International Journal of Theoretical Physics (2024) 63:55
https://doi.org/10.1007/s10773-024-05566-2

RESEARCH

A Quantum-inspired Approach to Pattern Recognition and Machine Learning. Part I

Maria Luisa Dalla Chiara[1] · Roberto Giuntini[2,3] · Giuseppe Sergioli[3]

Received: 22 March 2023 / Accepted: 25 January 2024 / Published online: 21 February 2024

Abstract

How are abstract concepts formed and recognized on the basis of a previous experience? It is interesting to compare the behavior of human minds and of artificial intelligences with respect to this problem. Generally, a human mind that abstracts a concept (say, *table*), from a given set of known examples creates a *table-Gestalt*: a kind of vague and out of focus image that does not fully correspond to a particular table with well determined features. Can the construction of a *gestaltic pattern* (that is so natural for human minds) be taught to an intelligent machine? This problem can be successfully discussed in the framework of a quantum-inspired approach to pattern recognition and to machine learning. The basic idea is replacing *classical datasets* with *quantum datasets* where objects are described by special quantum states, involving the characteristic uncertainty and ambiguity of the quantum theoretic formalism. In this framework, the intuitive concept of *Gestalt* can be simulated by the mathematical concept of *positive centroid of a given quantum dataset*. Accordingly, the crucial problem "how can we classify a new object on the basis of a previous experience?" can be dealt with in terms of some special *quantum similarity-relations* that may hold between the new object's state and the positive centroid of the quantum dataset under consideration. This allows us to define a particular quantum classifier, called *fidelity-classifier*, that admits the possibility of *uncertain answers*. Although recognition procedures are different for human and for artificial intelligences, there is a common method of "facing the problems" that seems to work in both cases.

Keywords Recognitions of concepts · Gestalt · Fidelity-classifiers

✉ Giuseppe Sergioli
giuseppe.sergioli@gmail.com

Maria Luisa Dalla Chiara
marisa.dallachiara@gmail.com

Roberto Giuntini
giuntini@unica.it

1 Dipartimento di Lettere e Filosofia, Universitá di Firenze, Via della Pergola 60, Firenze I-50121, Italy

2 TUM Institute for Advanced Study, Lichtenbergstr. 2a, Garching bei München 85748, Germany

3 Dipartimento di Pedagogia, Psicologia, Filosofia, Universitá di Cagliari, Via Is Mirrionis 1, Cagliari I-09123, Italy

1 Concept-recognitions for Human and for Artificial Intelligences

How are abstract concepts formed and recognized on the basis of some previous experience? This difficult question has been investigated with different methods by psychologists, neuroscientists, artificial intelligence researchers, logicians, philosophers. Interestingly enoug, neuroscientific researches have recently interacted with an important approach to psychology: the *Gestalt theory*, that had been proposed by Wertheimer, Kofka and Köhler in the early 20th century.[1] The basic idea of this theory can be sketched as follows: human perception and knowledge of objects is essentially based on our capacity of realizing a *Gestalt* (a *form*) of the objects in question: a *holistic image* that cannot be identified with the set of its component elements. A human mind that abstracts the concept *table* from a given set of concrete examples, generally creates a *table-Gestalt*, a kind of *vague and out of focus image* that does not fully correspond to a particular table with well determined features. When we *recognize* as a *table* a new object we have met in our environment, we generally make a *comparison* between

- the main *features* of the new object;
- the *table-Gestalt* that we had constructed in our mind.

Can such recognition-processes that are so natural for human minds be "taught" to an intelligent machine? Is it possible to *simulate* the intuitive notion of *Gestalt* by some adequate mathematical concepts? This question can be successfully investigated in the framework of a *quantum-inspired approach to pattern recognition and to machine learning*. Unlike some standard quantum approaches whose aim is designing quantum circuits to implement machine-learning processes by means of quantum computers, quantum-inspired approches to pattern recognition and to machine learning are theoretic studies that apply quantum-information concepts in order to investigate recognition and classification-questions arising in different fields of knowledge.[2]

Consider an *agent* (let us call her $Alice$) who is interested in a given *concept* $\mathcal{C}$ that may refer either to *concrete* or to *abstract* objects (say *table, triangle, beautiful*). The name $Alice$ may denote either a *human* or an *artificial intelligence*. We will use $Alice_H$ for a *Human Mind* and $Alice_M$ for an *Intelligent Machine*. $Alice$ will then correspond either to $Alice_H$ or to $Alice_M$. We suppose that $Alice$ (on the basis of her previous experience) has already *recognized* and *classified* a given set of objects for which the question "does the object under consideration verify the concept $\mathcal{C}$?" can be reasonably asked. We assume that the possible answers to this question are:

- *YES!*
- *NO!*
- *PERHAPS!*

As an example, $Alice$ might be a child who has already recognized (in the environment where she is living) the objects that are tables and the objects that are not tables. At the same time, she might have been doubtful about the *right* classification of some particular objects. For instance, she might have answered "PERHAPS!" to the question "is this food trolley a table?".

While $Alice_H$ may have *seen* the objects under consideration, *seeing* is of course more problematic for $Alice_M$. Thus, generally, one shall make recourse to some *theoretic representations* that faithfully describe the objects in question. As happens in physics, such theoretic

[1] See, for instance, [2].

[2] See [6].

representations can be identified with convenient mathematical objects that represent *object-states*.

In the classical approaches to pattern recognition and machine learning an *object-state* is usually represented as a vector

$$\overrightarrow{\mathbf{x}} = (x_1, \ldots, x_d),$$

that belongs to the real space $\mathbb{R}^d$ (where $d \geq 1$). Every component x_i of the vector $\overrightarrow{\mathbf{x}}$ is supposed to correspond to a possible value of an *observable* that is considered relevant for recognizing the concept $\mathcal{C}$. The number x_i is called a *feature* of the object represented by the vector $\overrightarrow{\mathbf{x}}$.

