$. The probability of a successful state discrimination of the states in R performing the measurement E is:

$$\mathbb{P}_E(R) = \sum_{i=1}^N p_i \mathrm{tr}(E_i \rho_i). \tag{3}$$

A complete characterization of the optimal measurement E_{opt} that maximizes the probability (3) for $R = \{(\rho_1, p_1), (\rho_2, p_2)\}$ is due to Helstrom [4]. Let $\Lambda := p_1\rho_1 - p_2\rho_2$ be the *Helstrom observable* whose positive and negative eigenvalues are, respectively, collected in the sets D_+ and D_-. Consider the two orthogonal projectors:

$$P_\pm := \sum_{\lambda \in D_\pm} P_\lambda, \tag{4}$$

where P_λ projects onto the eigenspace of λ. The measurement $E_{opt} := \{P_+, P_-\}$ maximizes the probability (3) that attains the *Helstrom bound*:

$$h_b(\rho_1, \rho_2) = p_1 \mathrm{tr}(P_+ \rho_1) + p_2 \mathrm{tr}(P_- \rho_2). \tag{5}$$

Helstrom quantum state discrimination can be used to implement a quantum-inspired binary classifier with promising performances. Let $\{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_M, y_M)\}$ be a training set with $\mathbf{x}_i \in \mathbb{C}^n$, $y_i \in \{1, 2\}$ $\forall i = 1, ..., M$. Assume to encode the data points into quantum states by means of $\mathbb{C}^n \ni \mathbf{x} \mapsto \rho_{\mathbf{x}} \in \mathfrak{S}(\mathsf{H})$, one can construct the quantum centroids ρ_1 and ρ_2 of the two classes $C_{1,2} = \{\mathbf{x}_i : y_i = 1, 2\}$:

$$\rho_{1,2} = \frac{1}{|C_{1,2}|} \sum_{\mathbf{x} \in C_{1,2}} \rho_{\mathbf{x}} \tag{6}$$

Let $\{P_+, P_-\}$ be the Helstrom measurement defined by the set $R = \{(\rho_1, p_1), (\rho_2, p_2)\}$, where the probabilities attached to the centroids are $p_{1,2} = \frac{|C_{1,2}|}{|C_1| + |C_2|}$. The *Helstrom classifier* applies the optimal measurement for the discrimination of the two quantum centroids to assign the label y to a new data instance $\mathbf{x}$, encoded into the state $\rho_{\mathbf{x}}$, as follows:

$$y(\mathbf{x}) = \begin{cases} 1 & \text{if} \quad \mathrm{tr}(P_+ \rho_{\mathbf{x}}) \geq \mathrm{tr}(P_- \rho_{\mathbf{x}}) \\ 2 & \text{otherwise} \end{cases} \tag{7}$$

A strategy to increase the accuracy in classification is given by the construction of the tensor product of q copies of the quantum centroids $\rho_{1,2}^{\otimes q}$ enlarging the Hilbert space where data are encoded. The corresponding Helstrom measurement is $\{P_+^{\otimes q}, P_-^{\otimes q}\}$, and the Helstrom bound satisfies:

$$h_b(\rho_1^{\otimes q}, \rho_2^{\otimes q}) \leq h_b\left(\rho_1^{\otimes (q+1)}, \rho_2^{\otimes (q+1)}\right) \qquad \forall q \in \mathbb{N}. \tag{8}$$

A larger Hilbert space in quantum encoding yields a better Helstrom bound and consequently a more accurate classifier, though it typically increases computational cost. Notably, when dealing with real input vectors, encoding them into Bloch vectors provides a method to effectively increase the Hilbert space dimension while potentially reducing time and space complexity.

Clearly, defining a quantum encoding is equivalent to select a feature map to represent feature vectors into a space of higher dimension. In the case of the considered quantum amplitude encoding $\mathbb{R}^2 \ni (x_1, x_2) \mapsto \rho_{(x_1,x_2)} \in \mathfrak{S}(\mathbb{C}^3)$, the nonlinear explicit injective function $\varphi : \mathbb{R}^2 \to \mathbb{R}^5$ to encode data into Bloch vectors can be defined as follows:

$$\varphi(x_1, x_2) := \frac{1}{x_1^2 + x_2^2 + 1} \left(2x_1x_2, 2x_1, 2x_2, x_1^2 - x_2^2, \frac{x_1^2 + x_2^2 - 2}{\sqrt{3}}\right). \tag{9}$$

