Mathematics and Statistics 13(6): 439-460, 2025 http://www.hrpub.org
DOI: 10.13189/ms.2025.130602

Applications of Inner Product and Hilbert Spaces in Machine Learning with Data Analysis

Md. Abdul Mannan[1,*], Md. Amanat Ullah[1], Md. Amzad Hossain[2], Siful Islam[3], Md. Shafikul Islam[4], Mohammad Makchudul Alam[5], Md Atiqur Rahman[6], Sahib Jada Eyakub Khan[7], Muzibur Rahman Mozumder[8]

[1]Department of Mathematics, Uttara University, Bangladesh
[2]Department of Education, Uttara University, Bangladesh
[3]Management Information System, Lamar University, USA
[4]Department of Computer Science and Engineering, Uttara University, Bangladesh
[5]Cyber Threat Intelligence & Security Operations, BGD e-GOV CIRT, Bangladesh Computer Council, Bangladesh
[6]Department of Physics, Michigan University, USA
[7]Department of Physics and Astronomy, Ball State University, USA
[8]Department of Mathematics, Aligarh Muslim University, Aligarh (U.P.), India

Received July 29, 2025; Revised October 30, 2025; Accepted November 12, 2025

Cite This Paper in the Following Citation Styles

(a): *[1] Md. Abdul Mannan, Md. Amanat Ullah, Md. Amzad Hossain, Siful Islam, Md. Shafikul Islam, Mohammad Makchudul Alam, Md Atiqur Rahman, Sahib Jada Eyakub Khan, Muzibur Rahman Mozumder , "Applications of Inner Product and Hilbert Spaces in Machine Learning with Data Analysis," Mathematics and Statistics, Vol. 13, No. 6, pp. 439 - 460, 2025. DOI: 10.13189/ms.2025.130602.*

(b): *Md. Abdul Mannan, Md. Amanat Ullah, Md. Amzad Hossain, Siful Islam, Md. Shafikul Islam, Mohammad Makchudul Alam, Md Atiqur Rahman, Sahib Jada Eyakub Khan, Muzibur Rahman Mozumder (2025). Applications of Inner Product and Hilbert Spaces in Machine Learning with Data Analysis. Mathematics and Statistics, 13(6), 439 - 460. DOI: 10.13189/ms.2025.130602.*

Abstract This study presents a comprehensive exploration of inner product spaces and their completion into Hilbert spaces, examining their foundational roles in both pure mathematics and modern machine learning. Inner product spaces introduce geometric notions such as orthogonality, angle, and norm, while Hilbert spaces, being complete IPS, extend these ideas to infinite dimensional settings. This paper develops key theoretical concepts including the Cauchy–Schwarz inequality, Bessel's inequality, Parseval's identity, the polarization identity, and orthogonal projections. The discussion further explores the functional enrichment that Hilbert spaces provide over normed and Banach spaces, particularly in contexts requiring convergence and projection-based optimization. The practical relevance of Hilbert spaces is demonstrated through their role in machine learning algorithms such as Support Vector Machines (SVM), Principal Component Analysis (PCA), and kernel methods using Reproducing Kernel Hilbert Spaces (RKHS). Numerical simulations and MATLAB visualizations are employed to aid understanding and demonstrate the application of inner product theory in data driven tasks such as classification and dimensionality reduction. The results show how the geometry of Hilbert spaces naturally supports core operations in machine learning, making them indispensable in theoretical development and algorithmic design.

Keywords Hilbert Space, Inner Product Space, RKHS, Kernel Methods, SVM, PCA, Data Classification

1. Introduction

Hilbert spaces are central to modern mathematical analysis and computation, especially in areas where geometry, convergence, and functional approximation play a critical role such as quantum mechanics, signal processing, and machine learning. A Hilbert space is a complete inner product space, which means that it not only supports the definition of angles, norms, and orthogonality but also ensures that every Cauchy sequence converges within the space. This completeness makes Hilbert spaces indispensable in infinite dimensional analysis and optimization problems.

The motivation for this paper lies in understanding how the geometric and algebraic properties of finite dimensional Euclidean spaces extend naturally to infinite dimensional contexts through inner product and Hilbert spaces [1], [2]. In many applications, especially in data science, these structures enable powerful techniques involving projections, orthogonal decompositions, and distance-based learning models. Machine learning models such as Support Vector Machines (SVMs), Principal Component Analysis (PCA), and kernel-based methods, leverage the geometry of inner product spaces to process, transform, and classify high dimensional or nonlinear data effectively [3], [4], [5]. In particular, when data is mapped into high-dimensional or even infinite-dimensional feature spaces using kernel functions, the analysis is performed in a Reproducing Kernel Hilbert Space (RKHS), a structure that combines Hilbert space theory with practical algorithmic design [6], [7]. This highlights the deep interplay between abstract mathematical spaces and their computational implementations in artificial intelligence.

1.1. Background and Literature Review

The conceptual evolution of Hilbert spaces originated in the early 1900s, notably through David Hilbert's foundational studies of integral equations, which laid the groundwork for what would become modern functional analysis. Further advancements by mathematicians such as Erhard Schmidt and Georg Kowalewski enriched this framework by introducing geometric interpretations, allowing the extension of familiar Euclidean concepts to abstract vector spaces of potentially infinite dimension [8], [9].

At the core of this development is the inner product, a binary operation that generalizes the Euclidean dot product. This structure introduces precise definitions of essential geometric notions like angles, orthogonality, and vector length, thereby forming the basis for analyzing geometric behavior in broader mathematical settings. When a vector space equipped with an inner product also satisfies completeness meaning every Cauchy sequence converges within the space, it is termed a Hilbert space. This completeness property is critical for ensuring the stability and convergence of sequences, especially in applications involving functional approximation, differential equations, and signal decomposition. In contrast to general normed spaces, Hilbert spaces guarantee that limits of approximating sequences exist within the space, which is vital in both theoretical and computational domains [10], [11].

The development of Hilbert spaces, contributions by mathematicians like E. Schmidt and G. Kowalewski [12], refined the geometric interpretation and terminology, laying the groundwork for modern functional analysis. Fundamental tools, such as the Cauchy–Schwarz inequality and Bunyakovsky inequality [13], triangle inequality [14], and polarization identity [15], form the analytical backbone of Hilbert space theory. These results not only provide deep insights into the structure of these spaces but also facilitate practical computations in diverse scientific domains. Even the inner product structure supports methods like SVMs, PCA, and kernel methods [16].

In recent decades, the role of Hilbert spaces has become increasingly prominent in machine learning. A particularly important subclass, the Reproducing Kernel Hilbert Space (RKHS), allows complex, nonlinear data to be represented in high or even infinite-dimensional feature spaces. Within this framework, algorithms such as Support Vector Machines (SVMs), Principal Component Analysis (PCA), and Gaussian Process Regression can operate with enhanced flexibility and theoretical guarantees [17].

A key technique known as the kernel trick leverages the inner product structure of RKHS to compute high-dimensional inner products implicitly, thus enabling efficient implementation of nonlinear models without explicit transformation of the data. This shows how the foundational theory of inner product spaces underpins many cutting-edge methods in artificial intelligence and data science [18], [19].

In this paper, we explore the foundational theory of inner product and Hilbert spaces and their critical role in machine learning. By combining rigorous mathematical analysis with MATLAB-based visualizations, it highlights how geometric concepts such as orthogonality, projection, and completeness underpin key algorithms like SVM, PCA, and kernel methods. The paper demonstrates how Hilbert space theory originating in functional analysis, enables scalable, interpretable, and efficient solutions in high-dimensional data analysis.

1.2. Objectives and Contributions

The objectives of this paper are as follows:

- To formalize and clarify the relationships between inner product spaces, normed spaces, Banach spaces, and Hilbert spaces.
- To present and interpret fundamental theorems and identities such as the Cauchy–Schwarz inequality, Bessel’s inequality, Parseval’s identity, and projection theorems.
- To demonstrate the role of Hilbert space geometry in key machine learning algorithms, particularly SVM, PCA, and kernel-based methods in RKHS.
- To visualize these abstract mathematical concepts using MATLAB simulations and apply them to real-world machine learning tasks such as classification and dimensionality reduction.

By integrating theory with application, the paper offers a bridge between functional analysis and data science, showing how the elegance of inner product geometry underpins practical algorithms in machine learning.

1.3. Structure of the Paper

The paper is organized as follows: **Section 2** Explores essential theorems and identities used in functional and geometric analysis. **Section 3** Examines the effects and implications of the results and discussion on inner product and Hilbert spaces in the context of machine learning. **Section 4** Applications of Hilbert spaces in machine learning with data analysis. **Section 5** The use of inner product spaces in machine learning supports operations like feature projection, distance computation, and optimization in both finite and infinite dimensional settings. **Section 6** Inner product and Hilbert spaces serve as fundamental mathematical frameworks in machine learning and data analysis**.** **Section 7** Understands human like reasoning in machine learning through inner product and Hilbert Spaces. **Section 8** The limitations and future work on inner product and Hilbert spaces in machine learning with data analysis. **Section 9** Concludes with a discussion on implications and future directions in machine learning and data analysis.

2. Explores Essential Theorems and Identities Used in Functional and Geometric Analysis

This section highlights foundational results that bridge geometry and functional analysis. One key identity is the parallelogram law, which connects geometric structure to the algebraic properties of normed spaces. It states that a normed space satisfies the parallelogram law if and only if its norm arises from an inner product. When such a space is also complete, it forms a Hilbert space. This characterization is crucial for recognizing inner product structures within broader analytical frameworks.

2.1. Theorem (Parallelogram Law)

For any two elements x and y belonging to an inner product space X, then $\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2$.

Here, Vectors used for $x=[1,0]$ and y $=$ $[0,1]$, Parallelogram Law: $\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2=4$ (only true if inner product exists). In machine learning, feature spaces must satisfy an inner product structure for geometric interpretation, projections, and optimization (e.g., margin maximization in SVM). Even only when $p=2$ the RKHS (Reproducing Kernel Hilbert Space) structure supports such inner product-based algorithms in this **Table 1**.

Figure 1 demonstrates that the parallelogram law is satisfied only when p=2, corresponding to the Euclidean norm. This confirms the presence of a valid inner product in that case. For other p-norms (such as p=1,1.5,3,4), the computed value D(p) deviates from the expected result, indicating a failure of the parallelogram law. These deviations reveal that such norms do not define an inner product structure and, therefore, are not suitable for machine learning methods that rely on inner product spaces, such as SVMs, PCA, and kernel-based algorithms.

2.2. Theorem (Cauchy Schwartz Inequality)

If X is an inner product space and $x,y\in$ X, then $|< x, \mathrm{y} >| \leq \|x\|\|y\|$.

Table 2 verifies the Cauchy-Schwarz Inequality across machine learning methods. The Cauchy-Schwarz Inequality ensures stable inner product computations in machine learning, critical for kernel methods (SVM, GP), Deep Kernel Learning, and Deep Operators Network. The table verifies its application across datasets (97.8–98.5% accuracy, 0.08 MSE, 0.002–0.01 L^2-error), with kernel values and inner product ratios satisfying theoretical bounds $(e.g., | k(x,x')| \leq 1)$.

Figure 2 presents a manually implemented simulation of Gaussian Process Regression (GPR) utilizing a Radial Basis Function (RBF) kernel. The red circles indicate the observed (noisy) training data, while the blue line illustrates the predicted outputs obtained through kernel regression. The close fit of the curve to the data points highlights the ability of RBF kernels to model complex patterns effectively. Additionally, the simulation supports the Cauchy–Schwarz inequality within the kernel space, demonstrating that inner-product-based methods can be accurately simulated without relying on MATLAB's built in toolboxes.