We will first discuss the problem "how is a concept $\mathcal{C}$ recognized on the basis of a previous experience?" in a classical framework. Suppose that (at a given time t_0) $Alice$ is interested in the concept $\mathcal{C}$. Her previous experience concerning $\mathcal{C}$ can be described by the formal notion of *classical three-valued* $\mathcal{C}$*-dataset*.[3]

Definition 1.1 Classical three-valued $\mathcal{C}$-dataset

A classical three-valued $\mathcal{C}$-dataset *(briefly,* classical 3$\mathcal{C}$-dataset*) is a sequence*

$$^{\mathcal{C}}CDS = \ (\mathbb{R}^d, \ ^{\mathcal{C}}St, \ ^{\mathcal{C}}St^+, \ ^{\mathcal{C}}St^-, \ ^{\mathcal{C}}St^? \ ),$$

where:

1. $^{\mathcal{C}}St$ *is a finite set of object-states* $\overrightarrow{\mathbf{x}}$ *in the space* $\mathbb{R}^d$*, for which the question "does the object described by* $\overrightarrow{\mathbf{x}}$ *verify the concept* $\mathcal{C}$*?" can be reasonably asked.*
2. $^{\mathcal{C}}St^+$ *is a subset of* $^{\mathcal{C}}St$*, consisting of all states that have been* positively classified *with respect to the concept* $\mathcal{C}$*. The elements of this set are called the* positive instances *of the concept* $\mathcal{C}$*.*
3. $^{\mathcal{C}}St^-$ *is a subset of* $^{\mathcal{C}}St$*, consisting of all states that have been* negatively classified *with respect to the concept* $\mathcal{C}$*. The elements of this set are called the* negative instances *of the concept* $\mathcal{C}$*.*
4. $^{\mathcal{C}}St^?$ *is a (possibly empty) subset of* $^{\mathcal{C}}St$*, consisting of all states that have been considered* problematic *with respect to* $\mathcal{C}$*. The elements of this set are called the* indeterminate instances *of the concept* $\mathcal{C}$*.*
5. *The three sets* $^{\mathcal{C}}St^+$, $^{\mathcal{C}}St^-$, $^{\mathcal{C}}St^?$ *are pairwise disjoint. Furthermore,* $^{\mathcal{C}}St^+ \cup {}^{\mathcal{C}}St^- \cup {}^{\mathcal{C}}St^? = {}^{\mathcal{C}}St$.

We indicate by n, n^+, n^-, $n^?$ *the cardinal numbers of the sets* $^{\mathcal{C}}St$, $^{\mathcal{C}}St^+$, $^{\mathcal{C}}St^-$, $^{\mathcal{C}}St^?$*, respectively.*

Particular examples of classical datasets are the *binary classical datasets*, where the set $^{\mathcal{C}}St^?$ of the indeterminate instances is empty. The elements of the set $^{\mathcal{C}}St^+ \cup {}^{\mathcal{C}}St^-$ will be called the *determinate* instances of the dataset $^{\mathcal{C}}CDS$.

Suppose that at a later time (t_1) $Alice$ "meets" a new object described by the object-state $\overrightarrow{\mathbf{y}}$. She shall find a rule that allows her to answer the question "does $\overrightarrow{\mathbf{y}}$ verify the concept $\mathcal{C}$?". And this answer shall refer to her previous knowledge that is represented by the classical 3$\mathcal{C}$-dataset

$$^{\mathcal{C}}CDS = \ (\mathbb{R}^d, \ ^{\mathcal{C}}St, \ ^{\mathcal{C}}St^+, \ ^{\mathcal{C}}St^-, \ ^{\mathcal{C}}St^? \ ).$$

[3] Although in most classical approaches to machine learning information is supposed to be dichotomic (every answer to a given question should be either "YES!" or "NO!") in some cases considering situations where the answer "PERHAPS!" is admitted may be more interesting and "realistic".

A winning strategy is based on the use of two special concepts: the *(classical) positive centroid* and the *(classical) negative centroid* of a given classical $3\mathcal{C}$-dataset.

Definition 1.2 *Classical centroids*
Consider a classical $3\mathcal{C}$-dataset

$$^{\mathcal{C}}CDS = (\mathbb{R}^d, \ ^{\mathcal{C}}St, \ ^{\mathcal{C}}St^+, \ ^{\mathcal{C}}St^-, \ ^{\mathcal{C}}St^?).$$

(1) The positive centroid *of $^{\mathcal{C}}CDS$ is the following vector of the space $\mathbb{R}^d$:*

$$\overrightarrow{\mathbf{x}}^+ = \sum_i \left\{ \frac{1}{n^+} \overrightarrow{\mathbf{x}_i} : \overrightarrow{\mathbf{x}_i} \in \ ^{\mathcal{C}}St^+ \right\}.$$

(2) The negative centroid *of $^{\mathcal{C}}CDS$ is the following vector of the space $\mathbb{R}^d$:*

$$\overrightarrow{\mathbf{x}}^- = \sum_i \left\{ \frac{1}{n^-} \overrightarrow{\mathbf{x}_i} : \overrightarrow{\mathbf{x}_i} \in \ ^{\mathcal{C}}St^- \right\}.$$

From an intuitive point of view the positive (negative) centroid of $^{\mathcal{C}}CDS$ can be regarded as the description of an *imaginary object*, whose state is determined by calculating the average-value of each feature for all positive (negative) instances of $^{\mathcal{C}}CDS$.

In order to face the classification-problem we will now introduce a special class of *similarity-relations* that allow us to compare any object-state living in the space $\mathbb{R}^d$ (of a given dataset $^{\mathcal{C}}CDS$) with the positive and with the negative centroid of $^{\mathcal{C}}CDS$. These particular similarity-relations can be defined in terms of a function that will be called *classical fidelity*.

Let us first recall the definition of *Euclidean distance*.

Definition 1.3 *Euclidean distance*
The Euclidean distance *on a space $\mathbb{R}^d$ is the binary function that associates to any pair of vectors $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$ of the space the following real number:*

$$d(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) = \| \overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}} \|$$

(where $\| \overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}} \|$ is the length of $\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{y}}$).

The Euclidean distance can be transformed, in a canonical way, into a new binary function, whose values are real numbers in the interval [0, 1]. We will call this function the *classical fidelity*.

Definition 1.4 *Classical fidelity*
The classical fidelity *on a space $\mathbb{R}^d$ is the binary function that associates to any pair of vectors $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$ of the space the following real number:*

$$CF(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) = \frac{1}{1 + d(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}})}.$$

From an intuitive point if view, the number $CF(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}})$ can be interpreted as a measure of the *degree of closeness* between the vectors $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$.

Apparently, the classical fidelity-values decrease with the increasing of the Euclidean distance. Furthermore, for any vectors $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$ we have:

- $CF(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) = 1$ iff $\overrightarrow{\mathbf{x}} = \overrightarrow{\mathbf{y}}$.

- $CF(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) \neq 0$.

By using the concept of *classical fidelity* we can now define in any space $\mathbb{R}^d$ the relation of *r-similarity* (where r is any real number in the interval $[0, 1]$).

Definition 1.5 *r-similarity*
Let $\overrightarrow{\mathbf{x}}$ *and* $\overrightarrow{\mathbf{y}}$ *be object-states of a space* $\mathbf{R}^d$ *and let* $r \in [0, 1]$.
The state $\overrightarrow{\mathbf{x}}$ *is called* r-similar *to the state* $\overrightarrow{\mathbf{y}}$ *(briefly,* $\overrightarrow{\mathbf{x}} \not\perp_r \overrightarrow{\mathbf{y}}$ *) iff* $r \leq CF(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}})$.