From a geometric point of view, the mapped feature vectors are points on the surface of a hyper-hemisphere. Class centroids, obtained by averaging feature vectors, lie within the hypersphere and, while not directly corresponding to density operators, can be rescaled to Bloch vectors. To boost classification accuracy, the dimension of the representation space can be increased by using q copies of quantum states in a tensor product, encoding data and centroids as $\rho^{\otimes q}$. Bloch encoding provides an efficient way to handle feature maps by offering an injective data encoding function that discards null and repeated elements from the Bloch vector. This drastically reduces storage requirements. Therefore, the Bloch representation allows for compact storage of the redundant information within $\rho^{\otimes q}$.

Let us consider a training set divided into the classes $C_1,...,C_M$, assume we have any training point $\mathbf{x}$ encoded into the Bloch vector $\mathbf{b}^{(\mathbf{x})}$ of a pure state on $\mathbb{C}^d$. The calculation of the centroid of the class C_i, within this quantum encoding, must take into account that the mean of the Bloch vectors $\mathbf{b}^{(i)} := \frac{1}{|C_i|}\sum_{\mathbf{x}\in C_i}\mathbf{b}^{(\mathbf{x})}$ does not represent a density operator in general. In fact, for $d > 2$ the points contained in the unit hypersphere of $\mathbb{R}^{d^2-1}$ are not in bijective correspondence with density matrices on $\mathbb{C}^d$. However, since any vector within the closed ball of radius $\frac{2}{d}$ gives rise to a density operator, a centroid can be defined in terms of a meaningful Bloch vector by a rescaling:

$$\widehat{\mathbf{b}}^{(i)} := \frac{2}{d|C_i|}\sum_{\mathbf{x}\in C_i}\mathbf{b}^{(\mathbf{x})}. \tag{10}$$

A method of quantum state discrimination for distinguishing more than two states $\{(\rho_1,p_1),...,(\rho_N,p_N)\}$ is the square-root measurement, also known as *Pretty-Good measurement*, defined by:

$$E_i = p_i\rho^{-\frac{1}{2}}\rho_i\rho^{-\frac{1}{2}}, \tag{11}$$

where $\rho = \sum_i p_i\rho_i$, PGM is the optimal minimum-error when states satisfy certain symmetry properties [2]. Clearly to distinguish between n centroids we need a measurement with at most n outcomes. It is sometimes optimal to avoid measurement and simply guess that the state is the a priori most likely state.

The optimal POVM $\{E_i\}_i$ for minimum-error state discrimination over

$$R = \{(\rho_1,p_1),...,(\rho_N,p_N)\}$$

satisfies the following necessary and sufficient Helstrom conditions [3]:

$$\Gamma - p_i\rho_i \geq 0 \qquad \forall i = 1,\dots,N, \tag{12}$$

where the Hermitian operator, also known as *Lagrange operator*, is defined by $\Gamma := \sum_i p_i\rho_i E_i$. It is also useful to consider the following properties which can be obtained from the above conditions:

$$E_j(p_j\rho_j - p_i\rho_i)E_i = 0 \quad \forall i,j. \tag{13}$$

For each i the operator $\Gamma - p_i\rho_i$ can have two, one, or no zero eigenvalues, corresponding to the zero operator, a rank-one operator, and a positive-definite operator, respectively. In the first case, we use the measurement $\{E_i = \mathtt{I}, E_{i\neq j} = 0\}$ for some i where $p_i \geq p_j\ \forall j$, i.e. the state belongs to the a priori most likely class. In the second case, if $E_i \neq 0$, it is a weighted projector onto the corresponding eigenstate. In the latter case, it follows that $E_i = 0$ for every optimal measurement.

Given the following Bloch representations:

$$\Gamma = \frac{1}{d}\Big(a\,\mathtt{I}_d + \sqrt{\frac{d(d-1)}{2}}\sum_{j=1}^{d^2-1} b_j\sigma_j\Big) \tag{14}$$

$$\rho_i = \frac{1}{d}\Big(\mathtt{I}_d + \sqrt{\frac{d(d-1)}{2}}\sum_{j=1}^{d^2-1} b_j^{(i)}\sigma_j\Big), \tag{15}$$

in order to determine the Lagrange operator in $\mathbb{C}^d$ we need d^2 independent linear constraints:

$$2p_i\Big(a - \widehat{\mathbf{b}}^{(i)}\cdot\mathbf{b} - \frac{p_i}{2}(1-|\widehat{\mathbf{b}}^{(i)}|^2)\Big) = a^2 - |\mathbf{b}|^2. \tag{16}$$