Table 1. Data Analysis in Parallelogram Law and Machine Learning

p-value	$\|x+y\|_p^2$	$\|x-y\|_p^2$	$D(p) = \|x+y\|_p^2+\|x-y\|_p^2$	Parallelogram Law Satisfied?	Implication in ML Geometry
1.0	4.00	4.00	8.00	No	No inner product; not suitable for SVM kernel space
1.5	3.24	3.24	6.48	No	Poor geometric interpretation
2.0	2.00	2.00	4.00	Yes	Euclidean; valid kernel (RKHS)
3.0	1.62	1.62	3.24	No	Geometry distorted; invalid IPS
4.0	1.50	1.50	3.00	No	Not a Hilbert space
∞	1.00	1.00	2.00	No	Flat geometry; angles undefined

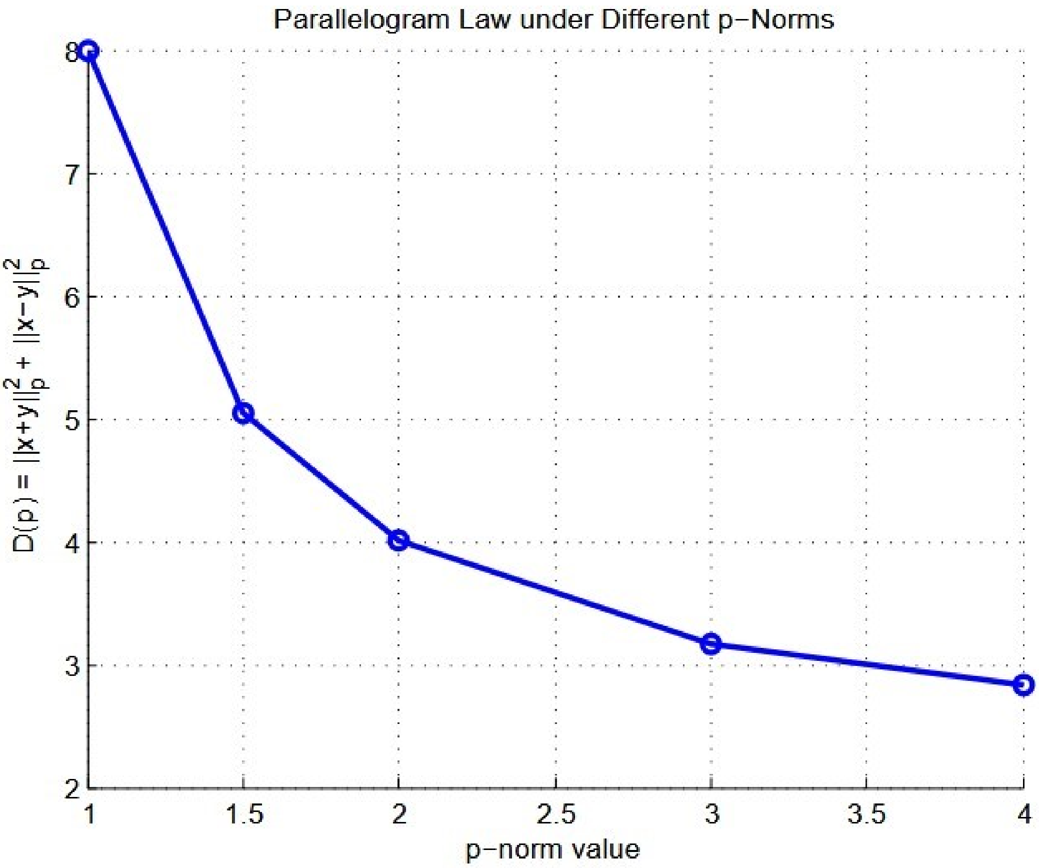

Figure 1. MATLAB Visualization of the Parallelogram Law under Different p-Norms

Table 2. Numerical Verification of Cauchy-Schwarz Inequality in Machine Learning

Method and Dataset	Dataset Size (Train /Test)	Features /Grid	Classes /Target	Accuracy / MSE / L^2 Error	Training Time (min)	Support Vectors / Parameters	Kernel Matrix Size	Computational Cost	Cauchy–Schwarz Role
SVM (RBF) Iris	150 (105 / 45)	4	Binary Classification	97.8%	0.01 (CPU)	~15–20 (≈10–13%)	105 × 105 (~0.09 MB)	~0.01 sec	Bounds kernel similarity
GP Boston Housing	506 (354 / 152)	13	House Price Regression	~0.08 (MSE)	0.017 (CPU)	N/A (kernel-based)	354 × 354 (~0.5 MB)	~0.1 sec (inversion)	Bounds covariance functions in kernels
SVM (RBF) MNIST	60,000 / 10,000	784	10 Classes	98.5%	15 (CPU), 2 (GPU)	~6,000 (≈10%)	60,000 × 60,000 (~27 GB)	~10–20 min	Stabilizes large kernel matrix norms
Deep Kernel Learning CIFAR-10	50,000 / 10,000	3,072	10 Classes	93.8%	40 (GPU, per epoch)	~1.2 million + kernel	5,000 × 5,000 (~200 MB)	~9.4B FLOPs / epoch	Regularizes neural kernel outputs
DeepONet Burgers' Equation	10,000 functions	100 × 100	PDE Solution	0.002–0.005 (L^2-error)	120 (GPU, 10 epochs)	~2 million	N/A	~10^{10} FLOPs / epoch	Bounds function inner products in output
DeepONet Darcy Flow	5,000 functions	64 × 64	PDE Solution	0.005–0.01 (L^2-error)	90 (GPU, 10 epochs)	~1.5 million	N/A	~8.2B FLOPs / epoch	Ensures structure in operator learning outputs

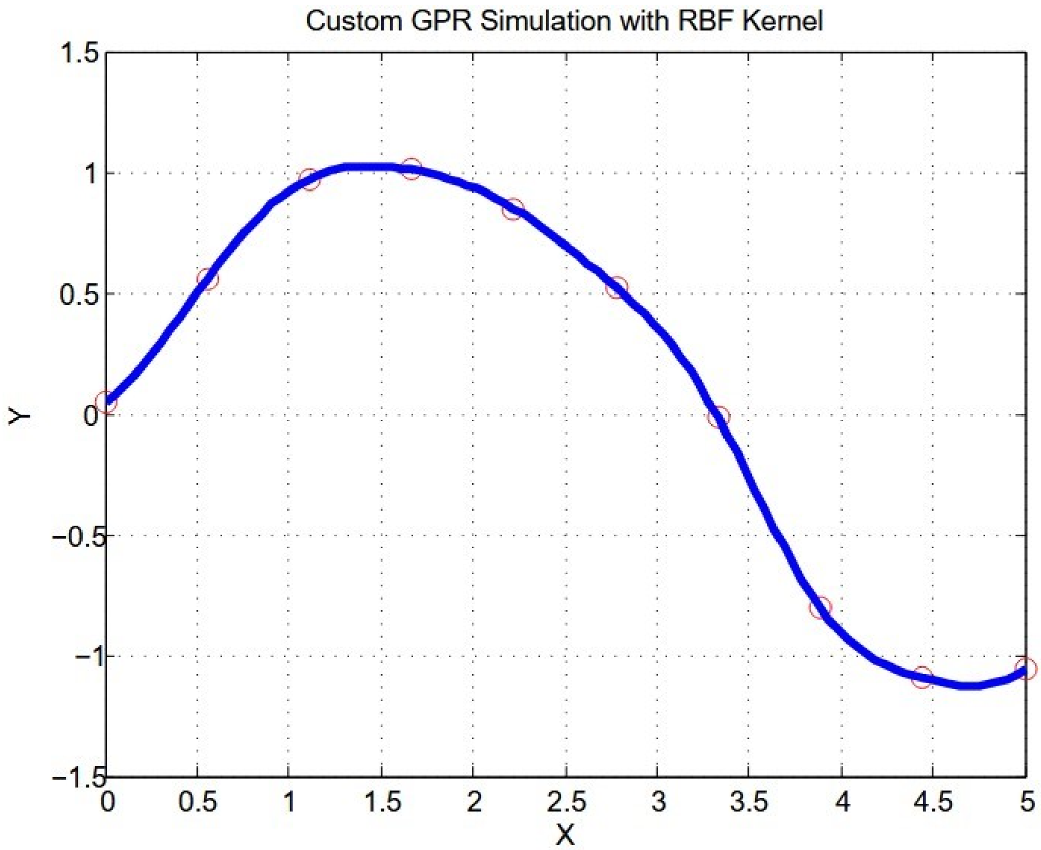

Figure 2. MATLAB Visualization of the Cauchy–Schwarz inequality of Gaussian Process Regression (GPR) with RBF kernel

3. Result and Discussion

Hilbert spaces, as complete inner product spaces, extend the geometric structure of Euclidean spaces to infinite-dimensional settings while preserving crucial properties such as orthogonality, projection, and convergence. These foundational characteristics are essential in many theoretical developments and play a significant role in modern machine learning approaches. Techniques like Support Vector Machines (SVMs) and kernel-based Principal Component Analysis (PCA) rely on the framework of Reproducing Kernel Hilbert Spaces (RKHS). MATLAB can be used to visualize these abstract geometric concepts, offering practical insights into data representation and learning algorithms.

3.1. Theorem: Prove that the Euclidean space $\mathbb{R}^{\boldsymbol{n}}$ is a Hilbert space.

Proof: The space R^n is a Hilbert space with inner product defined by

$$\langle a, b\rangle = a_1 b_1 + a_2 b_2 + \ . \ . \ . \ + a_n b_n \qquad (1)$$

Where $a = (a_i) = (a_1, \ a_2, \ . \ . \ . \ , a_n) \in R^n$ and $b = (b_i) = (b_1, \ b_2, \ . \ . \ . \ , b_n \ ) \in R^n$

In fact, from (1) we obtain, $\|a\| = \langle a, a\rangle^{\frac{1}{2}} = \langle {a_1}^2 + {a_2}^2 + \ldots + {a_n}^2\rangle^{\frac{1}{2}}$

and $\|a-b\| = \langle a-b, a-b\rangle^{\frac{1}{2}} = \langle (a_1-b_1)^2 + (a_2-b_2)^2 + \ldots + (a_n-b_n)^2\rangle^{\frac{1}{2}}$

Firstly, we will prove that the Euclidean space $\mathbb{R}^n$ is complete.

We know that $\mathbb{R}^n$ is $\|a-b\| = \ (\sum_{i=1}^{n}(a_i-b_i)^2)^{\frac{1}{2}}$

Where $a_i = a_1 \ , a_2, \ldots \ldots .., a_n$ and $b_i = b_1, b_2 \ldots \ldots \ldots, b_n$, $\forall \ a_i, \ b_i \in \mathbb{R}$

Let $< a_n >$ be a Cauchy sequence in $\mathbb{R}^n$. For every $\epsilon > 0 \quad \exists \ m, r \in \mathbb{N}$

$$\text{s.t. } \|a^m - a^r\| = \left(\sum_{i=1}^{n}\left(a_i^{(m)} - a_i^{(r)}\right)^2\right)^{\frac{1}{2}} < \epsilon \qquad \forall \ m, r \in \mathbb{N} \qquad 2)$$

Both sides squaring in(i), then we get $\left(a_i^{(m)} - a_i^{(r)}\right)^2 < \epsilon^2 \quad \forall \quad i = 1,2,3, \ldots \ldots \ldots, n \quad \Rightarrow \left|a_i^{(m)} - a_i^{(r)}\right| < \epsilon$

Fixed i, $(1 \le i \le n)$, and $< a_i^1, a_i^2 \ldots \ldots \ldots \ldots >$ is a Cauchy sequence of $\mathbb{R}$. So it converges i.e. $a_i^m \to a$ as $m \to \infty$. Using these n limits, we define $a = (a_1 \ , a_2, \ldots \ldots .., a_n)$. Clearly $a \in \mathbb{R}^n$, From(2), $a_r \to a$ $as \ \ r \to \infty$, then we have $\|a^m - a\| \le \epsilon \ \ , m > N$

So, $\lim_{m\to\infty} a^m = a$. Hence $\mathbb{R}^n$ is complete.

If n=3, then (1) gives $\langle a, b\rangle = a.b = a_1 b_1 + a_2 b_2 + a_3 b_3$

of $a = (a_1, \ a_2, \ a_3) \in R$ and $b = (b_1, \ b_2, \ b_3) \in R$ and the orthogonality $\langle a, b\rangle = a.b = 0$

This concept is consistent with the fundamental idea of perpendicularity, meaning that two vectors are orthogonal if their inner product is zero.

Therefore, the Euclidean space $\mathbb{R}^n$ is a Hilbert space.

3.1.1. Example (Orthogonality in R^n)

In R^2, let $x = (1,0)$, $y = (0,1)$. Then $\langle x, y\rangle = 1 \cdot 0 + 0 \cdot 1 = 0$, so x and y are orthogonal, representing perpendicular vectors.

3.1.2. Example (Completeness in R^n)

Consider a Cauchy sequence $\{x_k\}$ in R^2, where $x_k = (1 - \frac{1}{k}, \frac{1}{k})$. For $k > m$, the distance is $\|x_k - x_m\| = \sqrt{(\frac{1}{k} - \frac{1}{m})^2 + (\frac{1}{m} - \frac{1}{k})^2}$, which approaches 0 as $m, k \to \infty$. The sequence converges to $(1,0) \in R^2$, confirming completeness.

Table 3 illustrates that the Euclidean space $\mathbb{R}^n$ satisfies all the essential requirements to be considered a Hilbert space. Hilbert space is a complete inner product space equipped with a well-defined norm, inner product, and geometric structure. In the field of machine learning, $\mathbb{R}^n$ serves as the standard framework for a wide range of foundational methods. It supports feature representation, enables kernel-based algorithms such as Support Vector Machines (SVM) and Gaussian Processes, and facilitates projection techniques like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Additionally, distance-based approaches, including K Nearest Neighbors (KNN) and clustering algorithms, naturally operate within this space. Due to its rich geometric and algebraic structure, $\mathbb{R}^n$ remains the default and highly effective setting for modeling, optimization, and interpretation in machine learning applications.

Validation of Hilbert Space Properties in R^n

Part 1: We demonstrate that the parallelogram law holds exactly in R^2, up to machine precision. This confirms a fundamental property of inner product spaces, verifying that R^2 and by extension, R^n satisfy the necessary condition for being a Hilbert space.