One can easily check that this relation is reflexive, symmetric and generally non-transitive. Since $CF(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) \neq 0$, there are infinitely many r such that $\overrightarrow{\mathbf{x}} \not\perp_r \overrightarrow{\mathbf{y}}$. As we will see, this fact does not represent a shortcoming for our aims.

Given a dataset ${}^{\mathcal{C}}CDS$, it is useful to refer to a *threshold-value*

$$r^* \in (\frac{1}{2}, 1]$$

that is considered relevant for the dataset in question.

When $\overrightarrow{\mathbf{x}} \not\perp_{r^*} \overrightarrow{\mathbf{y}}$, we have:

$$CF(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) \geq r^*.$$

Thus, the *degree of closeness* between $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$ can be considered "sufficiently high". In other words, $\overrightarrow{\mathbf{x}}$ and $\overrightarrow{\mathbf{y}}$ are "sufficiently similar".

Now *Alice* has at her disposal the mathematical tools that allow her to face the *classification problem*. Suppose that *Alice*'s information about a concept $\mathcal{C}$ is the *classical* 3$\mathcal{C}$*-dataset*

$${}^{\mathcal{C}}CDS = \ (\mathbb{R}^d, \ {}^{\mathcal{C}}St, \ {}^{\mathcal{C}}St^+, \ {}^{\mathcal{C}}St^-, \ {}^{\mathcal{C}}St^? \ ),$$

whose *positive* and *negative centroids* are the object-states $\overrightarrow{\mathbf{x}}^+$ and $\overrightarrow{\mathbf{x}}^-$, respectively. Let r^* be a *threshold-value* (in the interval $(\frac{1}{2}, 1]$), which is considered relevant for ${}^{\mathcal{C}}CDS$. The main goal is defining a *classifier function*, that assigns to every state $\overrightarrow{\mathbf{y}}$ (which describes an object that *Alice* may meet) either the value $+$ (corresponding to the answer "YES!") or the value $-$ (corresponding to the answer "NO"!) or the value ? (corresponding to the answer "PERHAPS!").

Definition 1.6 *The three-valued classical fidelity-classifier*
The three-valued classical fidelity-classifier *(briefly* $3-CFC$*), determined by the classical dataset* ${}^{\mathcal{C}}CDS$ *and by the threshold-value* r^**, is the function* $\mathcal{C}l_{[{}^{\mathcal{C}}CDS, r^*]}$ *that satisfies the following condition for any object-state* $\overrightarrow{\mathbf{y}}$ *of the space* $\mathbb{R}^d$*:*

$$\mathcal{C}l_{[{}^{\mathcal{C}}CDS, r^*]}(\overrightarrow{\mathbf{y}}) = \begin{cases} +, \ \textit{if} \ \overrightarrow{\mathbf{y}} \not\perp_{r^*} \overrightarrow{\mathbf{x}}^+ \ \textit{and not} \ \overrightarrow{\mathbf{y}} \not\perp_{r^*} \overrightarrow{\mathbf{x}}^-; \\ \overrightarrow{\mathbf{y}} \not\perp_{r^*} \overrightarrow{\mathbf{x}}^- \ \textit{and not} \ \overrightarrow{\mathbf{y}} \not\perp_{r^*} \overrightarrow{\mathbf{y}}^+; \\ ?, \ \textit{otherwise.} \end{cases}$$

In other words, an object-state $\overrightarrow{\mathbf{y}}$ is classified

- as a positive instance, when it is "sufficiently similar" to the positive centroid and is not "sufficiently similar" to the negative centroid;
- as a negative instance, when it is "sufficiently similar" to the negative centroid and is not "sufficiently similar" to the positive centroid;
- as an indeterminate instance, otherwise.

Let us now turn to a quantum-inspired approach to pattern recognition and machine learning, which is based on the following idea: replacing *classical object-states* with *pieces of quantum information* (possible states of quantum systems that are storing the information in question). In this way, our mathematical description of objects acquires the peculiar *uncertainty* and *ambiguity* that characterize quantum states.[4]

In some situations it may be natural to start with a classical information represented by an *object-state* $\overrightarrow{\mathbf{x}}$. Then, the transition to a quantum pure state $|\psi\rangle$ can be realized by adopting an *encoding procedure*, that transforms our classical object-state into a quantum pure state:

$$\overrightarrow{\mathbf{x}} \ \mapsto \ |\psi\rangle_{\overrightarrow{\mathbf{x}}}.$$

Two important examples of a "natural" quantum encoding are the *amplitude encoding* and the *stereographic encoding*.

Definition 1.7 *Amplitude encoding*

Consider a classical object-state $\overrightarrow{\mathbf{x}} = (x_1, \ldots, x_d) \in \mathbb{R}^d$*. The* quantum-amplitude encoding *of* $\overrightarrow{\mathbf{x}}$ *is the following unit vector that lives in the space* $\mathbb{R}^{(d+1)}$*:*

$$AmpEnc(\overrightarrow{\mathbf{x}}) = \frac{(x_1, \ldots, x_d, 1)}{\|(x_1, \ldots, x_d, 1)\|}.$$

Definition 1.8 *Stereographic encoding*

Consider a classical object-state $\overrightarrow{\mathbf{x}} = (x_1, \ldots, x_d) \in \mathbb{R}^d$*. The* quantum stereographic encoding *of* $\overrightarrow{\mathbf{x}}$ *is the following unit vector that lives in the space* $\mathbb{R}^{(d+1)}$*:*

$$StEnc(\overrightarrow{\mathbf{x}}) = \frac{1}{\sum_{i=1}^{d}(x_i)^2 + 1}\left(2x_1, \ldots, 2x_d, \sum_{i=1}^{d}(x_i)^2 - 1\right).$$

Notice that, in both cases, the quantum encoding of a classical object state $\overrightarrow{\mathbf{x}}$ is a *quantum pure state* that preserves all *classical features* described by $\overrightarrow{\mathbf{x}}$.

Of course, one could also directly "reason" in a quantum-theoretic framewok, avoiding any reference to a (previously known) classical object-state $\overrightarrow{\mathbf{x}}$. In such a case, one can assume, right from the outset, that an *object-state* is represented by a quantum pure state $|\psi\rangle$ living in a given (finite-dimensional) Hilbert space $\mathcal{H}$. Such a state can be regarded as a *probabilistic answer* to a sequence of *quantum questions*:

$$(Q_1, \ldots, Q_q),$$

mathematically represented as projection operators of the space $\mathcal{H}$. The state $|\psi\rangle$ will assign to each question Q_i a probability-value according to the Born-rule:

$$Prob_{|\psi\rangle}(Q_i) = \mathrm{tr}(P_{|\psi\rangle} Q_i) \in [0, 1]$$

(where tr is the trace functional and $P_{|\psi\rangle}$ is the projection operator corresponding to the unit vector $|\psi\rangle$).

