Quantum Machine Learning in Feature Hilbert Spaces

Nick Fularczyk

2020

Overview

- Question: Does a quantum computer "help" in solving an SVM classification problem when the feature space becomes large, and the kernel functions become computationally expensive to estimate?
- Two independent teams try to answer the same question:
  - IBM
  - Xanadu

Review

Definition

Let $\mathcal{X}$ be a nonempty set, called the input set. A function $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ is called a kernel if for any finite subset $\{x^1, \ldots, x^M\} \subset \mathcal{X}$ with $M \geq 2$ and $c_1, \ldots, c_M \in \mathbb{C}$, $\sum_{m,m'=1}^{M} c_m c_{m'}^* k(x^m, x^{m'}) \geq 0$.

Definition

Let $\mathcal{F}$ be a Hilbert space, called the feature space, $\mathcal{X}$ an input set, and x a sample from the input set. A feature map is a map $\phi : \mathcal{X} \rightarrow \mathcal{F}$ from inputs to vectors in the Hilbert space. The vectors $\phi(x) \in \mathcal{F}$ are called feature vectors.

Review

Theorem

Let $\phi : \mathcal{X} \to \mathcal{F}$ be a feature map. The inner product of two inputs mapped to feature space defines a kernel via $k(x, x') := (\phi(x), \phi(x'))_{\mathcal{F}}$, where $(\cdot, \cdot)_{\mathcal{F}}$ is the inner product defined on $\mathcal{F}$.

Theorem

Let $\phi : \mathcal{X} \to \mathcal{F}$ be a feature map over an input set $\mathcal{X}$, giving rise to a complex kernel $k(x, x') = (\phi(x), \phi(x'))_{\mathcal{F}}$. The corresponding reproducing kernel Hilbert space has the form

$$\mathcal{R}_k = \{f : \mathcal{X} \to \mathbb{C} \mid f(x) = (w, \phi(x))_{\mathcal{F}} \; \forall \, x \in \mathcal{X}, w \in \mathcal{F}\}$$

Review

Theorem (Representer Thereom)

Let $\mathcal{X}$ be an input set, $k : \mathcal{X} \text{ x } \mathcal{X} \to \mathbb{R}$ a kernel, $\mathcal{D}$ a data set consisting of data pairs $(x^m, y^m) \in \mathcal{X} \text{ x } \mathbb{R}$ and $f : \mathcal{X} \to \mathbb{R}$ a class of model functions that live in the reproducing kernel Hilbert space $\mathcal{R}_k$ of k. Furthermore, assume we have a cost function $\mathcal{C}$ that quantifies the quality of a model by comparing predicted outputs $f(x^m)$ with targets y^m, and which has a regularisation term of the form $g(||f||)$ where $g : [0, \infty) \to \mathbb{R}$ is a strictly monotonically increasing function. Then any function $f^* \in \mathcal{R}_k$ that minimises the cost function $\mathcal{C}$ can be written as

$$f^*(x) = \sum_{m=1}^{M} \alpha_m k(x, x^m),$$

for some parameters $\alpha_m \in \mathbb{R}$.

Squeezing Example

- What is squeezing?
- What is the associated Hilbert space?
  - Fock space
  - Denote the basis as $\beta = \{|0\rangle, |1\rangle, \ldots\}$

Definition

A squeezed vacuum state of the electromagnetic field is defined as

$$|z\rangle = \frac{1}{\sqrt{cosh(r)}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{2^n n!} [-e^{i\varphi} tanh(r)]^n |2n\rangle$$

where $|n\rangle$ denotes the Fock basis and $z = re^{i\varphi}$ is the complex squeezing factor. Denote $|z\rangle = |(r, \varphi)\rangle$

Squeezing Example

If $(x_1, \ldots, x_N)^T \in \mathbb{R}^N$ then we can define the joint state of N squeezed vacuum modes as

$$\phi : x \to |(c, x)\rangle,$$

where $|(c, x)\rangle = |(c, x_1)\rangle \bigotimes \ldots \bigotimes |c, x_N\rangle \in \mathcal{F}$, where $\mathcal{F}$ is now a multimode Fock space and c is a fixed constant hyperparameter. The kernel associated with this feature map is

$$k(x, x'; c) = \prod_{i=1}^{N} \langle (c, x_i) | (c, x_i') \rangle$$

with

$$\langle (c, x_i) | (c, x_i') \rangle = \sqrt{\frac{sech(c) sech(c)}{1 - e^{i(x_i' - x_i)} tanh(c) tanh(c)}}$$