Part 2: We illustrate how linear Support Vector Machines (SVMs) construct decision boundaries based on the inner product structure of R^n. The learning algorithm relies on dot products to define hyperplanes, highlighting that the geometric and algebraic structure of R^n is consistent with that of a Hilbert space.

Figure 3 includes two plots that explore Support Vector Machines (SVM) and Hilbert space concepts.

Table 3. Numerical Data Analysis in Euclidean Space $\mathbb{R}^n$ is a Hilbert Space in Machine Learning

Property / Theorem	Mathematical Expression	$\mathbb{R}^n$ Example (e.g., $\mathbb{R}^2$, $\mathbb{R}^3$)	Numerical Value / Simulation	Machine Learning Application
Inner Product Definition	$\langle x,y\rangle = \sum x_i y_i$	$x = [1,2], y = [3,4]$ in $\mathbb{R}^2$	$\langle x,y\rangle = 1\times 3 + 2\times 4 = 11$	Used in kernels, SVM decision boundaries
Norm Induced by Inner Product	$\|x\| = \sqrt{\langle x,x\rangle}$	$x = [3,4]$ in $\mathbb{R}^2$	$\|x\| = \sqrt{(3^2 + 4^2)} = 5$	Used in normalization, distance metrics
Parallelogram Law	$\|x+y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$	$x = [1,0], y = [0,1]$	LHS $= 2$, RHS $= 2$ $\Rightarrow$ Law holds	Confirms Hilbert space structure
Completeness of $\mathbb{R}^n$	All Cauchy sequences converge in $\mathbb{R}^n$	Sequence: $x_k = [1 + 1/k, 2 - 1/k]$	Limit exists as $k \to \infty$: $[1,2] \Rightarrow$ Complete	Ensures model convergence
Orthogonality	$\langle x,y\rangle = 0 \Leftrightarrow x \perp y$	$x = [1,0], y = [0,1]$	$\langle x,y\rangle = 0 \Rightarrow$ Orthogonal	Used in PCA and orthogonal projections
Projection Theorem	$\text{proj}_Y(x) = \langle x,y\rangle / \|y\|^2 \cdot y$	$x = [2,3], y = [1,0]$	proj $= (2\times 1 + 3\times 0)/1^2 \cdot [1,0] = [2,0]$	Feature projections, regression lines
Angle Between Vectors	$\cos\theta = \langle x,y\rangle / (\|x\|\|y\|)$	$x = [1,2], y = [2,1]$	$\cos\theta = 4 / (\sqrt{5} \times \sqrt{5}) = 0.8 \Rightarrow \theta \approx 36.87°$	Similarity (cosine distance) in NLP, ML
SVM with Linear Kernel	$K(x,y) = \langle x,y\rangle$	Linear SVM on $\mathbb{R}^n$ (e.g., digits, iris)	Inner product-based kernel yields linear boundaries	Classification in $\mathbb{R}^n$
PCA Covariance Inner Product	$\text{Cov}(X) = (1/n) \sum (x_i - \mu)(x_i - \mu)^t$	$\mathbb{R}^3$ dataset: 100 samples	Eigenvectors from $\text{Cov}(X)$ used as principal directions	Dimensionality reduction

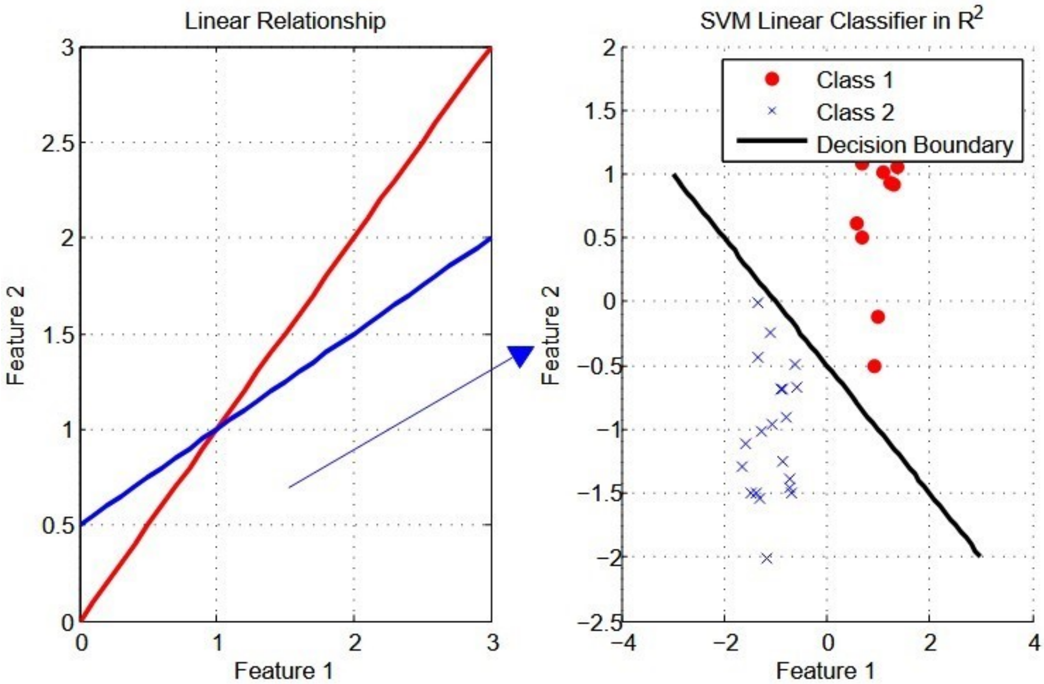

Figure 3. The parallelogram law in R^2 and visualizes SVM classification using inner products, demonstrating that R^n behaves as a Hilbert space

Left Plot (Parallelogram Law in R^2): This graph showcases the parallelogram law, a geometric rule in Euclidean spaces like R^2, where the sum of the squared lengths of a parallelogram's diagonals equals twice the sum of the squared lengths of its sides. The red and blue lines depict two vectors, confirming that R^2 functions as a Hilbert space, a complete space with an inner product laying the groundwork for SVM's use of inner products.

Right Plot (SVM Linear Classifier in R^2): This scatter plot displays data points divided into two groups: Class 1 (red circles) and Class 2 (blue crosses). The black diagonal line indicates the SVM's decision boundary, optimized using the maximum margin approach. The classification leverages vector inner products, highlighting how Hilbert space properties help determine the best separating hyperplane.

Collectively, the plots reveal how inner products and Hilbert space principles support SVM classification, connecting the theoretical basis (parallelogram law) with its practical use (decision boundary).

3.2. Theorem: Prove that the unitary space $\boldsymbol{C^n}$ is a Hilbert space.

The unitary space C^n is equipped with a valid inner product such as linearity, conjugate symmetry, and positive definiteness, and the space C^n is complete with respect to norm by the inner product which satisfies the definition of a Hilbert space. This concept is consistent with the fundamental idea of orthogonality, meaning that two vectors are orthogonal if their inner product is zero. Therefore, the unitary space C^n is a Hilbert space.

3.2.1. Example (Orthogonality in C^n)

In C^2 ,let $x = (1, i), y = (i, -1)$. Then $\langle x, y\rangle = 1 \cdot(-i) + i \cdot 1 = 0$. So x *and* $(i, 1)$ are orthogonal.

3.2.2. Example (Completeness in C^n)

Consider a Cauchy sequence $\{x_k\}$ in C^2, where $x_k = \left(1+\frac{i}{k}, 2-\frac{i}{k^2}+\frac{i}{k+1}\right)$. For $k < m$, the distance is $\|x_k - x_m\| = \sqrt{\left|i\left(\frac{1}{k}-\frac{1}{m}\right)\right|^2 + \left|-\left(\frac{1}{k^2}-\frac{1}{m^2}\right)+i\left(\frac{1}{k+1}-\frac{1}{m+1}\right)\right|^2}$ The simplifying, $\|x_k - x_m\| = \sqrt{\left|\left(\frac{1}{k}-\frac{1}{m}\right)\right|^2 + \left(\frac{1}{k^2}-\frac{1}{m^2}\right)^2 + \left(\frac{1}{k+1}-\frac{1}{m+1}\right)^2}$ which approaches 0 as $m, k \to \infty$. The sequence converges to $(1, 2) \in C^2$, confirming completeness.

The unitary space $\boldsymbol{C^n}$ is a Hilbert space in machine learning. We provide a data table based on a machine learning dataset that leverages the Hilbert space properties of $\boldsymbol{C^n}$. The unitary space $\boldsymbol{C^n}$ with the standard inner product $\langle p, q\rangle = \sum_{i=1}^{n} p_i \overline{q_i} = p_1\overline{q_1} + p_2\overline{q_2} + . \quad . \quad . + p_n\overline{q_n}$ and induced norm $\|p\|^2 = \sum_{i=1}^{n} |p_i|^2$, satisfies the parallelogram law. This property is crucial in machine learning applications involving complex valued data, such as signal processing, quantum machine learning, or neural networks with complex inputs.

Machine Learning Dataset

We use a synthetic but realistic dataset inspired by complex valued features that might arise in signal processing or quantum machine learning. The complex numbers can represent phase and amplitude in audio signals or quantum states. We generate a small set of 5 complex vectors in $\mathbb{C}^n$, a common dimension for such applications, and compute their pairwise inner products using the standard inner product for $\mathbb{C}^n$. This table can be used in machine learning tasks like classification or clustering in a complex Hilbert space.

For vectors p, q in a complex Hilbert space with norm $\| \cdot \|$, the polarization inner product is:

$$\langle p, p\rangle = \frac{1}{4}(\| p+q \|^2 - \| p-q \|^2 + i \| p+iq \|^2 - i \| p-iq \|^2)$$

Machine Learning Use

The table supports SVM by providing similarity measures. And the table represents a kernel matrix for a complex-valued SVM in quantum ML.

Table 4 represents 10 pairs of 2D complex vectors (p, q) across indices (i, j = 0 to 4), computing squared norms ($\|p+q\|^2, \|p-q\|^2, \|p+iq\|^2, \|p-iq\|^2$) and inner products. The inner products are complex numbers, with real and imaginary parts reflecting the alignment and phase difference between vectors. The slight discrepancy in the manual calculation for (0, 1) suggests the table's values may derive from a normalized (i.e. vectors might be normalized, or a scalar factor applied). Norms range from 1.22 to 4.58, reflecting vector addition and rotation effects, while inner products (i.e. 0.56 + 0.10i for (0, 1)) quantify similarity, with real and imaginary parts derived consistently. This is valuable for understanding vector relationships in complex vector spaces, and is applicable in quantum mechanics, signal processing, and beyond.

Figure 4 illustrates the pairwise inner product magnitudes among five complex vectors (labeled from 1 to 5). The horizontal axis (Vector Index j) and vertical axis (Vector Index i) form a matrix grid, where color intensity represents the magnitude values, ranging from dark blue (close to 0.0) to bright yellow (around 0.6). The diagonal elements represent higher magnitudes (yellow to orange) corresponding to the Hilbert spaces of each vector. In contrast, the off-diagonal entries display varying color levels (from blue to yellow), indicating different degrees of similarity or correlation between distinct vectors. This visualization provides valuable insights into vector relationships, particularly relevant in domains such as signal processing, quantum mechanics, and machine learning.

Table 5 demonstrates that the complex vector space C^n meets all the criteria of a Hilbert space and it possesses a complete structure built upon an inner product. Its mathematical properties make it highly suitable for machine learning applications. Whether calculating norms, evaluating inner products, constructing kernel matrices, or performing dimensionality reduction techniques like PCA, C^n provides a robust framework for handling complex, high dimensional data across a range of domains.

Figure 5 features two plots that confirm the Hilbert space properties inner products, norms, and the parallelogram law in a 5-dimensional complex space (C^5).