Now, by recalling the concept of *classical three-valued* $\mathcal{C}$*-dataset*, the concept of *quantum three-valued* $\mathcal{C}$*-dataset* can be defined in a natural way.

Definition 1.9 Quantum three-valued $\mathcal{C}$-dataset

A quantum three-valued $\mathcal{C}$-dataset *(briefly,* quantum $3\mathcal{C}$-dataset*) is a sequence*

$$^{\mathcal{C}}QDS = \ (^{\mathcal{C}}\mathcal{H}, \ ^{\mathcal{C}}St_q, \ ^{\mathcal{C}}St_q^{+}, \ ^{\mathcal{C}}St_q^{-}, \ ^{\mathcal{C}}St_q^{?} \ ),$$

[4] See, for instance, [3–5].

where:

1. ${}^{\mathcal{C}}\mathcal{H}$ *is a finite-dimensional Hilbert space associated to* $\mathcal{C}$.
2. ${}^{\mathcal{C}}St_q$ *is a finite set of pure states* $|\psi\rangle$ *of* ${}^{\mathcal{C}}\mathcal{H}$ *for which the question "does the object described by* $|\psi\rangle$ *verify the concept* $\mathcal{C}$*?" can be reasonably asked.*
3. ${}^{\mathcal{C}}St_q^+$ *is a subset of* ${}^{\mathcal{C}}St$*, consisting of all states that have been* positively classified *with respect to the concept* $\mathcal{C}$*. The elements of this set are called the* positive instances *of the concept* $\mathcal{C}$.
4. ${}^{\mathcal{C}}St_q^-$ *is a subset of* ${}^{\mathcal{C}}St$*, consisting of the* negative instances *of the concept* $\mathcal{C}$.
5. ${}^{\mathcal{C}}St_q^?$ *is a (possibly empty) subset of* ${}^{\mathcal{C}}St$*, consisting of the* indeterminate instances *of the concept* $\mathcal{C}$.
6. *The three sets* ${}^{\mathcal{C}}St_q^+$, ${}^{\mathcal{C}}St_q^-$, ${}^{\mathcal{C}}St_q^?$ *are pairwise disjoint. Furthermore,* ${}^{\mathcal{C}}St_q^+ \cup {}^{\mathcal{C}}St_q^- \cup {}^{\mathcal{C}}St_q^? = {}^{\mathcal{C}}St_q$.

As we have done in the classical case, we indicate by $n,\ n^+,\ n^-,\ n^?$ the cardinal numbers of the sets ${}^{\mathcal{C}}St_q$, ${}^{\mathcal{C}}St_q^+$, ${}^{\mathcal{C}}St_q^-$, ${}^{\mathcal{C}}St_q^?$, respectively.

The concepts of *quantum positive* and *quantum negative centroid* (of a given quantum dataset) can be now defined *mutatis mutandis* with respect to the classical case.

Definition 1.10 *Quantum centroids*

Consider a quantum $\mathcal{C}$*-dataset*

$${}^{\mathcal{C}}QDS = ({}^{\mathcal{C}}\mathcal{H},\ {}^{\mathcal{C}}St_q,\ {}^{\mathcal{C}}St_q^+,\ {}^{\mathcal{C}}St_q^-,\ {}^{\mathcal{C}}St_q^?\ ).$$

(1) The quantum positive centroid *of* ${}^{\mathcal{C}}QDS$ *is the following density operator of the space* ${}^{\mathcal{C}}\mathcal{H}$*:*

$$\rho^+ = \sum_i \left\{ \frac{1}{n^+} P_{|\psi_i\rangle} : |\psi_i\rangle \in {}^{\mathcal{C}}St^+ \right\}.$$

(2) The quantum negative centroid *of* ${}^{\mathcal{C}}QDS$ *is the following density operator of the space* ${}^{\mathcal{C}}\mathcal{H}$*:*

$$\rho^- = \sum_i \left\{ \frac{1}{n^-} P_{|\psi_i\rangle} : |\psi_i\rangle \in {}^{\mathcal{C}}St^- \right\}.$$

Notice that quantum centroids are density operators that do not generally correspond to pure states. Furthermore, a quantum centroid cannot be generally represented as the quantum encoding of a corresponding classical centroid.

The concept of *quantum positive centroid* seems to represent a "good" mathematical simulation for the intuitive idea of *Gestalt*. In fact, both the *quantum positive centroid* and the intuitive idea of *Gestalt* describe an *imaginary object*, representing a *vague, ambiguous idea* that $Alice$ has obtained as an abstraction from the "real" examples she has met in her previous experience. As happens in the case of the intuitive idea of *Gestalt*, a *quantum positive centroid*, represented by the density operator $\rho^+ = \sum_i \left\{ \frac{1}{n^+} P_{|\psi_i\rangle} : |\psi_i\rangle \in {}^{\mathcal{C}}St^+ \right\}$ *ambiguously alludes* to the concrete *positive instances* that $Alice$ had previously met.

We know that *human recognitions and classifications* are usually performed by means of a quick and mostly unconscious *comparison* between the main features of some *new objects* we have met and a *gestaltic pattern* that we had previously constructed in our mind. And any *comparison* generally involves the use of some similarity-relations that are mostly grasped in a vague and intuitive way by human intelligences. This typical human procedure can be formally represented in the framework of our quantum-inspired approach to pattern recognition.

As we have done in the classical case, we will first introduce a class of *similarity-relations*, that can be defined in terms of a quantum concept of *fidelity*.

Definition 1.11 *The (quantum) fidelity for pure states.*

The (quantum) fidelity *on a Hilbert space* $\mathcal{H}$ *is defined as the function* F *that assigns to any pair* $|\psi\rangle$ *and* $|\varphi\rangle$ *of pure states of* $\mathcal{H}$ *the real number*

$$F(|\psi\rangle, |\varphi\rangle) = |\langle \psi \mid \varphi \rangle|^2$$

(where $\langle \psi \mid \varphi \rangle$ *is the inner product of* $|\psi\rangle$ *and* $|\varphi\rangle$*).*

The definition of *fidelity* can be generalized to the case of density operators, which may represent either pure or mixed states.