Squeezing Example

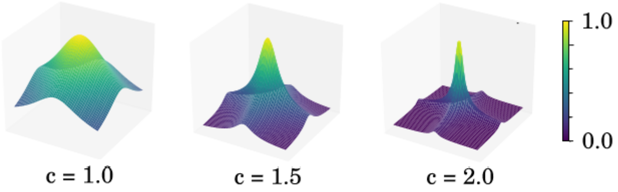
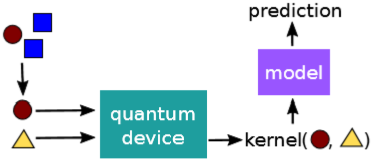


FIG. 4. Shape of the squeezing kernel function $\kappa_{\mathrm{sq}}(x, x')$ from Equation (7) for different squeezing strength hyperparameters c. The input x is fixed at $(0, 0)$ and x' is varied. The plots show the interval $[-1, 1]$ on both horizontal axes.

Overview of Quantum Kernel Estimator

- Needs to implement a feature-embedding circuit which is a map $U_\phi(x)$ such that $U_\phi(x)|0\cdots0\rangle = |\phi(x)\rangle$
- Needs to estimate inner products between quantum states
- Input these estimates into a model which runs on a classical computer



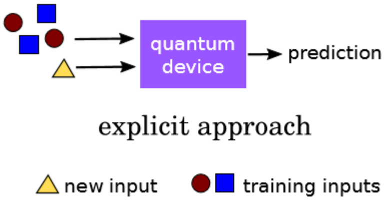
implicit approach

Overview of Quantum Variational Circuit

- Goal is to find a state $|w\rangle$ such that

$$f(x; w) = \langle w|\phi(x)\rangle$$

- $|w\rangle$ is prepared by a variational circuit, $W(\theta)$, that depends on trainable parameters θ to give $|w(\theta)\rangle = W(\theta)|0\rangle$
- Paper follows a slightly more general approach and computes the state $W(\theta)U_{\phi}|0\cdots0\rangle$ and then uses measurements to determine the output of the model.



Quantum Kernel Estimator based on two-dimensional Squeezing

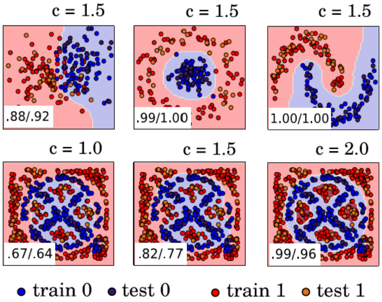
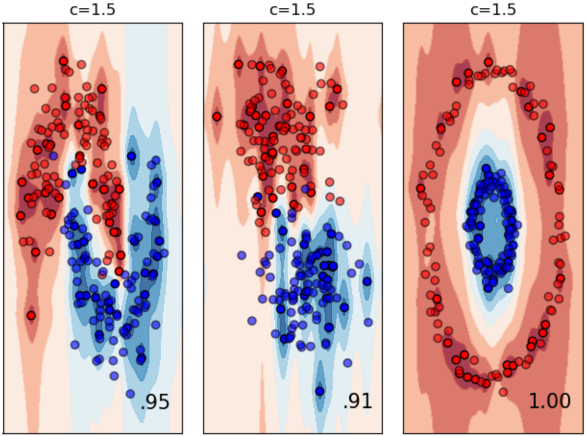


FIG. 5. Decision boundary of a support vector machine with the custom kernel from Eq. (7). The shaded areas show the decision regions for Class 0 (blue) and Class 1 (red), and each plot shows the rate of correct classifications on the training set/test set. The first row plots three standard 2-dimensional datasets: 'circles', 'moons' and 'blobs', each with 150 test and 50 training samples. The second row illustrates that increasing the squeezing hyperparameter c changes the classification performance. Here we use a dataset of 500 training and 100 test samples. Training was performed with python's *scikit-learn* SVC classifier using a custom kernel which implements the overlap of Eq. (8).

Quantum Kernel Estimator based on two-dimensional Squeezing



References

- Schuld, M. and Killoran, N. Quantum machine learning in feature Hilbert spaces. Phys. Rev. Lett. 122, 040504 (2019). https://arxiv.org/abs/1803.07128
- Havlíček, V., Córcoles, A.D., Temme, K. et al. Supervised learning with quantum-enhanced feature spaces. Nature 567, 209–212 (2019).