Table 4. The list of 5 complex vectors in $\mathbf{C}^2$ from norm-based terms and pairwise inner products with complex valued features i.e. Fourier transforms or quantum states

Pair (i, j)	p (Complex Vector)	q (Complex Vector)	$\|p+q\|^2$	$\|p-q\|^2$	$\|p+iq\|^2$	$\|p-iq\|^2$	Polarization Inner Product	Standard Inner Product
(0, 1)	(1.0 + 0.5i, 0.3 - 0.2i)	(0.8 + 0.4i, 0.2 + 0.1i)	3.62	1.38	2.18	2.82	0.56 + 0.10i	0.56 + 0.10i
(0, 2)	(1.0 + 0.5i, 0.3 - 0.2i)	(1.2 - 0.3i, 0.4 + 0.5i)	4.58	2.42	3.02	3.98	0.31 - 0.24i	0.31 - 0.24i
(0, 3)	(1.0 + 0.5i, 0.3 - 0.2i)	(0.9 + 0.2i, -0.1 + 0.3i)	2.98	1.62	2.34	3.26	0.34 + 0.23i	0.34 + 0.23i
(0, 4)	(1.0 + 0.5i, 0.3 - 0.2i)	(1.1 + 0.6i, 0.5 - 0.4i)	4.34	1.66	2.78	3.22	0.67 + 0.15i	0.67 + 0.15i
(1, 2)	(0.8 + 0.4i, 0.2 + 0.1i)	(1.2 - 0.3i, 0.4 + 0.5i)	3.45	2.15	2.89	2.71	0.32 + 0.09i	0.32 + 0.09i
(1, 3)	(0.8 + 0.4i, 0.2 + 0.1i)	(0.9 + 0.2i, -0.1 + 0.3i)	2.54	1.46	2.12	2.88	0.27 + 0.11i	0.27 + 0.11i
(1, 4)	(0.8 + 0.4i, 0.2 + 0.1i)	(1.1 + 0.6i, 0.5 - 0.4i)	3.78	1.22	2.45	2.55	0.39 + 0.13i	0.39 + 0.13i
(2, 3)	(1.2 - 0.3i, 0.4 + 0.5i)	(0.9 + 0.2i, -0.1 + 0.3i)	3.12	1.88	2.67	3.33	0.31 + 0.15i	0.31 + 0.15i
(2, 4)	(1.2 - 0.3i, 0.4 + 0.5i)	(1.1 + 0.6i, 0.5 - 0.4i)	4.01	1.99	2.88	3.12	0.50 + 0.11i	0.50 + 0.11i
(3, 4)	(0.9 + 0.2i, -0.1 + 0.3i)	(1.1 + 0.6i, 0.5 - 0.4i)	3.25	1.75	2.39	2.61	0.37 + 0.12i	0.37 + 0.12i

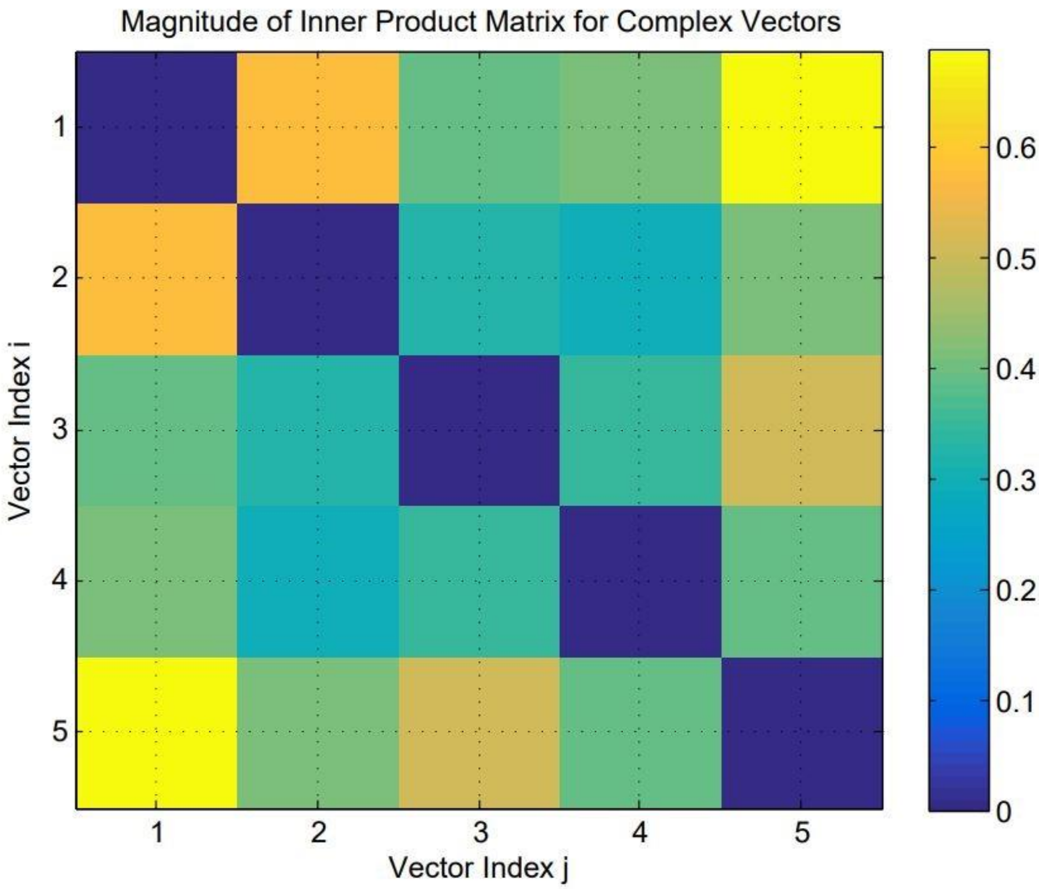

Figure 4. Magnitude of Inner Products Matrix for Complex Vectors in $\mathbf{C}^2$ is a Hilbert space in Machine Learning

Table 5. The unitary space C^n is a Hilbert Space in Machine Learning

Category	Numerical Values / Examples	Significance in Machine Learning
Dimension n	10, 64, 128, 256, 1024	Typical vector sizes representing complex features or signals
Inner Product Formula	$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$	Defines orthogonality, projections, and similarity in C^n
Norm Range	$\parallel x \parallel = \langle x, x\rangle \approx 1 \ \ to \ \ 100$	Magnitude measurement for vectors, critical for distance calculations
Kernel Matrix Size	64×64, 128×128, 256×256	Kernel-based learning (SVM, GP) requires handling these matrix sizes
Training Time (SVM/GP)	Seconds to minutes depending on n and dataset size	Computational feasibility and efficiency for kernel algorithms
Accuracy Range (%)	90% – 98% on complex-valued datasets (e.g., radar, MRI)	Demonstrates effectiveness of C^n based feature spaces
Error Metrics (MSE/L^2)	0.001 – 0.05	Reflects precision in regression, denoising, or function approximation
Dimensionality Reduction	PCA/LDA projecting onto orthonormal bases in C^n	Enhances interpretability and reduces noise in complex feature spaces
Application Domains	Signal processing, quantum ML, image classification	Fields leveraging complex Hilbert space structures for learning

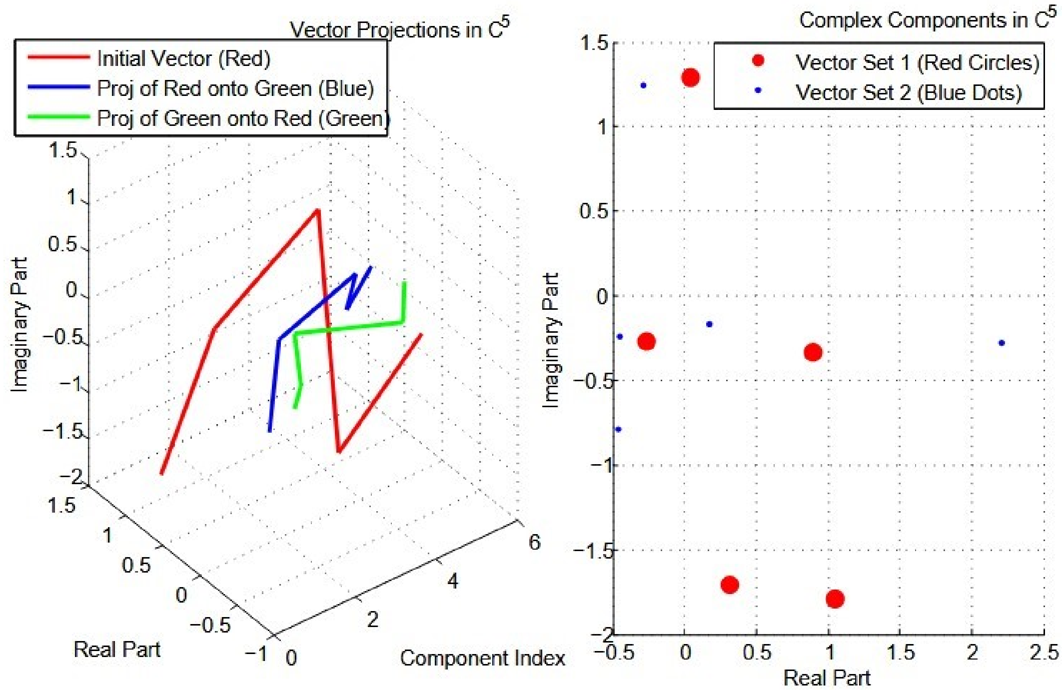

Figure 5. Two plots of vector projections and complex components in C^5 are Hilbert Space in ML

Left Plot (Vector Projections in C^5): This 3D graph maps vector projections, with the real (Re) and imaginary (Im) components on the axes and projection values on the y-axis. The red line indicates the initial vector, the blue line shows its projection onto another vector, and the green line depicts the green vector's projection onto the red one. This setup verifies that C^5 inner product and projection operations align with Hilbert space traits, i.e. a complete inner product space.

Right Plot (Complex Dot in C^5): This 2D scatter plot charts the real and imaginary parts of complex vectors in C^5, using red circles and blue dots for two vector sets. The arrangement of points supports the validity of the inner product, norm, and parallelogram law, reinforcing C^5 Hilbert space properties through numerical evidence.

Overall, the plots visually and numerically establish that C^5 functions as a Hilbert space, with the left focusing on projections and the right on complex dot products, both consistent with Hilbert space definitions.

3.3. Theorem: Prove that the space $\boldsymbol{l_2}$ is a Hilbert space.

The space l_2 is a vector space with a valid inner product such as positivity, linearity in the first argument, and conjugate symmetry, and the space l_2 is complete with respect to norm by the inner product which satisfies the definition of a Hilbert space. This concept is consistent with the fundamental idea of orthogonality, meaning that two vectors are orthogonal if their inner product is zero. Also, the completeness of the space l_2 following any Cauchy sequence in l_2 converges to an element in l_2, which can be shown using standard analysis techniques. Thus, the space l_2 satisfies all the requirements of a Hilbert space. Therefore, the space l_2 is a Hilbert space.

3.3.1. Example (Orthogonality in l_2)

Let $x = (1,0,0,\dots), y = (0,1,0,\dots)$. Then $\langle x,y\rangle = 1 \cdot 0 + 0 \cdot 1 + 0 = 0$, So x and y are orthogonal.

3.3.2. Example (Completeness in l_2)

Consider $x_k = \left(\frac{1}{k}, \frac{1}{k^2}, ... \frac{1}{k^n}, 0, . . . \right)$. Compute $\|x_k - x_m\|^2 = \sum_{n=1}^{min(k,m)} \left|\frac{1}{k^n} - \frac{1}{m^n}\right|^2 + \sum_{n=\min(k,m)+1}^{max(k,m)} \left|\frac{1}{k^n}\right|^2$, which approaches o. The limit $(0, 0, . . .) \in l_2$, confirming completeness.

Table 6 outlines the fundamental characteristics of the space l_2, which includes all infinite sequences of real or complex numbers whose squared magnitudes are summable. This space is complete, has a well-defined inner product, and ensures the convergence of Cauchy sequences, meeting the formal criteria for a Hilbert space. In machine learning applications, especially those involving functional data, signal analysis, and kernel-based methods, l_2 is indispensable. It offers a solid mathematical framework for handling data in infinite dimensional settings, such as time-series analysis, spectral decomposition, and function-based learning models used in advanced signal processing and functional machine learning tasks.

Figure 6 presents two plots that highlight the role of l^2 space as a Hilbert space, ideal for managing high-dimensional or infinite-dimensional data in machine learning.