Definition 1.12 *Fidelity for density operators*

Consider a Hilbert space $\mathcal{H}$*. The* fidelity *on* $\mathcal{H}$ *is the function* F *that assigns to any pair* ρ *and* σ *of density operators of* $\mathcal{H}$ *the real number*

$$F(\rho, \sigma) := \operatorname{tr}\left(\sqrt{\sqrt{\rho}\,\sigma\sqrt{\rho}}\right)^2.$$

This definition represents a "good" generalization of the concept of fidelity for pure states. For, one can show that:

$$F(P_{|\psi\rangle}, P_{|\varphi\rangle}) = |\langle \psi \mid \varphi \rangle|^2.$$

It is interesting to recall the main properties of the fidelity-function, which play an important role in many applications:

1. $F(\rho, \sigma) \in [0, 1]$.
2. $F(\rho, \sigma) = F(\sigma, \rho)$.
3. $F(\rho, \sigma) = 0$ iff $\rho\sigma$ is the null operator.
4. $F(\rho, \sigma) = 1$ iff $\rho = \sigma$.[5]

From a physical point of view, the fidelity-function can be regarded as a form of *symmetric conditional probability*: $F(\rho, \sigma)$ represents the probability that a quantum system in state ρ can be transformed into a system in state σ, and vice versa.

As happens in the classical case, the quantum concept of *fidelity* allows us to define in any Hilbert space $\mathcal{H}$ a class of similarity-relations, called r*-similarities*, where r is any real number in the interval $[0, 1]$.

Definition 1.13 *Quantum* r*-similarity*

Let ρ *and* σ *be two density operators of a Hilbert space* $\mathcal{H}$ *and let* $r \in [0, 1]$.

The state ρ *is called* r-similar *to the state* σ *(briefly,* $\rho \not\perp_r \sigma$*) iff* $r \leq F(\rho, \sigma)$.

Now $Alice$ has at her disposal the mathematical tools that allow her to face the *classification problem* in the quantum case. Suppose that $Alice$'s information about a concept $\mathcal{C}$ is the *quantum* $\mathcal{C}$*-dataset*

$$^{\mathcal{C}}QDS = \left(^{\mathcal{C}}\mathcal{H},\ ^{\mathcal{C}}St_q,\ ^{\mathcal{C}}St_q^{+},\ ^{\mathcal{C}}St_q^{-},\ ^{\mathcal{C}}St_q^{?}\right),$$

[5] Notice that property 4 ($F(\rho, \sigma) = 1$ iff $\rho = \sigma$) holds for pure states when they are dealt with as special cases of density operators (i.e. as projections over 1-dimensional closed subspaces). Pure states dealt with as unit vectors only satisfy the weaker condition: $|\psi\rangle = |\varphi\rangle \implies F(|\psi\rangle, |\varphi\rangle) = 1$.

whose *positive* and *negative centroids* are the states ρ^+ and ρ^-, respectively. Let r^* be a *threshold value* in the interval $(\frac{1}{2}, 1]$, that is considered relevant for ${}^{\mathcal{C}}QDS$. Like in the classical case, the main goal is defining a *classifier function*, that assigns to every state σ (which describes an object that $Alice$ may meet) either the value $+$ (corresponding to the answer "YES!") or the value $-$ (corresponding to the answer "NO"!) or the value ? (corresponding to the answer "PERHAPS!").

Definition 1.14 *The three-valued quantum fidelity-classifier*

The three-valued quantum fidelity-classifier *(briefly,* $3 - QFC$*), determined by the quantum dataset* ${}^{\mathcal{C}}QDS$ *and by the threshold-value* r^**, is the function* $\mathcal{C}l_{[{}^{\mathcal{C}}QDS, r^*]}$ *that satisfies the following condition for any state* σ *of the space* ${}^{\mathcal{C}}\mathcal{H}$*:*

$$\mathcal{C}l_{[{}^{\mathcal{C}}QDS, r^*]}(\sigma) = \begin{cases} +, & \text{if } \sigma \not\perp_{r^*} \rho^+ \text{ and not } \sigma \not\perp_{r^*} \rho^-; \\ -, & \text{if } \sigma \not\perp_{r^*} \rho^- \text{ and not } \sigma \not\perp_{r^*} \rho^+; \\ ?, & \text{otherwise.} \end{cases}$$

In other words, an object-state σ is classified

- as a positive instance, when it is "sufficiently similar" to the positive centroid and is not "sufficiently similar" to the negative centroid;
- as a negative instance, when it is "sufficiently similar" to the negative centroid and is not "sufficiently similar" to the positive centroid;
- as an indeterminate instance, otherwise.

Although recognition procedures are different for human and for artificial intelligences, there is a common method of "facing the problems" that seems to work in both cases. Using quantum-theoretic concepts represents a great advantage in order to investigate the relationships between the behaviors of human and of artificial intelligences. The intuitive concept of *Gestalt* can hardly been simulated in a classical framework; for, the characteristic *ambiguity* of a *quantum positive centroid* is not shared by the corresponding notion of *classical positive centroid*. As we have seen, in the classical case a *positive centroid* represents an *exact object-state*, that is obtained by calculating the average-values of the values that all *positive instances* assign to the observables under consideration. Thus, unlike the quantum case, a *classical positive centroid* describes an *imaginary object* that is characterized by *precise features*.

2 An Empirical Simulation

We will now illustrate a simple *empirical simulation*, based on the fidelity-classifier, both in the classical and in the quantum case. We suppose that $Alice$ is interested in a concept $\mathcal{C}$ that describes a kind of flower (say, the *rose*). At the initial time (t_0) she has a classical three-valued information concerning a given set of instances of different flowers. Every flower is supposed to be characterized by two *features* that concern the *petal length* and the *petal width*, respectively. Thus, the classical object-state that describes a particular flower in the set of instances under consideration will be a vector $\vec{x} = (x_i, x_j)$ of the real space $\mathbb{R}^2$.

We suppose that the numbers occurring in our empirical simulation are the following:
n^+ (the number of the positive instances) = 352;
n^- (the number of the negative instances) = 359;
$n^?$ (the number of the indeterminate instances) = 339.

Hence, $n = n^{+} + n^{-} + n^{?} = 1050$.

Accordingly, $Alice$' s classical information at the initial time is illustrated by Fig. 1, where the red points correspond to the *positive instances*, while the blue points and the green points correspond to the *negative instances* and to the *indeterminate instances*, respectively.

