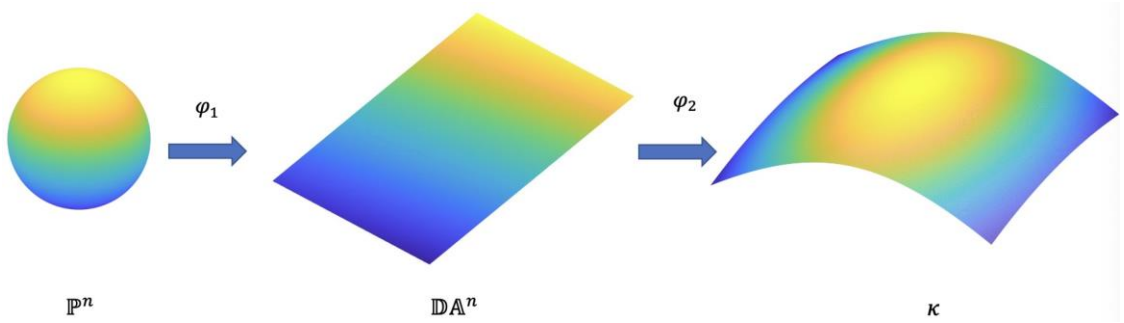
Towards Kernelizing the Classifier for Hyperbolic Data

Meimei YANG, Qiao LIU, Xinkai SUN, Na SHI, Hui XUE

Frontiers of Computer Science , DOI: 10.1007/s11704-022-2457-y

Problems & Ideas

- Problems of classification of hierarchical data:
  - To encode the hierarchical data in Euclidean space and then train a Euclidean classifier leads to a performance drop due to distortion of data embedding in the Euclidean space.
  - Although much progress has been made in hyperbolic embedding hierarchical data, existing hyperbolic kernel methods to classify the hierarchical data still lead to distortion and the classification results are not so good.
- Ideas: Two novel kernel formulations in the hyperbolic space, with one being positive definite (PD) and another one being indefinite, to solve the classification tasks in hyperbolic space.



The mapping process form Poincaré ball (a model of hyperbolic space) to the Kre ĭn space induced from DA-Sigmoid kernel

Main Contributions

- Contributions:
  - Using a positive definite kernel defined in the Drury-Arveson (DA) space, a special reproducing kernel Hilbert space (RKHS), thus the barrier from hyperbolic space to Euclidean space is broken.
  - An indefinite kernel mapping from hyperbolic space to Kreĭn spaces, with DA space as the bridge between them, providing new classification methods of hyperbolic data.

Dataset \ Kernel	Dim	Hyperbolic					Euclidean			
		DA-Sigmoid	H-Poly	H-RBF	H-Laplace	H-Binomial	E-Sigmoid	E-RBF	E-Laplace	E-Poly
Facebook	#2	**85.5**±0.5	69.1±0.4•	73.9±1.3•	74.6±1.8•	66.0±1.3•	71.5±1.0•	67.2±0.4•	67.2±0.3•	66.0±0.9•
	#5	**90.5**±0.3	82.3±0.6•	87.4±0.5•	87.6±0.5•	81.8±3.6•	86.4±0.9•	85.1±0.5•	85.3±0.6•	82.3±0.7•
	#10	**90.7**±0.5	89.3±0.6•	87.7±1.3•	88.3±0.5•	87.3±3.5•	88.8±0.7•	86.6±1.2•	86.9±0.6•	86.2±0.6•
	#25	91.0±0.5	**91.3**±0.4○	88.0±1.6•	88.2±0.4•	86.5±1.9•	90.2±0.1	88.0±0.8•	88.0±0.5•	88.8±0.7•
Terrorist	#2	59.6±2.1	52.6±1.3•	59.0±4.7	60.5±3.9	51.3±3.0•	59.0±2.1	60.1±2.3	**60.6**±1.5	51.5±1.0•
	#5	**68.1**±1.6	53.2±2.2•	64.6±3.3•	66.6±2.0•	57.6±7.7•	63.2±1.5•	62.9±1.4•	64.4±1.5•	56.4±1.2•
	#10	**68.2**±2.3	59.6±1.5•	65.9±1.5	66.0±1.9•	61.4±8.0•	67.9±1.3	65.5±2.3•	66.3±2.1•	64.9±1.5•
	#25	**67.8**±1.5	61.6±1.2•	64.8±1.9•	66.0±2.3•	60.9±5.4•	67.6±1.6	63.2±1.5•	63.0±2.2•	66.0±1.0•
Wiki	#2	**58.4**±0.8	40.2±0.5•	38.6±1.9•	41.0±2.4•	27.5±3.1•	31.3±2.0•	23.6±1.3•	23.9±0.9•	19.3±0.7•
	#5	**71.6**±0.9	63.7±0.7•	63.0±2.1•	64.8±1.8•	47.7±4.3•	62.6±0.8•	58.8±1.6•	59.3±1.4•	45.1±3.1•
	#10	**73.2**±0.7	68.9±0.7•	66.5±1.5•	67.1±1.8•	49.8±5.1•	67.9±1.4•	66.3±1.0•	65.4±1.7•	54.1±4.1•
	#25	**74.0**±1.3	72.2±0.9	67.9±1.6•	68.9±1.6•	55.5±4.2•	72.0±0.8•	69.0±1.6•	68.1±1.3•	65.1±1.4•
AC	#2	**79.1**±0.6	73.6±0.9•	74.5±2.2•	75.5±1.1•	69.6±3.2•	73.4±1.0•	71.9±1.4•	72.5±1.0•	66.9±4.0•
	#5	**87.2**±0.9	79.1±0.7•	81.8±9.9•	86.1±1.1•	82.9±2.1•	78.7±0.5•	77.6±0.8•	79.2±0.5•	73.6±1.9•
	#10	**89.8**±0.4	88.0±0.8•	86.8±0.6•	87.6±0.6•	86.8±1.1•	85.3±1.0•	84.4±1.1•	84.9±0.8•	79.7±2.6•
	#25	89.6±0.6	**89.7**±0.6	85.4±5.0•	87.3±0.5•	87.3±1.1•	87.3±0.3•	85.9±0.8•	86.7±0.4•	84.2±0.6•
Cora ML	#2	**71.5**±0.6	42.2±3.9•	67.2±1.3•	68.8±0.9•	51.7±3.5•	61.3±0.8•	56.5±0.6•	57.0±0.7•	45.9±2.2•
	#5	**84.9**±0.6	75.5±0.2•	83.3±0.7•	83.6±1.0•	79.5±1.7•	73.8±1.3•	71.0±0.6•	71.3±0.5•	57.2±1.6•
	#10	**86.1**±0.3	82.7±0.4•	84.7±0.6•	85.0±0.4•	81.7±2.9•	77.1±1.5•	76.1±0.8•	76.7±0.9•	71.6±1.0•
	#25	**86.4**±0.5	85.7±0.6•	84.8±0.5•	85.1±0.5•	83.6±1.3•	79.3±0.8•	69.5±1.3•	77.7±1.4•	76.4±1.2•
Avg.ACC.		**78.7**	71.1	73.8	74.9	67.8	72.2	69.5	70.2	65.0
Top1 Times		17	2	0	0	0	0	0	1	0

Average ACC of node classification on real-world datasets. DA-Sigmoid kernel our proposed indefinite kernel to classify hyperbolic data. 'Hyperbolic' refers to the hyperbolic embedding of the hierarchical structure, and then classifying the hyperbolic data. 'Euclidean' refers to the classification of Euclidean embedded data.