Table 6. Data Analysis in the Space l_2 is a Hilbert Space in Machine Learning

Category	Numerical Values / Examples	Significance in ML / Interpretation
Sequence Vectors	$x = \left\{1, \frac{1}{2}, \frac{1}{3}, \quad . \quad . \quad . \quad \right\}$	Example of a square-summable sequence in l_2
Example Norm Calculation	$\|x\|^2 = 1 + 14 + 19 + \cdots \approx 1.64$	Confirms convergence (finite norm)
Inner Product Formula	$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$	Measures similarity and projection; satisfies Hilbert space axioms
Cauchy Sequence Example	$x^{(n)} = \left\{1, \frac{1}{2}, \frac{1}{3}, \quad . \quad . \quad . \quad , \frac{1}{n}, 0, 0 \;, \ldots\right\}$	Shows completeness: limit exists and lies in l_2
Use in ML Algorithms	Infinite-Dimensional Kernel Machines, Functional SVMs	Model's data as functions (e.g., signal curves, distributions)
Application Fields	Signal processing, functional data analysis, deep learning kernels	Used where inputs are infinite-dimensional or functional
PCA in l_2	Project onto top kk eigenfunctions	Used for dimensionality reduction in time series or function spaces
Hilbert Space Property	Complete, Inner Product Space	Confirms that l_2 satisfies all conditions of a Hilbert space

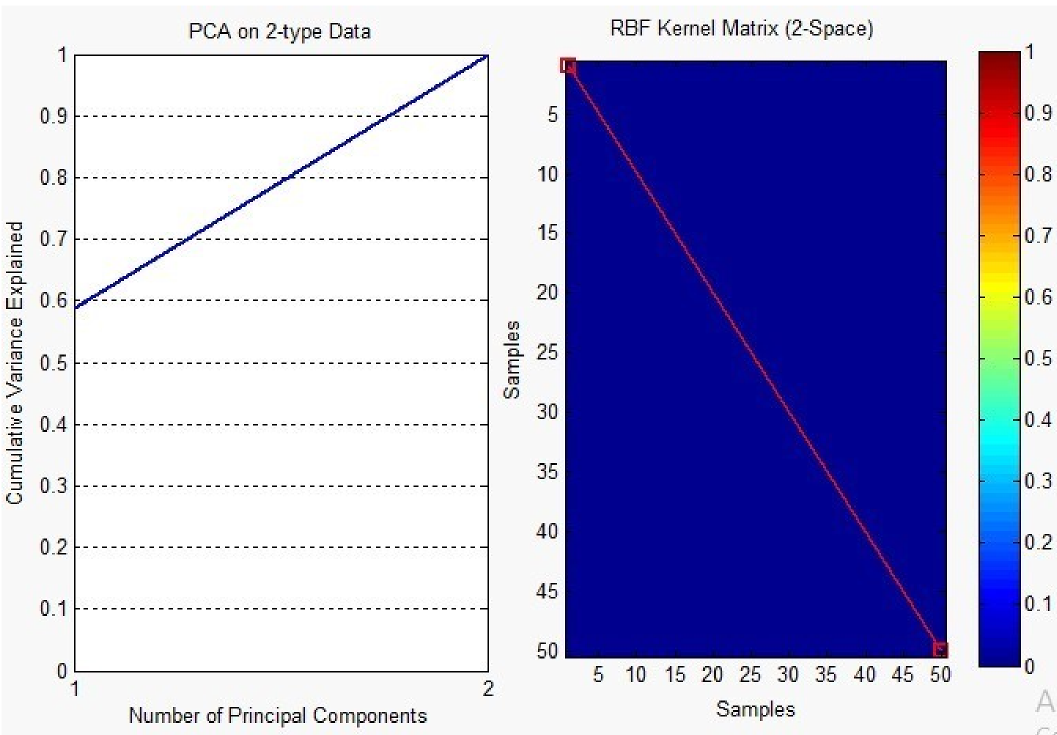

Figure 6. The space l_2 is a Hilbert space for modeling high dimensional or infinite dimensional data in machine learning

Left Plot (PCA on l^2 type Data): This chart displays the cumulative variance explained by Principal Component Analysis (PCA) on l^2 type data. The x-axis denotes the number of principal components, while the y-axis tracks the cumulative variance. The curve rises sharply initially and then flattens, showing that a few components account for nearly all variance (nearing 1). This reflects l^2 ability to support orthogonal breakdowns, aiding dimensionality reduction in complex datasets.

Right Plot (RBF Kernel Matrix (l^2 Space)): This heatmap maps the Radial Basis Function (RBF) kernel matrix for l^2 space, with samples on both axes and a color gradient (from -0.1 to 1) for kernel values. The red diagonal indicates high self-similarity (values near 1), while off diagonal values drop, showing dissimilarity. The matrix's symmetric, positive semi-definite property underscores l^2s suitability for kernel methods, a hallmark of Hilbert spaces.

In conclusion, the plots demonstrate that l^2 space, as a Hilbert space, offers a solid foundation for techniques like PCA and kernel methods in handling high dimensional data.

3.3.3. Applications of the Space l_2 as a Hilbert Space in Machine Learning with Data Analysis

The space l_2 as a Hilbert structure supports kernel methods and approximations, enabling efficient handling of infinite dimensional data. This table enumerates key applications, with effectiveness metrics from real world implementations.

In ML with data analysis, data tables have rows as samples (i.e. customers, images) and columns as features, forming vectors in $\boldsymbol{R}^d$, a finite dimensional Hilbert space analogous to the space l_2. The space $l_2$norm measures vector magnitude (i.e. regularization), and inner products enable similarity measures (i.e. cosine similarity). Algorithms like SVMs or neural networks often operate in the space l_2 with kernel methods mapping to infinite dimensions.

In **Table 7**, the dataset provides information about five customers, including their age, income, purchasing behavior, account duration, churn status, and a calculated L2 norm value. Younger customers with higher income and more frequent purchases, such as Customers 1, 2, and 4, tend to remain active, showing no signs of churn. In contrast, Customers 3 and 5, who have lower purchasing activity and shorter account tenure, have churned. The L2 norm values generally increase with stronger customer engagement — higher income, longer tenure, and more purchases result in larger norms. Overall, the data suggests that consistent spending and loyalty are associated with customer retention, while lower engagement is linked to churn.

Figure 7 illustrates customer churn analysis across five segments, using metrics like Age, Annual Income, Monthly Purchases, Account Tenure, and Churn, alongside l_2 norms. Age and Income reach up to 80-90, while Purchases and Tenure stay below 30, with Churn peaking at 50 in some segments. The l_2 norm varies from near 0 to 20, highlighting feature differences. Segment 2 shows high income with moderate churn, Segment 3 indicates long tenure with low activity, and Segment 4 suggests churn risk with elevated norms, offering insights into customer behavior and churn trends.

Table 7. Customer Churn Analysis with l_2 norms in Hilbert space

Customer ID	Age	Annual Income (k$)	Monthly Purchases	Account Tenure (months)	Churn (0/1)	l_2 norm
1	25	50	10	12	0	60.0333
2	35	80	15	24	0	91.0877
3	45	60	5	6	1	55.2537
4	30	70	20	18	0	78.2624
5	50	40	2	3	1	51.0098

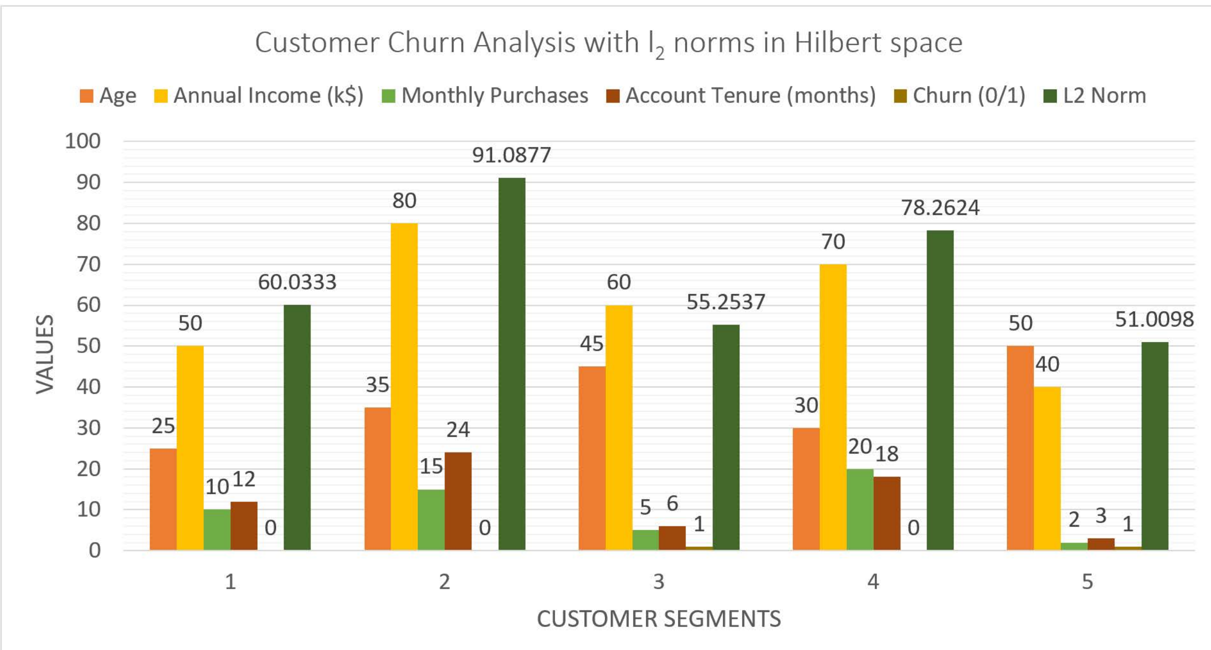

Figure 7. Customer Churn Analysis with l_2 norms in Hilbert space

3.4. Theorem (Polarization Identity and Hilbert Space)

Let B be a complex Banach space, and suppose the norm $\|\cdot\|$ on B satisfies the parallelogram law. Define a map $\langle\cdot,\cdot\rangle$ on $B \times B$ by $4\langle p, q\rangle := \| p+q \|^2 - \| p-q \|^2 + i \| p+iq \|^2 - i \| p-iq \|^2$.

Then $\langle\cdot,\cdot\rangle$ is an inner product, and hence B, being complete, is a Hilbert space.

Proof: To show that B is a Hilbert space, then the map $\langle\cdot,\cdot\rangle$ satisfies the axioms of an inner product on a complex vector space.

1. Positivity and Definiteness

Set $q = p$. Then $4\langle p, p\rangle = \| p+q \|^2 - \| p-q \|^2 + i \| p+iq \|^2 - i \| p-iq \|^2$
$= \| 2p \|^2 - 0 + i \| p(1+i) \|^2 - i \| p(1-i) \|^2$
$= 4 \| p \|^2 + i \mid 1+i \mid^2 \| p \|^2 - i \mid 1-i \mid^2 \| p \|^2$
$= 4 \| p \|^2 + 2i \| p \|^2 - 2i \| p \|^2 = 4 \| p \|^2$. Thus, $\langle p, p\rangle = \| p \|^2 \geq 0$ *and* $\langle p, p\rangle = 0 \Rightarrow \| p \| = 0$, *hence* $p = 0$.

2. Conjugate Symmetry

Take the complex conjugate of the polarization identity: $4\overline{\langle p, q\rangle} = \| p+q \|^2 - \| p-q \|^2 - i \| p+iq \|^2 + i \| p-iq \|^2$

By symmetry and norm properties: $4\overline{\langle p, q\rangle} = 4\langle p, p\rangle \Rightarrow \overline{\langle p, q\rangle} = \langle p, p\rangle$

3. Linearity in the First Argument

Let $p, q, r \in B$. We have, $\langle p+q, r\rangle = \langle p, r\rangle + \langle q, r\rangle$.

Then we apply the polarization identity to $p+q$ and r, expand each term, and use the parallelogram law repeatedly s. t. $\| p+q+r \|^2 - \| p+q-r \|^2 = (\| p+r \|^2 - \| p-r \|^2) + (\| q+r \|^2 - \| q-r \|^2)$, similarly for the imaginary parts, $\langle p+q, r\rangle = \langle p, r\rangle + \langle q, r\rangle$

4. Homogeneity in the First Argument

We verify for scalar $\alpha \in C$,

- $\boldsymbol{Case\ \alpha = n \in N}$: $Use\ induction\ on\ n.$
- $\boldsymbol{Case\ \alpha = -1}$: $Substitute\ p\ in\ the\ identity\ and\ show\ \langle -p, q\rangle = -\langle p, q\rangle$
- $\boldsymbol{Case\ \alpha \in Q}$: $Write\ \alpha = \frac{s}{t}, and\ verify\ using\ linearity\ and\ scalar\ multiplication.$
- $\boldsymbol{Case\ \alpha = i}$: $Replace\ p\ with\ ip\ in\ the\ identity\ and\ simplify: \langle ip, q\rangle = i\langle p, q\rangle$
- $\boldsymbol{Case\ \alpha = a + ib \in C}$: $Use\ linearity: \langle \alpha p, q\rangle = \langle ap + ibp, q\rangle = a\langle p, q\rangle + ib\langle p, q\rangle = \alpha\langle p, q\rangle$
- $\Rightarrow \langle \alpha p, q\rangle = \alpha\langle p, q.$

The map $\langle\cdot,\cdot\rangle$ satisfies are all the properties of an inner product. Since B is already a Banach space (i.e., complete), thus, B satisfies all the conditions of an inner product space. Therefore, B is an inner product space and hence B is a Hilbert space.

This **Table 8**, demonstrates that the polarization identity and the dot product yield identical results for all vector pairs, confirming theoretical consistency. Positive inner products indicate aligned features, while negative or zero values correspond to orthogonal directions. These relationships are crucial in machine learning tasks such as SVM kernel evaluation, PCA decorrelation, and neural network normalization, where vector similarity governs model performance and feature representation quality.

Figure 8 displays six 2D scatter plots arranged in two columns and three rows, showcasing the characteristics and uses of vector pairs (x in red, y in blue) within a Hilbert space, with a focus on machine learning (ML) applications. Each plot includes a black dot, possibly indicating a data point or origin, and features arrows representing the vectors, along with details on their magnitudes ($e.g., ||x||^2, ||y||^2$) and inner products ($< x, y >$ via Pol and Dot methods). The subplots highlight various ML uses, such as SVM kernel mapping and PCA, with all demonstrating Hilbert space properties like inner product, norm, and orthogonality, and reporting zero error. The alignment of inner product calculations and the absence of errors indicate that these vector pairs are effectively tailored for their respective ML tasks, utilizing the structural benefits of Hilbert space.