This information can be represented as a particular three-valued classical dataset, having the following form:

$$^{\mathcal{C}}CDS = \;(\mathbb{R}^2,\; {}^{\mathcal{C}}St,\; {}^{\mathcal{C}}St^{+},\; {}^{\mathcal{C}}St^{-},\; {}^{\mathcal{C}}St^{?}\;).$$

An interesting parameter is represented by the *indeterminacy rate* of the dataset ${}^{\mathcal{C}}CDS$, which is defined as follows:

$$IR({}^{\mathcal{C}}CDS) := \;\frac{n^{?}}{n}.$$

From an intuitive point of view, the number $IR({}^{\mathcal{C}}CDS)$ measures the *degree of uncertainty* of $Alice$'s information. In the case of our example we have:

$$IR({}^{\mathcal{C}}CDS) = 0.323.$$

An important question arises: can $Alice$ control the *reliabilty* of her *fidelity-classifier* $Cl_{[{}^{\mathcal{C}}CDS, r^{*}]}$, which is based on the dataset ${}^{\mathcal{C}}CDS$ and on the choice of a threshold value r^{*}?

In order to answer this question $Alice$ can apply the standard *supervised procedure*. First of all she randomly splits the set ${}^{\mathcal{C}}St$ of all instances of her dataset ${}^{\mathcal{C}}CDS$ into two proper subsets:

- the *training set* ${}^{\mathcal{C}}St_{Train}$;

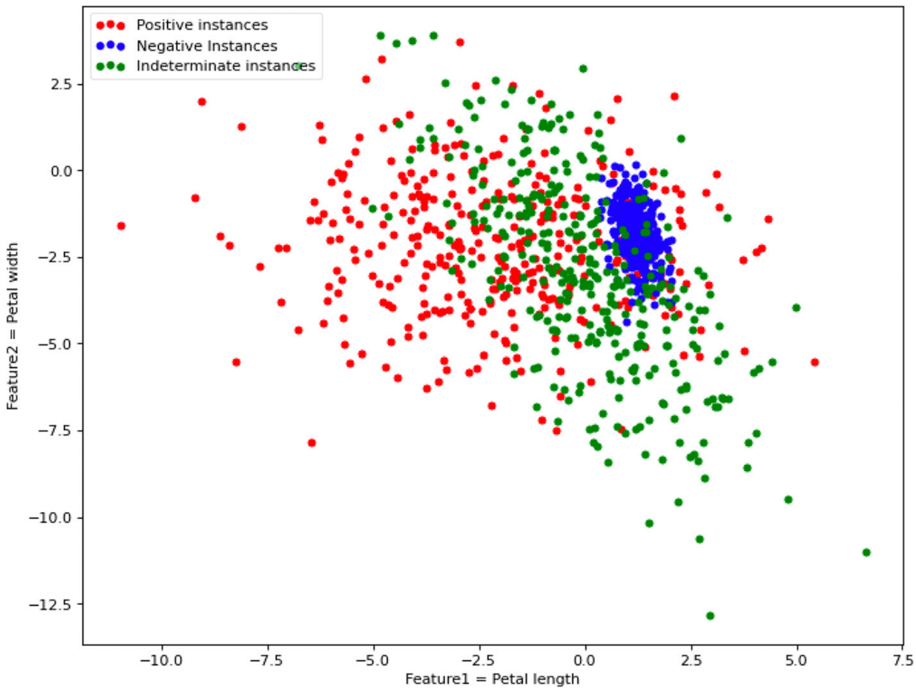

Fig. 1 $Alice$'s classical information at the initial time

- the *test set* ${}^{\mathcal{C}}St_{Test}$.

We assume that the training set ${}^{\mathcal{C}}St_{Train}$ represents the 80% of the original set ${}^{\mathcal{C}}St$, while the test set ${}^{\mathcal{C}}St_{Test}$ corresponds to the 20% . This gives rise to two new ("smaller") datasets, called the *training dataset* and the *test dataset*, that will be indicated as follows:

- ${}^{\mathcal{C}}CDS_{Train} = (\mathbb{R}^2,\ {}^{\mathcal{C}}St_{Train},\ {}^{\mathcal{C}}St^{+}_{Train},\ {}^{\mathcal{C}}St^{-}_{Train},\ {}^{\mathcal{C}}St^{?}_{Train}\ )$.
- ${}^{\mathcal{C}}CDS_{Test} = (\mathbb{R}^2,\ {}^{\mathcal{C}}St_{Test},\ {}^{\mathcal{C}}St^{+}_{Test},\ {}^{\mathcal{C}}St^{-}_{Test},\ {}^{\mathcal{C}}St^{?}_{Test}\ )$.

As expected, the training dataset ${}^{\mathcal{C}}CS_{Train}$ will have its own positive and negative centroids, called the *training positive centroid* ($\overrightarrow{\mathbf{x}}^{+}_{Train}$) and the *training negative centroid* ($\overrightarrow{\mathbf{x}}^{-}_{Train}$), respectively. Of course, generally, the two training centroids will be different from the centroids of the original dataset ${}^{\mathcal{C}}CDS$.

At a later time t_1, $Alice$ applies the fidelity-classifier $\mathcal{C}l_{[{}^{\mathcal{C}}CDS_{Train},r^*]}$ (based on the training dataset and on the threshold value r^*) to all *determinate* elements of the test dataset (i.e. to all instances that are either positive or negative). As a result, every input $\overrightarrow{\mathbf{y}} \in {}^{\mathcal{C}}St^{+}_{Test} \cup {}^{\mathcal{C}}St^{-}_{Test}$ will be classified either as a positive or as a negative or as an indeterminate instance.

By referring to this classification (performed by $\mathcal{C}l_{[{}^{\mathcal{C}}CDS_{Train},r^*]}$), we introduce the following terminology. We say that that an instance $\overrightarrow{\mathbf{y}}$ of the test dataset ${}^{\mathcal{C}}CDS_{Test}$ represents

- a *true positive instance* iff $\overrightarrow{\mathbf{y}}$ is a positive instance in the test dataset and has been classified as a positive instance (by the classifier $\mathcal{C}l_{[{}^{\mathcal{C}}CDS_{Train},r^*]}$);
- a *false positive instance* iff $\overrightarrow{\mathbf{y}}$ is a negative instance in the test dataset and has been classified as a positive instance;
- a *true negative instance* iff $\overrightarrow{\mathbf{y}}$ is a negative instance in the test dataset and has been classified as a negative instance;
- a *false negative instance* iff $\overrightarrow{\mathbf{y}}$ is a positive instance in the test dataset and has been classified as a negative instance;
- a *false indeterminate instance* iff $\overrightarrow{\mathbf{y}}$ is a determinate instance in the test dataset and has been classified as an indeterminate instance.

Consider now the following five numbers (which can be easily calculated in the case of our empirical simulation):

(1) the number TP of the true positive instances;
(2) the number TN of the true negative instances;
(3) the number FP of the false positive instances;
(4) the number FN of the false negative instances;
(5) the number FI of the false indeterminate instances.