A measurement with more than d^2 outcomes can always be decomposed as a probabilistic mixture of measurements with at most d^2 outcomes. Therefore, if the number of classes is greater than or equal to d^2 and we get d^2 linearly independent equations, we construct the Lagrange operator and derive the optimal measurements. From the geometric point of view, we obtain the unit vectors corresponding to the rank-1 projectors $E_i = \frac{1}{d}\Big(\mathtt{I}_d + \sqrt{\frac{d(d-1)}{2}}\sum_{j=1}^{d^2-1} n_j^{(i)}\sigma_j\Big)$ where $\mathbf{n}^{(i)} = \frac{\widehat{\mathbf{b}}^{(i)} - a\mathbf{b}}{|\widehat{\mathbf{b}}^{(i)} - a\mathbf{b}|} \in \mathbb{R}^{d^2-1}$ giving the POVM of the measurement. It is also possible to further partition the classes in order to increase the number of centroids and of the corresponding equations. The classification is carried out in this way: an unlabeled point $\widehat{\mathbf{x}}$ is associated with the first label y such that $\mathbf{b}^{(\widehat{\mathbf{x}})}\cdot\mathbf{n}^{(y)} = \max_i \mathbf{b}^{(\widehat{\mathbf{x}})}\cdot\mathbf{n}^{(i)}$.

3 LOCAL QUANTUM-INSPIRED CLASSIFIERS

Our implementation of the quantum state discrimination classifiers begins by employing a k-nearest neighbors (kNN) approach to select the k closest training samples to the unclassified data point. The kNN algorithm itself is a simple classifier that operates through these steps:

1. calculating the distance between the test sample and all training samples using a chosen metric;

2. identifying the k training samples with the smallest distances;

3. assigning the class label based on the majority class among these k neighbors.

We proceed by first using kNN to extract the closest elements to a test instance, followed by a quantum-inspired classification instead of majority voting. We investigate two scenarios: either applying kNN in the original input space (e.g., via Euclidean distance) and then encoding the k neighbors for quantum classification, or encoding the entire dataset into density matrices and then using kNN with a quantum operator distance to find the k neighbors. In this latter case, the distance metric we employ is the *Bures distance*, a quantum generalization of the Fisher information and a distance linked to super-fidelity. The Bures distance is defined by:

$$d_B(\rho_1,\rho_2)=\sqrt{2\left(1-\sqrt{\mathscr{F}(\rho_1,\rho_2)}\right)}, \qquad (17)$$

where the fidelity between density operators is given by $\mathscr{F}(\rho_1,\rho_2)=\left(\mathrm{tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right)^2$. Let us note that the fidelity reduces to $\mathscr{F}(\rho_1,\rho_2)=\langle\psi_1|\rho_2|\psi_1\rangle$ when $\rho_1=|\psi_1\rangle\langle\psi_1|$. Therefore the Bures distance between the pure state ρ_1 and the arbitrary state ρ_2 can be expressed in term of the Bloch representation as follows:

$$d_B(\rho_1,\rho_2)=\sqrt{2\left(1-\sqrt{\frac{1}{d}\left(1+(d-1)\mathbf{b}^{(1)}\cdot\mathbf{b}^{(2)}\right)}\right)} \qquad (18)$$

where $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$ are the Bloch vectors of ρ_1 and ρ_2 respectively and d is the dimension of the Hilbert space of the quantum encoding. The special form of the Bures distance, expressed in terms of Bloch vectors as in (18), is relevant for our purpose because data vectors are encoded into pure states and the quantum centroids are calculated as Bloch vectors of mixed states in general.