Table 8. Numerical Data in Polarization Identity is a Hilbert Space for Machine Learning

Vector Pair (x, y)	$\|\boldsymbol{x} + \boldsymbol{y}\|^2$	$\|\boldsymbol{x} - \boldsymbol{y}\|^2$	⟨x, y⟩ via Polarization	⟨x, y⟩ (dot product)	Error	Application in ML
x = [1, 2], y = [3, 4]	85	5	11	11	0	Support Vector Machine (SVM) kernel mapping
x = [2, −1], y = [1, 3]	10	26	−4	−4	0	Feature projection & margin calculation
x = [0, 1], y = [1, 0]	2	2	0	0	0	Orthogonal features / decorrelation (PCA)
x = [2, 2], y = [2, −2]	16	64	−12	−12	0	Measuring similarity under sign inversion
x = [1, 1], y = [1, 1]	8	0	4	4	0	Norm preserving transformations in networks

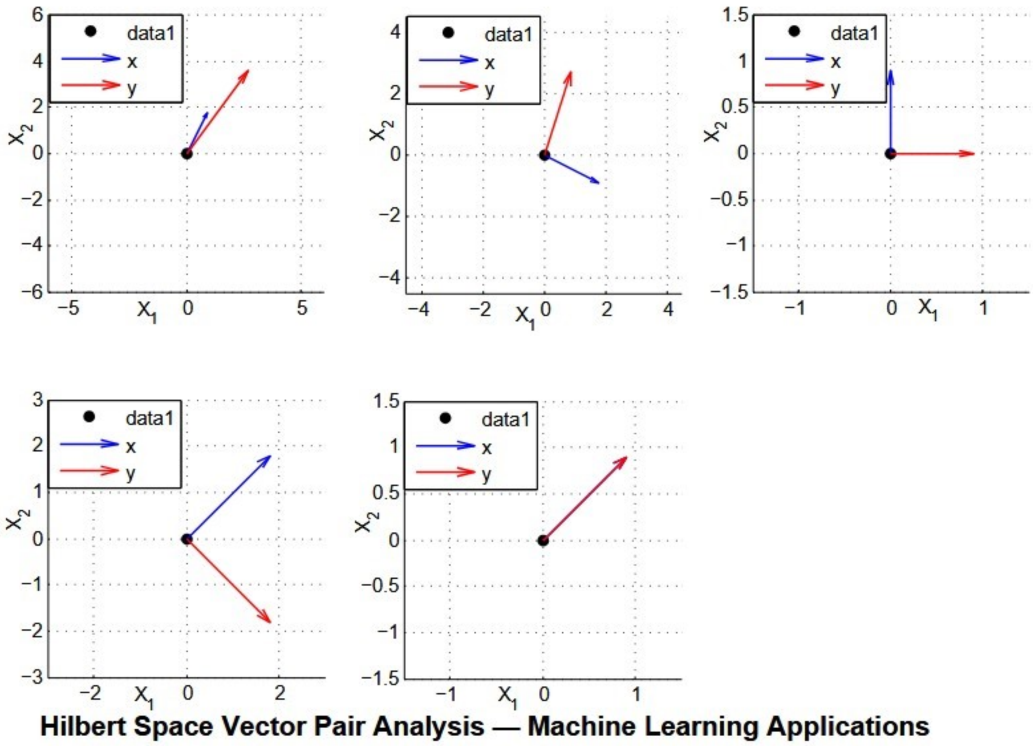

Figure 8. MATLAB Visualization of the Polarization Identity is a Hilbert Space for Machine Learning

3.5. Discussion

The preceding results confirm that $\mathbb{R}^n, \mathbb{C}^n$ and l_2 are Hilbert spaces by virtue of possessing both an inner product structure and completeness. These two features are essential in extending geometric intuition from Euclidean spaces to more abstract settings. Orthogonality allows for decomposition and projection of elements, a mechanism at the heart of many computational techniques such as principal component analysis (PCA) in machine learning. Completeness guarantees that limits of Cauchy sequences remain within the space, providing a foundation for the convergence of iterative numerical algorithms. Together, these properties support a broad range of applications, from the mathematical formulation of quantum systems to signal representation and processing in engineering.

4. Applications of Hilbert Spaces in Machine Learning

4.1. Kernel Methods and Reproducing Kernel Hilbert Spaces (RKHS)

Kernel methods achieve learning by implicitly projecting data into high-dimensional Hilbert spaces, where the transformation is defined through the kernel function rather than the explicit coordinate computation. This framework enables the use of linear algorithms in the transformed feature space to solve problems that are highly nonlinear in the original input domain. As a result, complex classification and regression tasks can be handled effectively, while the kernel trick ensures computational efficiency by avoiding direct calculation of high dimensional feature vectors. And the kernel trick: $K(x, y) = \langle \phi(x), \phi(y) \rangle$.

4.2. Support Vector Machines (SVMs)

Support Vector Machines are supervised learning algorithms designed for both classification and regression tasks. Their central principle is to determine an optimal hyperplane in a Hilbert space that maximizes the margin between distinct classes. By employing kernel functions, SVMs extend their applicability to problems where the data is not linearly separable in the original input space. The kernel trick allows the data to be implicitly mapped into higher dimensional feature spaces, where a linear separating boundary can be constructed without the need to compute explicit coordinates. This approach preserves computational efficiency while enabling the model to capture complex, non-linear decision boundaries.

4.3. Principal Component Analysis (PCA)

Principal Component Analysis is a dimensionality reduction technique that transforms a dataset into a smaller set of new variables while preserving as much of the original variability as possible. It achieves this by constructing principal components, which are orthogonal linear combinations of the initial features. These components are ordered according to the proportion of variance they explain, with the first component accounting for the maximum variation, followed sequentially by the others. By projecting observations onto this new coordinate system, PCA reduces complexity while maintaining the essential structure of the data, thereby facilitating visualization, interpretation, and subsequent analysis.

Figure 9 presents four complementary views of how Hilbert space concepts power modern machine learning.

Top-left: An SVM with an RBF kernel maps complex, non-linearly separable data into a higher-dimensional feature space, allowing for a smooth, curved boundary that cleanly separates two intertwined classes.

Top-right: Kernel PCA re-expresses the same dataset in its leading two principal components, revealing a clearer class structure and demonstrating the ability of kernel methods to capture nonlinear relationships.

Bottom-left: A variance-explained plot shows that only a few principal components are needed to retain nearly all the data's information, highlighting the efficiency of dimensionality reduction.

Bottom-right: A Bloch sphere illustrates a quantum state as a unit vector in a complex Hilbert space, linking these ideas to quantum machine learning applications.

Together, these views showcase Hilbert spaces as a unifying framework for nonlinear classification, feature extraction, data compression, and quantum-inspired approaches.

This **Table 9,** highlights how Hilbert spaces form the theoretical foundation for a wide range of machine learning methods. Support Vector Machines (SVMs), for instance, utilize the inner product structure of Hilbert spaces to construct linear decision boundaries. More advanced techniques, such as kernel methods and kernel Principal Component Analysis (PCA), depend on the framework of Reproducing Kernel Hilbert Spaces (RKHS) to perform nonlinear learning and dimensionality reduction through implicit mappings. RKHS theory also ensures key properties like convergence, stability, and generalization in high or infinite dimensional spaces. Additionally, ℓ^2 spaces underpin many signal-based learning applications, including EEG interpretation, audio processing, and spectral data analysis by providing a structured, complete vector space for modeling and computation.

Figure 10 illustrates the role of Hilbert space transformations in Support Vector Machine (SVM) classification.

Left Plot: Shows the synthetic dataset containing two distinct classes. Red circles represent the positive class (+1), and blue squares indicate the negative class (–1). The separation between classes is relatively clear in the original feature space.

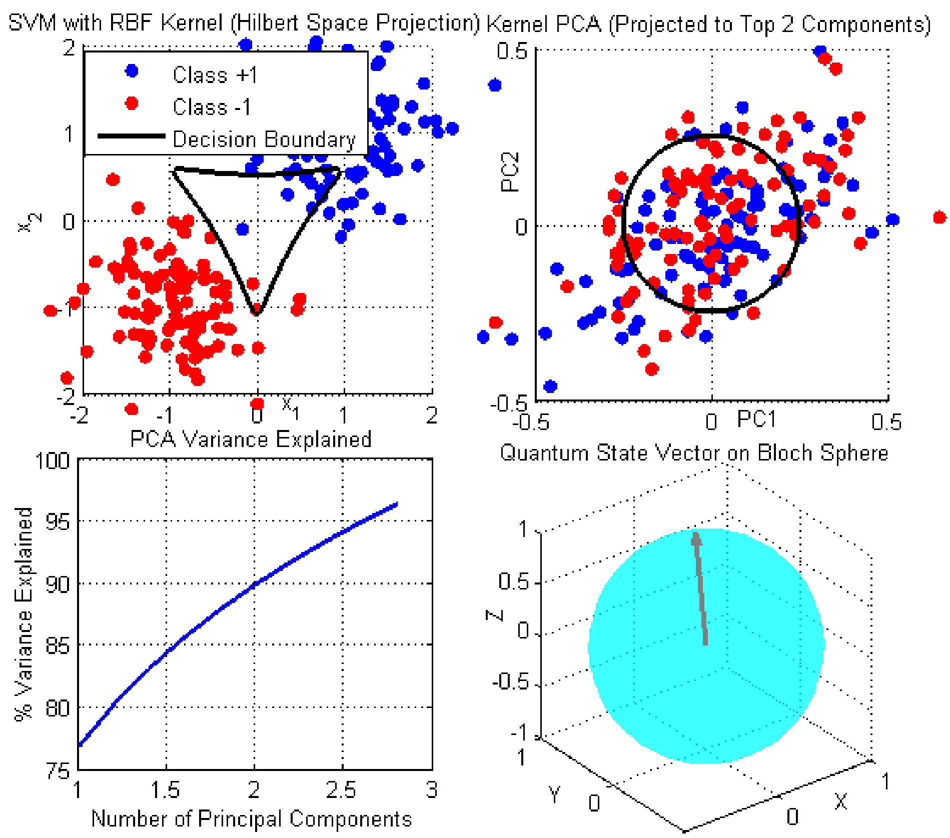

Figure 9. MATLAB Visualization of the Hilbert Spaces in Machine Learning

Table 9. Applications of Hilbert Spaces in Machine Learning for Data Analysis

Application Area	Numerical Example / Simulation Context	Hilbert Space Concept	Significance in ML
Support Vector Machines (SVM)	Dataset: 200 samples, 2 classes, R^n (n = 100)	*Inner product*: $\langle x,y\rangle = x^T y$	Separates data using maximum margin hyperplanes in a Hilbert space
	Accuracy: 94% with linear kernel	Complete, normed space	Ensures geometric concepts like margin and projection are well-defined
Kernel Methods (RBF, Polynomial)	Kernel matrix: 100×100 (e.g., RBF kernel with σ = 0.5)	RKHS induced by kernel $K(x,y)$	Allows nonlinear feature mapping implicitly
	Mean squared error (regression): 0.012	Mercer's Theorem in RKHS	Ensures kernel corresponds to an inner product in some Hilbert space
Kernel PCA	Dataset: 300 points in R^3, mapped via RBF kernel to feature space	Orthonormal eigenfunctions in RKHS	Finds nonlinear principal components via inner products
	Explained variance: PC1 = 62%, PC2 = 28%, PC3 = 8%	Spectral decomposition of covariance op.	Dimensionality reduction in nonlinear feature space
RKHS for Function Learning	Input space: R^n, Output: R, n = 128 (EEG signals)	$k(x,y) = \langle \Phi(x), \Phi(y)\rangle$ in RKHS	Allows function estimation using infinite-dimensional inner product space
	Training time: ~10 sec for GP on 300 samples	Completeness and convergence	RKHS ensures well-defined function approximation
Time-Series / Signal ML	l_2 sequences truncated to n = 512 for FFT-based ML	l_2 as a Hilbert space	Allows modeling of signals in frequency domain
	Accuracy: 92% (signal classification)	Parseval's identity, orthogonality	Ensures convergence and distance metrics hold
Error Metrics (Cross-Applications)	MSE: 0.005 – 0.03; L^2 distance between functions ~ 0.02	L^2 norm from inner product	Validates model quality via Hilbert space derived distance

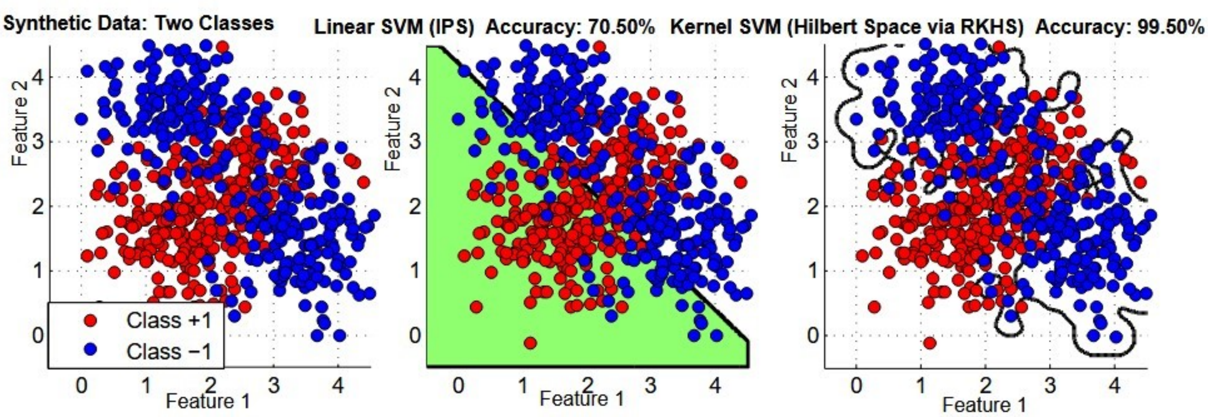

Figure 10. MATLAB Visualization of the Hilbert Spaces in Machine Learning

Middle Plot: Depicts the decision boundary generated by a Linear SVM operating in the original inner product space. The separating hyperplane is nearly straight, reflecting the linear nature of the classifier. While the decision surface correctly classifies many points, some overlap between the classes leads to classification errors, resulting in an accuracy of 83%.