On this basis, the *accuracy* of $Alice$'s classifier $\mathcal{C}l_{[{}^{\mathcal{C}}CDS,r^*]}$ can be defined as a function of these numbers:

$$Acc(\mathcal{C}l_{[{}^{\mathcal{C}}CDS,r^*]}) := \frac{TP + TN}{TP + TN + FP + FN + FI}.$$

In some situations it may be interesting to distinguish the *accuracy* of a given classifier from the *balanced accuracy*, that also depends on the cardinal number n^+ of the set of the positive instances and on the cardinal number n^- of the set of the negative instances of the dataset under consideration. The *balanced accuracy* of the classifier $\mathcal{C}l_{[{}^{\mathcal{C}}CDS,r^*]}$ is defined as follows:

$$BalAcc(\mathcal{C}l_{[{}^{\mathcal{C}}CDS,r^*]}) := \frac{1}{2}\left(\frac{TP}{n^+} + \frac{TN}{n^-}\right).$$

Another significant parameter is the *indeterminacy rate* of the classifier $\mathcal{C}l_{[^{\mathcal{C}}CDS_{Train}, r^*]}$ that is defined in the following way:

$$IR(\mathcal{C}l_{[^{\mathcal{C}}CDS_{Train}, r^*]}) := \frac{FI}{n_{Test}},$$

where n_{Test} is the cardinality of the set of all instances of the test-dataset. In the case of our example we obtain:

$$IR(\mathcal{C}l_{[^{\mathcal{C}}CDS_{Train}, r^*]}) = 0.502.$$

Notice that $IR(\mathcal{C}l_{[^{\mathcal{C}}CDS_{Train}, r^*]})$ is a parameter that regards the classification-process, while $IR(^{\mathcal{C}}CDS)$ represents an intrinsic property of the original dataset $^{\mathcal{C}}CDS$.

By definition, both the accuracy and the balanced accuracy of the classifier $\mathcal{C}l_{[^{\mathcal{C}}CDS, r^*]}$ depend on the choice of the threshold-value r^*. Figure 2 shows how the accuracy-values vary with the possible values of r^*. The accuracy reaches its maximum value (0.537) when $r^* = 0.501$. Also the maximum value of the balanced accuracy (0.554) is reached when $r^* = 0.501$.

After having acted as a classical epistemic agent, at time t_2, $Alice$ decides to transform her initial classical information into a quantum $\mathcal{C}$-data set (via a *stereographic encoding*). In this way, every classical object-state

$$\vec{x} = (x_i, x_j) \in \mathbb{R}^2$$

is transformed into a pure state $|\psi\rangle_{\vec{x}}$ of the Hilbert space $\mathbb{R}^3$. The result is a three-valued quantum $\mathcal{C}$-dataset

$$^{\mathcal{C}}QDS = (^{\mathcal{C}}\mathcal{H}, \ ^{\mathcal{C}}St_q, \ ^{\mathcal{C}}St_q^+, \ ^{\mathcal{C}}St_q^-, \ ^{\mathcal{C}}St_q^?),$$

where:

- $^{\mathcal{C}}\mathcal{H}$ is the Hilbert space $\mathbb{R}^3$;
- $n^+ = 352, \quad n^- = 359, \quad n^? = 339.$

$Alice$'s quantum information (at time t_2) is illustrated by Fig. 3, where only the positive instances (represented by the red points) and the negative instances (represented by the blue points) have been considered.

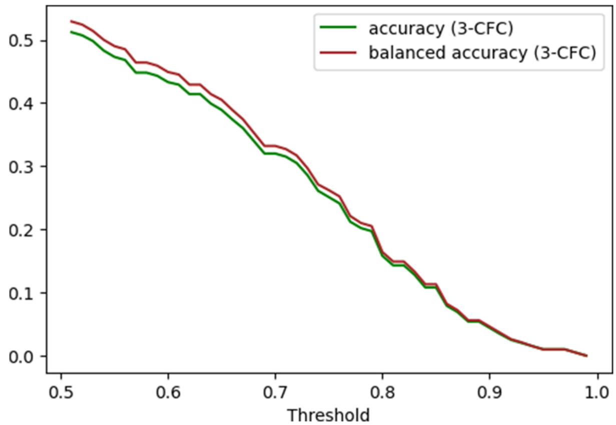

Fig. 2 How the accuracy-values vary with the threshold-values in the classical case