An alternative distance can be defined via super-fidelity

$$d_G(\rho_1,\rho_2)=\sqrt{1-\mathscr{G}(\rho_1,\rho_2)}, \qquad (19)$$

where the super-fidelity between density operators is given by

$$\mathscr{G}(\rho_1,\rho_2)=\mathrm{tr}\rho_1\rho_2+\sqrt{(1-\mathrm{tr}\rho_1^2)(1-\mathrm{tr}\rho_2^2)}.$$

Notice that the super-fidelity reduces to $\mathscr{G}(\rho_1,\rho_2)=\langle\psi_1|\rho_2|\psi_1\rangle$ when $\rho_1=|\psi_1\rangle\langle\psi_1|$. The inner distance between the corresponding Bloch vectors represents the angle θ between the unit vectors $(\mathbf{b}^{(1)},\sqrt{1-|\mathbf{b}^{(1)}|^2})$ and $(\mathbf{b}^{(2)},\sqrt{1-|\mathbf{b}^{(2)}|^2})$, which is normalized to be 1: $\hat{D}_G\left(\mathbf{b}^{(1)},\mathbf{b}^{(2)}\right)=\frac{\arccos\left(\mathbf{b}^{(1)}\cdot\mathbf{b}^{(2)}+\sqrt{(1-|\mathbf{b}^{(1)}|^2)(1-|\mathbf{b}^{(2)}|^2)}\right)}{\pi}$. For pure states the inner distance corresponds to the Fubini-Study distance.

In Algorithm 1, the locality is imposed by running the kNN on the input space finding the training vectors that are closest to the test element, then there is the quantum encoding into pure states and a quantum-inspired classifier (Helstrom, PGM, geometric Helstrom) is locally executed over the restricted training set. In Algorithm 2, the test element and all the training elements are encoded into Bloch vectors of pure states then a kNN is run w.r.t. the Bures distance to find the nearest neighbors in the space of the quantum representation, then a quantum-inspired classifier is executed with the training instances corresponding to the closest quantum states.

Algorithm 1 *Local quantum-inspired classification based on kNN in the input space before the quantum encoding. The distance can be: Euclidean, Manhattan, Chessboard, Canberra, Bray-Curtis.*

Require: Dataset X of labeled instances, unlabeled point $\hat{\mathbf{x}}$
Ensure: Label of $\hat{\mathbf{x}}$
find the k nearest neighbors $\mathbf{x}_1,...,\mathbf{x}_k$ to $\hat{\mathbf{x}}$ in X w.r.t. the Euclidean distance
encode $\hat{\mathbf{x}}$ into a pure state $\rho_{\hat{\mathbf{x}}}$
for $j=1,...,k$ **do**
  encode $\mathbf{x}_j$ into a pure state $\rho_{\mathbf{x}_j}$
end for
run the quantum-inspired classifier with training points encoded into $\{\rho_{\mathbf{x}_j}\}_{j=1,...,k}$.

Algorithm 2 *Local quantum-inspired classification based on kNN in the Bloch representation after the quantum encoding. The distance can be: Bures, Super-Fidelity, Inner.*

Require: Dataset X of labeled instances, unlabeled point $\hat{\mathbf{x}}$
Ensure: Label of $\hat{\mathbf{x}}$
encode $\hat{\mathbf{x}}$ into a Bloch vector $\mathbf{b}^{(\hat{\mathbf{x}})}$ of a pure state
for $\mathbf{x}\in X$ **do**
  encode $\mathbf{x}$ into a Bloch vector $\mathbf{b}^{(\mathbf{x})}$ of a pure state
end for
find the k nearest neighbors to $\mathbf{b}^{(\hat{\mathbf{x}})}$ in $\{\mathbf{b}^{(\mathbf{x})}\}_{\mathbf{x}\in X}$ w.r.t. the distance D_B
run the quantum-inspired classifier over the k nearest neighbors.

A local quantum-inspired classifier can be defined without quantum state discrimination but considering a *nearest mean classification* like the following: after the quantum encoding we perform a kNN selection and calculate the centroid of each class considering only the nearest neighbors to the test element, finally we assign the label according to the nearest centroid as schematized in Algorithm 3.

Algorithm 3 *Local quantum-inspired nearest mean classifier.*

Require: Training set X divided into n classes C_i, unlabeled point $\hat{\mathbf{x}}$
Ensure: Label of $\hat{\mathbf{x}}$
encode $\hat{\mathbf{x}}$ into a Bloch vector $\mathbf{b}^{(\hat{\mathbf{x}})}$ of a pure state
for $\mathbf{x} \in X$ **do**
encode $\mathbf{x}$ into a Bloch vector $\mathbf{b}^{(\mathbf{x})}$ of a pure state
end for
find the neighborhood $K = \{\mathbf{b}^{(\mathbf{x}_1)}, ..., \mathbf{b}^{(\mathbf{x}_k)}\}$ of $\mathbf{b}^{(\hat{\mathbf{x}})}$ w.r.t. the distance D_B
for $i = 1, ..., n$ **do**
construct the centroid $\widehat{\mathbf{b}}^{(i)} = \frac{2}{d|C_i^k|} \sum_{\mathbf{x} \in C_i^k} \mathbf{b}^{(\mathbf{x})}$ where $C_i^k := \{\mathbf{x} \in C_i : \mathbf{b}^{(\mathbf{x})} \in K\}$
end for
find the closest centroid $\widehat{\mathbf{b}}^{(l)}$ to $\frac{2}{d}\mathbf{b}^{(\hat{\mathbf{x}})}$ w.r.t. the distance D_B
return label of the class C_l