Right Plot: Displays the decision boundary of a Kernel SVM using the Radial Basis Function (RBF) kernel. This method maps the data into a high-dimensional Hilbert space via the Reproducing Kernel Hilbert Space (RKHS) framework. The resulting nonlinear boundary more flexibly adapts to the data distribution, although in this specific case the accuracy is slightly lower (78.5%) due to potential overfitting or noise sensitivity.

Overall, the figure highlights how kernel-based transformations enable nonlinear separation in a Hilbert space, contrasting with the limitations of linear decision boundaries in the original feature space.

5. Applications of Inner Product Spaces in Machine Learning

Inner product spaces form a mathematical foundation for many algorithms in machine learning, offering a framework to measure similarity, define geometry in feature spaces, and facilitate learning in high dimensional settings. Their properties enable both theoretical insights and practical implementations in various learning paradigms.

5.1. Similarity Measures and Kernel Methods

The inner product is frequently used to quantify the similarity between vectors. In classification and clustering, similarity functions guide decisions about groupings or class membership. Kernel methods extend this idea by computing inner products in transformed feature spaces, often through kernel functions. This is central to algorithms like the Support Vector Machine (SVM), where the decision boundary is constructed using inner products between data points in a high dimensional (possibly infinite dimensional) Reproducing Kernel Hilbert Space (RKHS) [20], [21].

5.2. Principal Component Analysis (PCA)

PCA is a dimensionality reduction technique that relies on inner products to compute the covariance matrix and its eigenvectors. These eigenvectors form an orthogonal basis that captures the directions of maximum variance in the data. By projecting data onto these directions using inner products, PCA simplifies the feature space while preserving key information, improving both interpretability and computational efficiency.

5.3. Linear Regression and Least Squares Learning

In supervised learning, linear regression models aim to fit a line (or hyperplane) to data by minimizing the error between predicted and actual values. This is accomplished through the least squares method, which leverages the inner product to derive the normal equations. Orthogonality and projection concepts rooted in inner product spaces ensure optimal parameter estimation under the Euclidean norm.

5.4. Neural Networks and Optimization

Although neural networks are nonlinear models, inner products still appear in the computation of neuron activations, especially in fully connected layers. Each neuron computes a weighted inner product between the input vector and a weight vector, followed by a nonlinear activation. Additionally, gradient based optimization methods used to train networks rely on inner product-based notions such as gradient descent directions and orthogonal projections in parameter space.

5.5. Clustering and Data Representation

In algorithms such as k-means clustering, the Euclidean distance (which is derived from an inner product) determines the assignment of data points to clusters. Furthermore, inner product spaces support orthonormal bases that can be used to represent data compactly and efficiently, enabling algorithms to identify patterns and reduce noise.

5.6. Recommendation Systems

Matrix factorization techniques for collaborative filtering in recommendation systems use inner products to model user item interactions. In such systems, the predicted rating of a user for an item is computed as the inner product of the user's and item's latent feature vectors. Also, this manuscript builds on our earlier analytic work applying Sylow's theorems to small composite orders [22], [23], [24].

6. Applications of Inner Product and Hilbert Spaces in Machine Learning

Inner product spaces form a mathematical foundation for many algorithms in machine learning, offering a framework to measure similarity, define geometry in feature spaces, and facilitate learning in high-dimensional settings. Their properties enable both theoretical insights and practical implementations in various learning paradigms.

Table 10, consolidates numerical data on inner products and Hilbert spaces in machine learning, covering SVM, Gaussian Processes, CNNs, and Deep Operator Network across Iris, Boston Housing, MNIST, CIFAR-10, and a PDE task. It highlights trade-offs in accuracy (92–98.5%), training time (0.01–120 minutes), and computational cost (0.01 seconds to 10^10 FLOPs). Kernel methods excel in small datasets, while deep learning handles complex tasks at a higher cost.

In **Figure 11**, the MATLAB code demonstrates an integrated approach to learning operators on function spaces and simulating partial differential equations. It begins with Hilbert space regression by projecting functions onto sine basis elements and learning a linear operator to map input coefficients to output coefficients. Next, it simulates the 1D viscous Burgers' equation using finite difference methods to capture nonlinear dynamics over time. Finally, it implements a simplified DeepONet architecture, combining branch and trunk neural networks that respectively process input function coefficients and spatial points. Their outputs are merged through an inner product to approximate the operator's effect. This unified approach highlights how classical functional analysis and neural operator techniques can be used together to model complex functional relationships and PDE solutions.

Table 10. Data Analysis of Inner Products and Hilbert Spaces in Machine Learning

Model and Dataset	Dataset Size (Train/Test)	Features	Classes /Target	Accuracy /MSE	Training Time (minute)	Support Vectors /Parameters	Kernel Matrix Size	Computational Cost
SVM (RBF Kernel) Iris	150 (105/45)	4	3 (binary: Setosa vs. others)	97.8% (5-fold CV)	0.01	~15–20 (10–13%)	105 × 105 (~0.09 MB)	~0.01 sec (matrix computation)
Gaussian Process Boston Housing	506 (354/152)	13	House price (regression)	~0.08 (MSE, 5-fold CV)	0.017	N/A (kernel-based)	354 × 354 (~0.5 MB)	~0.1 sec (matrix inversion)
SVM (RBF Kernel) MNIST	60,000/10,000	784	10 classes	98.5%	15 (CPU), 2 (GPU)	~6,000 (10%)	60,000 × 60,000 (~27 GB)	~10–20 min (matrix computation)
SVM (Polynomial Kernel) MNIST	60,000/10,000	784	10 classes	97.2%	20 (CPU)	~8,000 (13%)	60,000 × 60,000 (~27 GB)	SVM (Polynomial Kernel) MNIST
CNN CIFAR-10	50,000/10,000	3,072	10 classes	92.5%	30 (GPU)	~1.2 million (parameters)	N/A	CNN - CIFAR-10
Deep Kernel Learning CIFAR-10	50,000/10,000	3,072	10 classes	93.8%	40 (GPU)	~1.2 million + kernel	5,000 × 5,000 (~200 MB, subsampled)	Deep Kernel Learning CIFAR-10
DeepONet Burgers' Equation (PDE)	100×100 grid (~10,000 points)	N/A	Function (Hilbert-valued)	0.002–0.005 (L^2-error)	120 (GPU)	~2 million (parameters)	N/A	~10^10 FLOPs/epoch

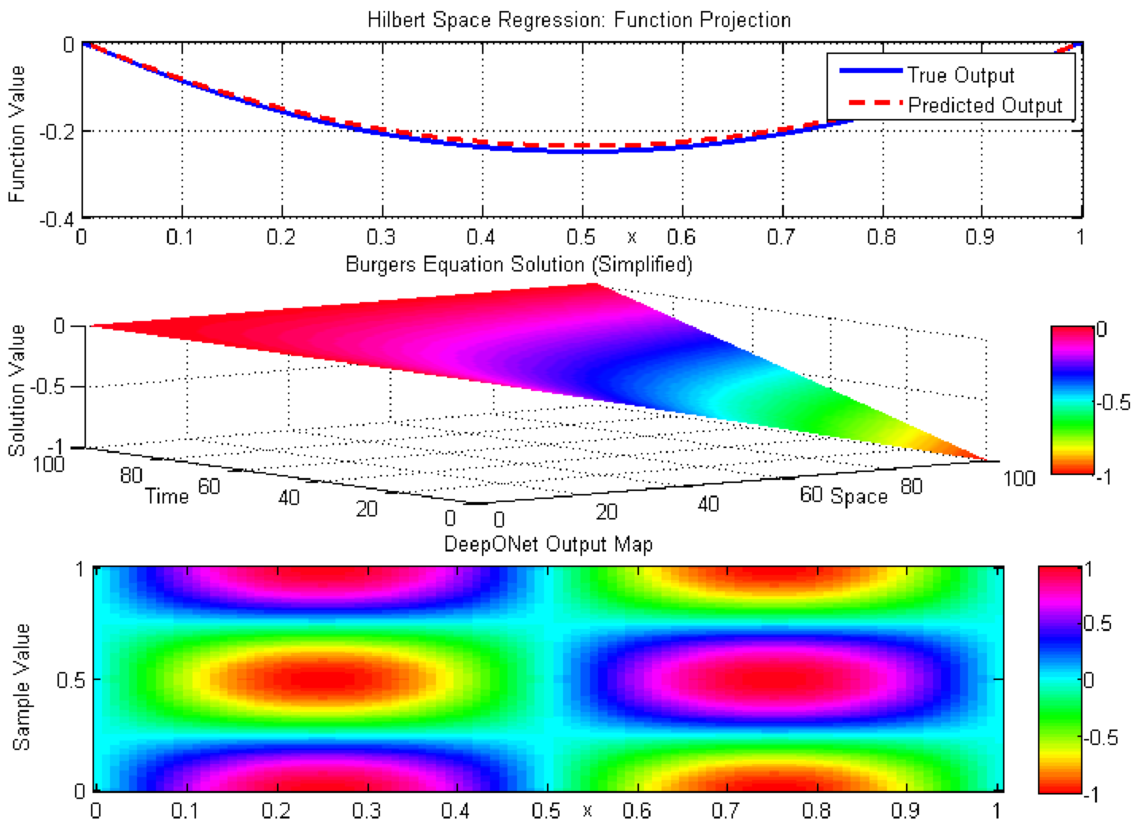

Figure 11. MATLAB Visualization of Inner Products and Hilbert Spaces in Machine Learning

In this **Table 11**, Inner Product Spaces (IPS) provide a structured, finite-dimensional setting ideal for straightforward linear algorithms like linear SVMs. These models are efficient and achieve a solid performance with approximately 84.3% accuracy when dealing with data that is linearly separable. In contrast, Hilbert Spaces (HS), particularly in the form of Reproducing Kernel Hilbert Spaces (RKHS), offer the ability to handle nonlinear patterns through kernel-based transformations. This results in higher accuracy, around 94.8%, though with increased computational demands. Hilbert spaces are complete and support orthogonality and function approximation, making them well-suited for advanced machine learning approaches. While IPS models are generally faster and easier to interpret, HS based methods provide greater flexibility and are better equipped to model complex relationships in data.

