Proposition 1. *Jaccard index defined for two finite nonempty sets A, B in a universe Ω with finite elements is a kernel:*

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}. \tag{1}$$

Proof. We need several preliminary results in order to prove the above proposition.

Lemma 2. *Closure properties of kernels:*

- *If k_1 and k_2 are kernels, then so is $ak_1 + bk_2$ for scalars $a, b \geq 0$.*
- *If k_1 and k_2 are kernels, then so is $k_1 k_2$.*

See lecture note for a proof.

Lemma 3. *If $k(x, y)$ is a kernel bounded from above: $\max_{x,y} k(x, y) = \max_x k(x, x) < D < \infty$, then*

$$k'(x, y) = \frac{1}{D - k(x, y)}, \tag{2}$$

is also a kernel.

Proof. Using series expansion:

$$k'(x, y) = \frac{1}{D} \sum_{n=0}^{\infty} \left(\frac{k(x, y)}{D} \right)^n. \tag{3}$$

The expansion converges because $\max_{x,y} k(x, y) < D$. With Lemma 2, $k'(x, y)$ is a kernel. (One might wonder whether Lemma 2 works for countable infinite many terms. The answer is positive as long as the series converges, because essentially we only need to prove the positive semi-definiteness of the kernel.) □

One last result we need is

Lemma 4. *For two nonempty sets A, B with finite elements, $k(A, B) = |A \cap B|$ and $k(A, B) = |\bar{A} \cap \bar{B}|$ are kernels.*

Proof. Use bit encoding of the set. For $k(A, B) = |A \cap B|$, explicitly construct feature map $\phi(A)$ as follows: $\phi(A)$ is a vector such that $[\phi(A)]_i = 1$ if i-th element is in A and $[\phi(A)]_i = 0$ otherwise. Then $k(A, B) = \phi(A)^T \phi(B)$. For $k(A, B) = |\bar{A} \cap \bar{B}|$, use $\psi = 1 - \phi$ as the feature map. □

Note $A \cup B = \overline{\bar{A} \cap \bar{B}}$, where $\bar{A} \equiv \Omega - A$ is the complement of set A. It follows $|A \cup B| = |\Omega| - |\bar{A} \cap \bar{B}|$. Using Lemma 3 and 4, $1/|A \cup B|$ is a kernel. Then using Lemma 2 again, $|A \cap B|/|A \cup B|$ is a kernel. □

Remark. What is the feature map of the Jaccard index? It should be very hard to write down explicitly. Let us take a step back and consider the feature map of $k(A, B) = 1/|A \cup B|$, which is also proved to be a kernel:

$$k(A, B) = \frac{1}{|A \cup B|} = \frac{1}{|\Omega| - |\bar{A} \cap \bar{B}|} = \frac{1}{|\Omega|} \sum_{n=0}^{\infty} \left(\frac{|\bar{A} \cap \bar{B}|}{|\Omega|} \right)^n. \tag{4}$$

Since we already know the feature map of $|\bar{A} \cap \bar{B}|$ as ψ, the feature map of $|\bar{A} \cap \bar{B}|^n$ can simply be the tensor product $\psi^{\otimes n}$ in a $|\Omega|n$-dimensional space. Because the infinite summation is involved, the feature mapping of $1/|A \cup B|$ is also in a countable infinite dimensional space:

$$\varphi = \frac{1}{\sqrt{|\Omega|}} \sum_{n=0}^{\infty} \frac{\psi^{\otimes n}}{\sqrt{|\Omega|^n}}, \tag{5}$$

where the summation is the direct sum.

With the feature map of $|A \cap B|$: ϕ (finite-dimensional) and $1/|A \cup B|$: φ (infinite-dimensional), the feature map of Jaccard index can be constructed similarly using tensor product: $\phi \otimes \varphi$ and is infinite-dimensional. Note that this construction does not necessarily exclude a finite-dimensional feature map because feature map is not unique.

Proposition 5. *Jaccard index defined for two finite nonempty sets A, B in a universe Ω with infinite elements is a kernel:*

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}. \tag{6}$$

Proof. Now since $\max_x k(x,x) = |\Omega| = \infty$, we cannot use Lemma 3. The work around is to use another representation of $|A \cup B|$ and the following lemma:

Lemma 6. *The following index for two nonempty sets A, B is a kernel:*

$$H(A,B) = \frac{1}{|A| + |B|}. \tag{7}$$

Proof. According to Mercer's theorem, we need to prove its positive semi-definiteness. In particular, it suffices to prove any $n \times n$ matrix M, where $n \in \mathbb{N}$ and $M_{i,j} = 1/(i+j)$ is positive semi-definite. This is exactly a Hilbert matrix and its positive semi-definiteness is well-known. To prove this fact, for all c_i, c_j:

$$\sum_{i,j=1}^{n} c_i M_{i,j} c_j = \sum_{i,j=1}^{n} \frac{c_i c_j}{i+j} = \sum_{i,j=1}^{n} c_i c_j \int_0^1 t^{i+j-1} dt = \int_0^1 t \left(\sum_{i=1}^{n} c_i t^{i-1} \right)^2 dt \geq 0. \tag{8}$$

Hence we have proved $H(A,B)$ is a kernel. Note that one can also use another trick $1/(i+j) = \int_0^{+\infty} e^{-(i+j)t} dt$ in the proof. $\square$

Decompose $|A \cup B|$ as

$$\frac{1}{|A \cup B|} = \frac{1}{|A| + |B| - |A \cap B|} = \frac{1}{|A| + |B|} \frac{1}{1 - \frac{|A \cap B|}{|A| + |B|}} = \frac{H(A,B)}{1 - H(A,B)|A \cap B|}. \tag{9}$$

Using Lemma 2, 3, 4 and 6, $1/|A \cup B|$ is a kernel even if the universe Ω has infinite elements. Then use Lemma 2 again, $J(A,B)$ is a kernel. $\square$