4 RESULTS AND DISCUSSION

In this section, we describe some results obtained by the implementation of the local quantum-inspired classifiers with several distances compared to well-known classical algorithms. In particular, we consider the SVM with different kernels: linear, radial basis function, and sigmoid. Then, we run a random forest, a naive Bayes classifier, and the logistic regression. In order to compare the results with previous papers, we take into account the following benchmark datasets from PMLB public repository [5]. For each dataset we randomly select 80% of the data to create a training set and use the residual 20% for the evaluation. We repeated the same procedure 10 times and calculated the average accuracy using the code available at github.com/leporini/classification. Certainly, it is possible to compare the performances based on different statistic indices including Matthews correlation coefficient, F-measure, Cohen's parameter.

We observe that the performances of the local quantum-inspired classifiers turn out to be definitely more accurate, where the hyperparameter k is set equal to the number of classes in the dataset. This value is reasonable to construct the centroids of the classes. In particular, Algorithm 1 with the Euclidean distance is the most accurate classifier for the datasets *analcatdata_boxing1, analcatdata_happiness, biomed, prnn_fglass, wine_recognition*, while with Manhattan distance is best for *analcatdata_aids, analcatdata_japansolvent, breast_cancer, iris, tae*, with Chessboard distance is best for *analcatdata_cyyoung9302, analcatdata_lawsuit*, and with Bray-Curtis distance is best for *analcatdata_bankruptcy, appendicitis*. Algorithm 2 with the Bures distance outperforms Algorithm 1 and 3 for *analcatdata_dmft* and produces the same accuracy for *labor*. Algorithm 3 with the Bures distance is the most accurate classifier for *analcatdata_asbestos, new_thyroid, phoneme, prnn_synth*.

5 CONCLUSIONS

This paper centers on the practical implementation of classification algorithms that rely on quantum state discrimination. A key innovation is the introduction of a local approach for executing the classifier. Specifically, after partitioning the training set, the k nearest data points to the test element are encoded into Bloch vectors and subsequently used to determine the quantum centroid for each class.

The proposed methodology introduces a family of classifiers due to the flexibility in choosing both the strategy for defining locality within the training set and the quantum state discrimination procedure. Both the local classification approach and the quantum-inspired data encoding and processing warrant further exploration to fully understand their impact on machine learning.

In a forthcoming article, we will provide a formal complexity analysis of the algorithms with respect to dataset size, number of features, and Hilbert space dimensionality. We will also compare performance against more advanced classical methods, such as deep neural networks, to better position the benefits of quantum-inspired approaches. Furthermore, we will present a thorough error analysis and sensitivity study. This will specifically focus on how the hyperparameter k in our local strategy influences the results.

6 REFERENCES

[1] Leporini, R.; Pastorello, D. An efficient geometric approach to quantum-inspired classifications. *Sci. Rep.* **2022**, *12*, 8781. https://doi.org/10.1038/s41598-022-12392-1.

[2] Mochon, C. Family of generalized pretty good measurements and the minimal-error pure-state discrimination problems for which they are optimal. *Phys. Rev. A* **2006**, *73*, 032328. https://doi.org/10.1103/PhysRevA.73.032328.

[3] Bae, J. Structure of minimum-error quantum state discrimination. *New J. Phys.* **2013**, *15*, 073037. https://doi.org/10.1088/1367-2630/15/7/073037.

[4] Helstrom, C.W. Quantum detection and estimation theory. *J. Stat. Phys.* **1969**, *1*, 231–252.

[5] Romano, J.D.; Le, T.T.; La Cava, W.; Gregg, J.T.; Goldberg, D.J.; Chakraborty, P.; Ray, N.L.; Himmelstein, D.; Fu, W.; Moore, J.H., PMLB v1.0: an open source dataset collection for benchmarking machine learning methods, *arXiv* **2020**, arXiv:2012.00058.