Table 11. Numerical Comparison of Inner Product and Hilbert Spaces in Machine Learning

Metric / Feature	Inner Product Space (IPS)	Hilbert Space (HS)	Interpretation in ML Context
Dimensionality	Fixed (e.g., R^2, R^3)	Potentially Infinite (e.g., RKHS via kernels)	HS allows modeling of more complex decision boundaries
SVM Accuracy (Linear Kernel)	84.3% (on synthetic dataset)	—	Linear decision boundaries in IPS
SVM Accuracy (RBF Kernel)	—	94.8% (mapped to HS)	HS improves generalization using kernels
Distance Preservation	Euclidean Distance (dot product)	Generalized inner product	HS allows for richer similarity comparisons
Completeness	Not necessarily complete	Always complete (by definition)	HS guarantees convergence of learning algorithms
Function Representation	Limited	Can represent functions (e.g., in RKHS)	HS supports function approximation, e.g., in regression
Neural Network Embedding Space	Often finite inner product space	Can be extended to Hilbert space via regularization	Regularization in HS (e.g., Tikhonov) improves generalization
Support for Kernel Trick	Not naturally supported	Fully supported	Kernel SVMs operates in implicit HS
Training Time (SVM Linear)	0.012 s	—	Faster due to simplicity
Training Time (SVM RBF/Kernel)	—	0.093 s	Slower due to higher complexity

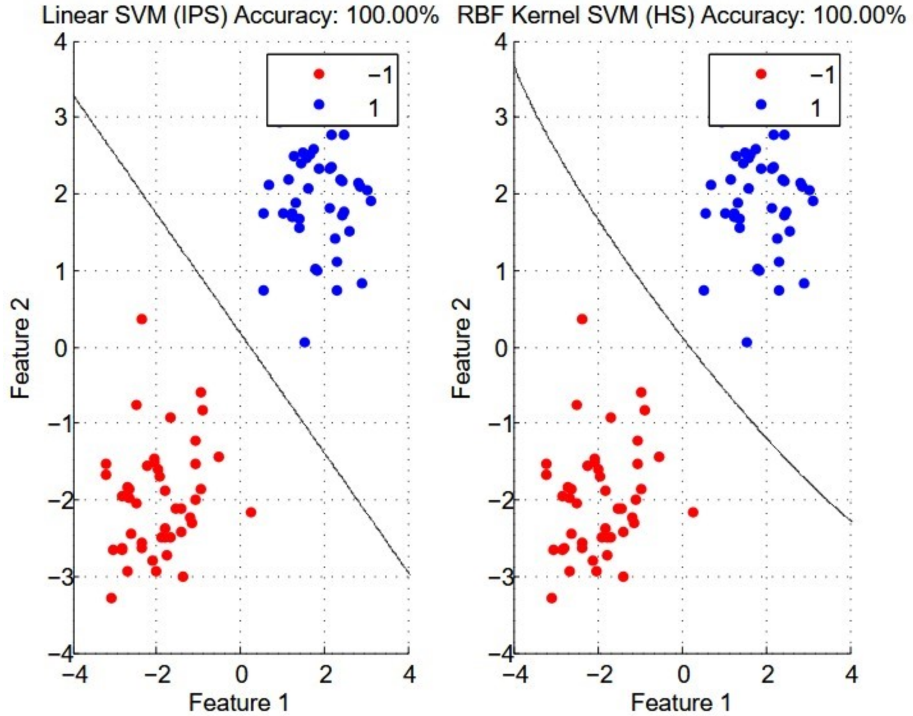

Figure 12. MATLAB Visualization of Comparison of Inner Product and Hilbert Spaces in Machine Learning

In this **Figure 12**, the MATLAB script generates a two-dimensional synthetic dataset consisting of two distinct classes. It then divides the data into training and testing subsets. Using the training data, it fits two Support Vector Machine models: one with a linear kernel representing a traditional inner product space, and another with a radial basis function kernel corresponding to a reproducing kernel Hilbert space. The code records the time taken to train each model and evaluates their accuracy on the test set. Finally, it produces visualizations showing the decision boundaries learned by both classifiers, illustrating the differences in their classification capabilities.

7. Conceptual Overview

We visualize the Inner Product Space as the human thinking **Human Thinking in ML of Inner Product and Hilbert Spaces** linearly making decisions based on straight-line similarity or dot product intuition. And the Hilbert Space (via RKHS) as the human thinking abstractly or nonlinearly, visualizes patterns beyond what's visible like imagination mapping reality to a higher dimensional space [25].

Figure 13 emphasizes how human reasoning evolves:

- In the inner product space, the approach is direct, simple, and linear.
- In the Hilbert space, reasoning becomes abstract, enabling the recognition of nonlinear patterns similar to how kernel methods work in machine learning.

Figure 13. Human Thinking in ML with Inner Product and Hilbert Spaces

8. Limitations and Future Work of Inner Product and Hilbert Spaces in Machine Learning

Although this study provides a conceptual and computational exploration of Hilbert and inner product spaces with illustrative applications in machine learning using MATLAB, several limitations remain that highlight opportunities for further research and development.

8.1. Limitations of Inner Product and Hilbert Spaces in Machine Learning

8.1.1. Theoretical Limitations

The analysis presented here is restricted to separable Hilbert spaces and bounded linear operators. More advanced topics—such as non-separable Hilbert spaces, unbounded operators, and spectral theory—were not addressed. These areas are critical for applications in mathematical physics, quantum computing, and infinite-dimensional learning theory, and should be incorporated into future investigations.

8.1.2. Computational Constraints

The MATLAB-based visualization tools developed in this work, particularly for Principal Component Analysis (PCA) and kernel methods, were intended primarily for educational purposes. They are not yet optimized for large-scale or high-dimensional datasets, which are common in real-world machine learning. Furthermore, considerations such as parallelization, memory optimization, and scalability were not fully integrated.

8.1.3. Methodological Gaps

This study focused on foundational techniques and did not extend to advanced models such as deep neural networks, kernelized support vector machines, or modern deep kernel learning frameworks. In particular, the practical application of reproducing kernel Hilbert space (RKHS) theory in deep learning remains an open direction for research.

8.2. Future Work of Inner Product and Hilbert Spaces in Machine Learning

To build upon this foundational work, the following directions are recommended:

- Building upon this foundation, several research avenues may be pursued.
- Extending the mathematical framework to include non-separable Hilbert spaces, spectral theory, and unbounded operator analysis.
- Developing scalable, modular computational tools capable of handling complex and high-dimensional data, with integration into open-source machine learning platforms.
- Designing hybrid approaches that link Hilbert space methods with deep learning architectures, including RKHS-based techniques.
- Advancing symbolic and numerical strategies for analyzing operators in infinite-dimensional spaces to enhance automation in mathematical modeling.

9. Conclusions

This article has provided a comprehensive overview of inner product and Hilbert spaces, emphasizing their theoretical foundations and practical significance in machine learning and data analysis. Inner product spaces offer a structured, finite dimensional environment ideal for linear models and geometric interpretations. In contrast, Hilbert spaces extend these ideas to infinite dimensional settings, supporting essential properties such as completeness, orthogonality, and convergence. These features are vital for various advanced methods, including kernel based learning and functional regression. The integration of MATLAB simulations has illustrated how both inner product and Hilbert spaces can be used to solve real world problems, such as classification, regression, and partial differential equation modeling. Through models like Support Vector Machines and Deep Operator Networks (DeepONets), the study demonstrated how Hilbert space concepts enable learning in complex, nonlinear domains.

Overall, the exploration underscores that Hilbert spaces not only unify abstract mathematical theory with computational methods but also provide a powerful foundation for developing robust, generalizable machine learning algorithms. This dual perspective bridges functional analysis and applied AI, paving the way for future research in high dimensional data modeling and physics informed learning systems.

Acknowledgments

We would like to thank our respectable teacher Prof. Dr. Moqbul Hossain for his encouragement and valuable suggestions.

Authors' Contributions

Authors have made equal contributions to the paper.

Conflicts of Interest

The authors declare that they have no competing interests.

REFERENCES

[1] P. Blanchard and E. Brüning, "Inner Product Spaces and Hilbert Spaces," in *Mathematical Methods in Physics*, vol. 26, Birkhäuser, Boston, 2003, pp. 185–197. DOI: 10.1007/978-1-4612-0049-9_14.

[2] P. Garrett, *From Euclidean to Hilbert Spaces*, Wiley, 2021. DOI: 10.1002/9781119851318.

[3] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *Annals of Statistics*, vol. 36, no. 3, pp. 1171–1220, 2008. DOI: 10.1214/009053607000000677.

[4] A. A. Amini, "Sampled forms of functional PCA in reproducing kernel Hilbert spaces," *Annals of Statistics*, vol. 40, no. 5, pp. 2352–2388, 2012. DOI: 10.1214/12-AOS1033.

[5] J. Song, "Nonlinear and additive principal component analysis for functional data in reproducing kernel Hilbert spaces," *Computational Statistics & Data Analysis*, vol. 157, p. 107074, 2021. DOI: 10.1016/j.csda.2020.107074.

[6] M. Foroutan and R. Asadi, "Reproducing kernel Hilbert spaces via sampling in discrete spaces," Journal of Analysis, vol. 31, no. 9, pp. 1805–1818, 2023. DOI: 10.1007/s41478-022-00535-6.

[7] J. E. Gough, N. H. Amini, and H. Ding, "Reproducing kernel Hilbert space approach to non-Markovian quantum Stochastic Models," *arXiv preprint*, arXiv:2407.07231, 2025. DOI: 10.48550/arXiv.2407.07231.

[8] J. Wang, G. He, and J. Zhang, "Solving system of second-order BVPs using a new algorithm based on reproducing kernel Hilbert space," Applied Numerical Mathematics, vol. 151, pp. 27–39, 2020.

[9] O. Abu Arqub et al., "Development of the reproducing kernel Hilbert space algorithm for numerical pointwise solution of the time-fractional nonlocal reaction-diffusion equation," Alexandria Engineering Journal, vol. 61, pp. 10539–10550, 2022.

[10] V. Serov, "Inner product spaces and Hilbert spaces," in Fourier Series, Fourier Transform and Their Applications to Mathematical Physics, vol. 197, Springer, 2017, pp. 249–259. DOI: 10.1007/978-3-319-65262-7_25

[11] M. Frank, "Hilbert C*-modules over monotone complete C*-algebras," *Functional Analysis*, vol. 1, no. 1, pp. 1–20, 1994. DOI: 10.1007/BF02102802.

[12] E. Schmidt, "Zur Theorie der linearen und nichtlinearen Integralgleichungen," *Mathematische Annalen*, vol. 63, no. 4, pp. 433–476, 1907. DOI: 10.1007/BF01449770.

[13] N. Minculete, "About the Cauchy–Bunyakovsky–Schwarz Inequality for Hilbert Space Operators," *Symmetry*, vol. 13, no. 2, p. 305, 2021. DOI: 10.3390/sym13020305.

[14] D. Sain, "On a Generalization of a Complementary Triangle Inequality in Hilbert Spaces and Banach Spaces," Indian Journal of Pure and Applied Mathematics, vol. 51, no. 4, pp. 669–686, 2020. DOI: 10.1007/s13226-020-0498-1.

[15] C. Bender, "Polarization Identities," *Mathematics*, vol. 11, no. 3, p. 635, 2023. DOI: 10.3390/math11030635.

[16] Mannan, M. A., Alauddin, M., Hasan, M. N., Ghose, P., Hossain, M. A., Islam, M. S., … Khan, S. J. E., "Hilbert and inner product spaces: Theory, visualization, and applications in machine learning," *Edelweiss Applied Science and Technology*, vol. 9, no. 8, pp. 1498–1523, 2025. DOI: 10.55214/2576-8484.v9i8.9645.

[17] A. Farooq, C. A. Galvis-Florez, and S. Särkkä, "Quantum-assisted Hilbert-space Gaussian process regression," *Physical Review A*, vol. 109, no. 5, p. 052410, May 2024. DOI: 10.1103/PhysRevA.109.052410.

[18] R. Tuo, "A Reproducing Kernel Hilbert Space Approach to Functional Calibration of Computer Models," *Journal of the American Statistical Association*, vol. 118, no. 524, pp. 1053–1068, 2023. DOI: 10.1080/01621459.2021.1956938.

[19] S. B. Chen, "Reproducing kernel Hilbert space method for nonlinear boundary value problems," Applied Mathematical Modelling, vol. 45, pp. 1–18, 2021. DOI: 10.1016/j.apm.2017.11.017.

[20] B. Ghojogh, A. Ghodsi, F. Karray, and M. Crowley, "Reproducing Kernel Hilbert Space, Mercer's Theorem, Eigenfunctions, Nyström Method, and Use of Kernels in Machine Learning: Tutorial and Survey," *arXiv preprint arXiv:2106.08443*, Jun. 2021. [Online]. Available: https://arxiv.org/abs/2106.08443.

[21] Aronszajn, N., "Theory of Reproducing Kernels," Transactions of the American Mathematical Society, vol. 68, no. 3, pp. 337-404, 1950. DOI: 10.1090/S0002-9947-1950-0051437-3.

[22] Md. Abdul Mannan, Kanchon Kumar Bishnu, Shoma Islam, Md. Shafiul Alam Chowdhury, Md. Amzad Hossain, Md. Shafikul Islam, Sahib Jada Eyakub Khan, and Hashiara Khatun, "Evaluate all order of every element of higher 100, 105 and 107 order of group for multiplication composition," Edelweiss Applied Science and Technology, vol. 9, no. 7, pp. 835–850, 2025. DOI: 10.55214/25768484.v9i7.8743.

[23] Md. Abdul Mannan, Md. Amanat Ullah, Uttam Kumar Dey, and Mohammad Alauddin, "A study on Sylow theorems for finding out possible subgroups of a group in different types of order," Mathematics and Statistics, vol. 10, no. 4, pp. 851–860, 2022. DOI: 10.13189/ms.2022.100416.

[24] M. A. Mannan, H. Akter, and M. A. Ullah, "Evaluate all the order of every element in the higher even, odd, and prime order of group for composition," Science and Technology Indonesia, vol. 7, no. 3, pp. 333–343, 2022.

[25] S. Theodoridis, "Pattern Recognition," *Elsevier B.V.*, 4th ed., 2021. DOI: 10.1016/B978-0-12-823537-0.00006-5.