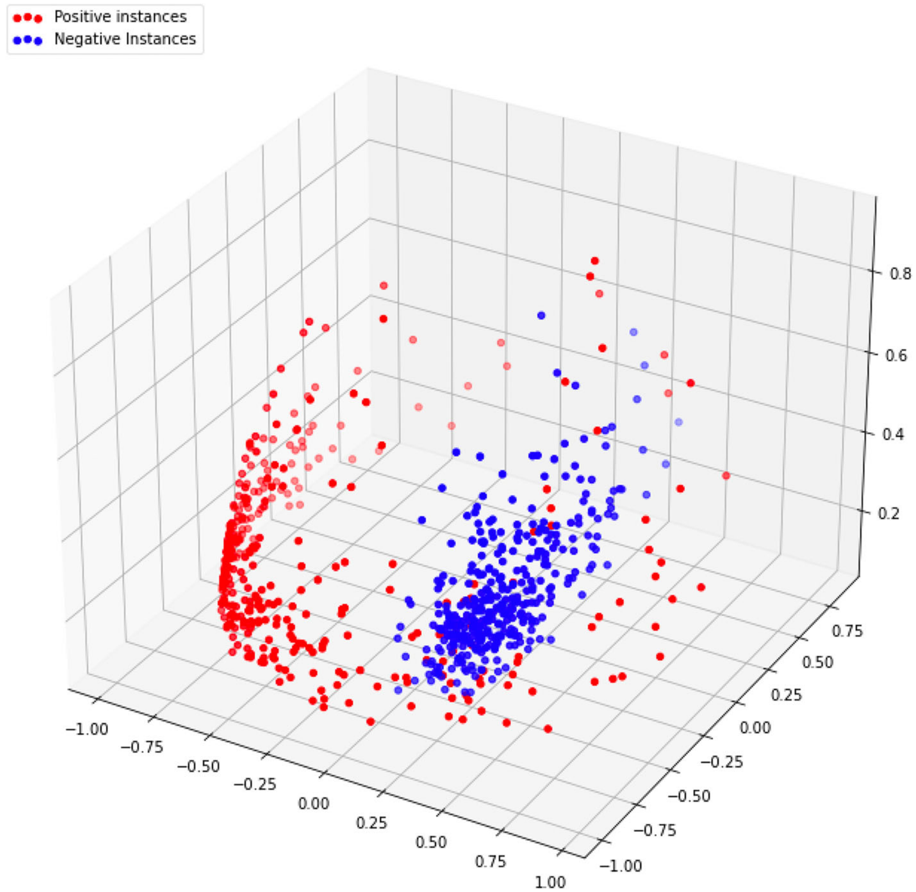

Fig. 3 *Alice*'s quantum information at time t_2

Our empirical simulation clearly shows why the concept of *quantum positive centroid* represents a "good" simulation of the intuitive concept of *Gestalt*. Both $Alice_H$ and $Alice_M$ have met $n^+(= 352)$ particular flowers that have been recognized as *roses*. On this basis, $Alice_H$ has constructed in her mind a *rose-Gestalt*: a kind of out of focus image that preserves an ambiguous memory of some general, vague features of the concrete flowers she had previously seen. This characteristic human procedure can be emulated by $Alice_M$, by referring to the mathematical concept of *quantum positive centroid* (of the quantum $\mathcal{C}$-dataset ${}^{\mathcal{C}}QDS$), where the mixed state

$$\rho^+ = \sum_i \left\{ \frac{1}{352} P_{|\psi_i\rangle} : |\psi_i\rangle \in {}^{\mathcal{C}}St^+ \right\}$$

represents a vague information that ambiguously alludes to all instances that $Alice_M$ had previously recognized as *roses*.

Alice can now proceed like in the classical case in order to control the *reliabilty* of her quantum fidelity-classifier $\mathcal{C}l_{[{}^{\mathcal{C}}QDS, r^*]}$, based on the quantum dataset ${}^{\mathcal{C}}QDS$ and on the choice of a threshold value r^*. The quantum concepts of *accuracy* and of *balanced accuracy*

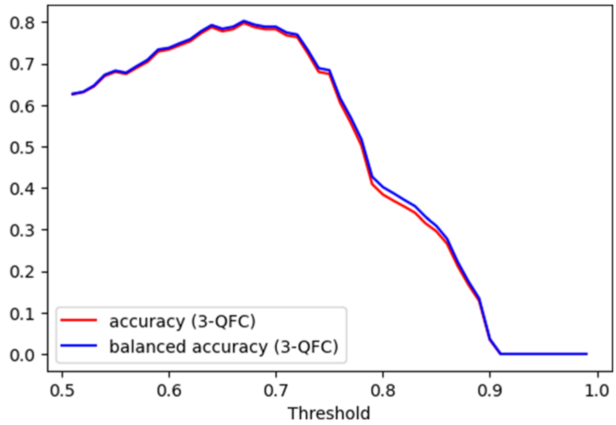

Fig. 4 How the accuracy-values vary with the threshold-values in the quantum case

are defined in the expected way. Figure 4 shows how the accuracy-values vary with the possible values of r^* in the quantum case. Both accuracies reach their maximum value when $r^* = 0.67$. The accuracy's maximum value is 0.798, while the balanced accuracy's maximum value is 0.803.

On this basis we can compare the accuracies of our quantum classifiers with the accuracies of the corresponding classical classifiers. Figure 5 clearly shows the supremacy of quantum classifiers.

As happens in the classical case, we can also define the indeterminacy rate of the quantum classifier $\mathcal{C}l_{[\mathcal{C}_{QDS_{Train}}, r^*]}$. One can show that:

$$IR(\mathcal{C}l_{[\mathcal{C}_{QDS_{Train}}, r^*]}) = 0.138.$$

Thus, in the quantum case the degree of uncertainty of $Alice$'s classification has decreased with respect to the classical case (where $IR(\mathcal{C}l_{[\mathcal{C}_{CDS_{Train}}, r^*]}) = 0.502$).

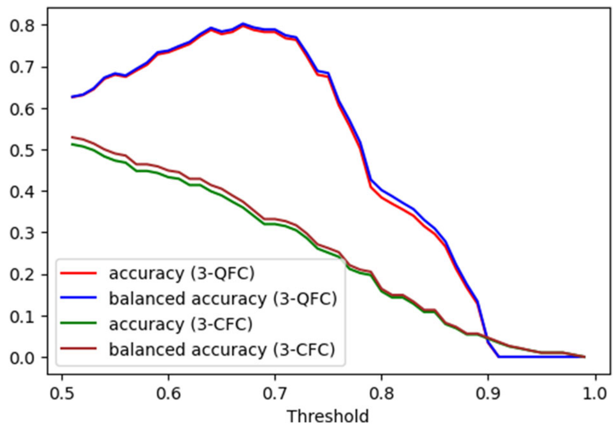

Fig. 5 The supremacy of quantum classifiers

From a logical point of view, the main ideas of a quantum approach to pattern recognition can be naturally reconstructed in the framework of a special version of *first-order quantum computational semantics*.[6] But this is a longer story that will be told in another paper.

Acknowledgements R. Giuntini and G. Sergioli are partially supported by the projects: i) "Ubiquitous Quantum Reality (UQR): understanding the natural processes under the light of quantum-like structures", funded by Fondazione di Sardegna (code: F73C22001360007); ii) "CORTEX The COst of Reasoning: Theory and EXperiments", funded by the Ministry of University and Research (Prin 2022, cod. 2022ZLLR3T): iii) "Quantum Models for Logic, Computation and Natural Processes (Qm4Np)" funded by the Ministry of University and Research (Prin-Pnrr 2022 cod. P2022A52CR). R. Giuntini is partially funded by the TÜV SÜD Foundation, the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder, as well as by the Technical University of Munich-Institute for Advanced Study.

Author Contributions All authors contributed equally to the conceptualization of the work, the execution of the experiments, and the writing of the manuscript.

Declarations

Competing interests The authors declare no competing interests.

References

1. Dalla Chiara, M.L., Giuntini, R., Leporini, R., Sergioli, S.: Quantum Computation and Logic. How Quantum Computers have inspired Logical Investigations, Springer, Berlin (2018)
2. Ehrenstein, W.H., Spillmann, L., Sarris, W.: Gestalt issues in modern neuroscience. Axiomathes **13**, 433–458 (2003)
3. Schuld, M., Petruccione, F.: Supervised Learning with Quantum Computers. Springer, Berlin (2018)
4. Sergioli, G., Bosyk, G.M., Santucci, E., Giuntini, R.: A Quantum-inspired version of the classification problem. Int. J. Theoretical Phys. **56**, 3880–3888 (2017)
5. Sergioli, G., Giuntini, R., Freytes, G.: A new quantum approach to binary classification. PLOS ONE **14**(5), e0216224 (2019)
6. Sergioli, G.: Quantum and Quantum-like Machine Learning. A Note on Similarities and Differences. Soft Comput. **24**, 10247–10255 (2019)

[6] See [1].