Beyond Euclidian Embeddings
(Introduction to Hyperbolic Embeddings)

Why Go Beyond Euclidean Embeddings?

- Euclidean Embeddings are great:
  - We have great understanding of the Euclidean geometry
  - All our prediction models are designed for (and crucially depend on!) Euclidean spaces

- So why bother embedding with other geometries?
  - Certain types of data (graphs, networks, etc) cannot be embedded in Euclidean spaces isometrically but can be embedded in other geometries!

- Hold on, what about Nash's theorem?
  - Recall: any (compact) n-manifold can be embedded isometrically in Euclidean space of dimension 2n+1.
  - So… if we can embed into another geometry, ie an "abstract" manifold, then can't we embed it in Euclidean space??

Towards Other Interesting Geometries

- Recall the Euclidian postulates:
  - A straight line segment can be drawn joining any two points.
  - Any straight line segment can be extended indefinitely in a straight line.
  - Given any straight line segment, a circle can be drawn having the segment as radius and one endpoint as center.
  - All right angles are congruent.
  - [*Parallel Postulate*] Given any straight line and a point not on it, there "exists one and only one straight line which passes" through that point and never intersects the first line, no matter how far they are extended.

can we relax the parallel postulate? e.g.

- multiple distinct straight lines (passing through the point) are all parallel?
- no straight line (passing through the point) is parallel?

(this gives rise to other geometries)

Towards Other Interesting Geometries

Modified parallel postulate: Given any straight line L and a point p not on it,

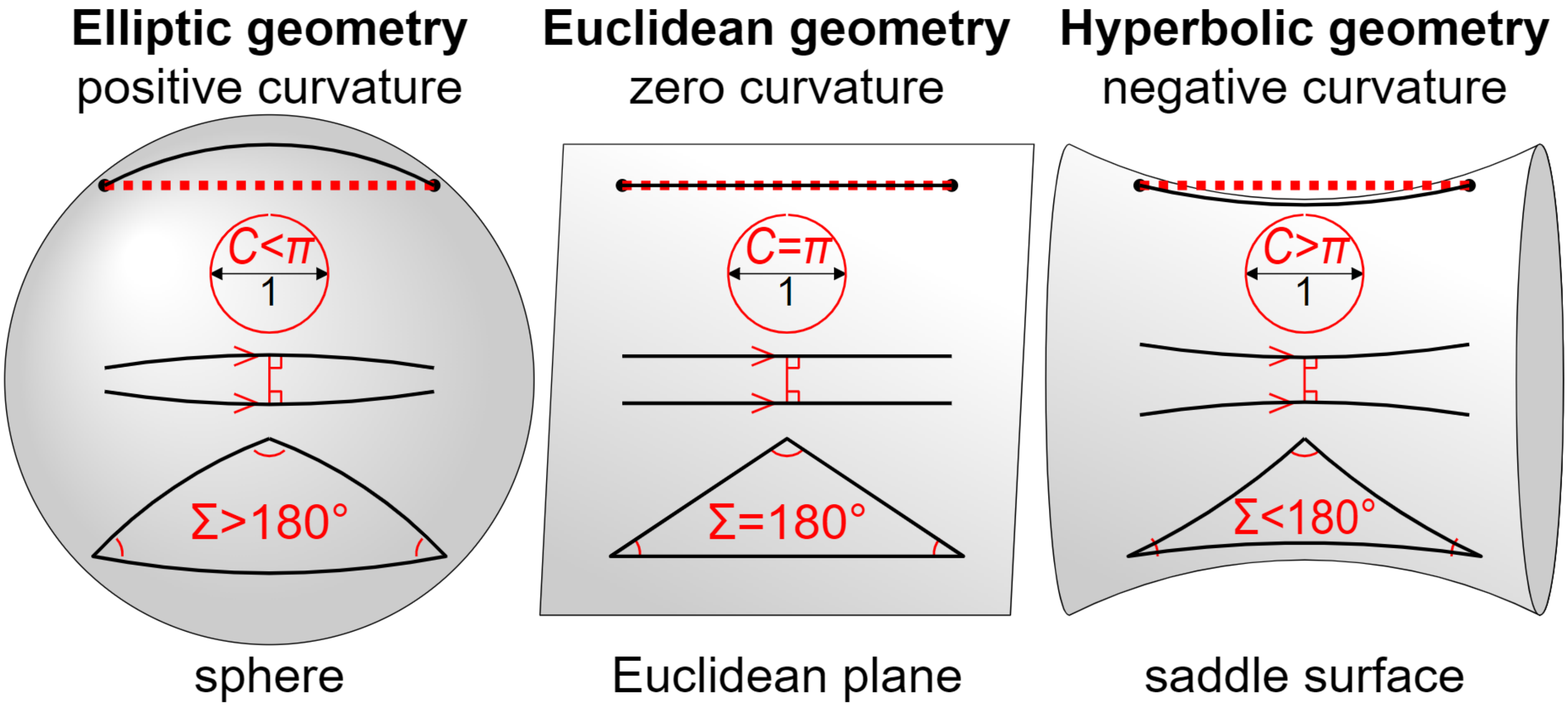
- (option 1) No straight line (passing through p) never intersects L
  (ie ALL straight lines through p intersect L)

  this gives rise to Elliptical geometry

- (option 2) Multiple (distinct) straight lines (passing through p) never intersect L

  this gives rise to Hyperbolic geometry

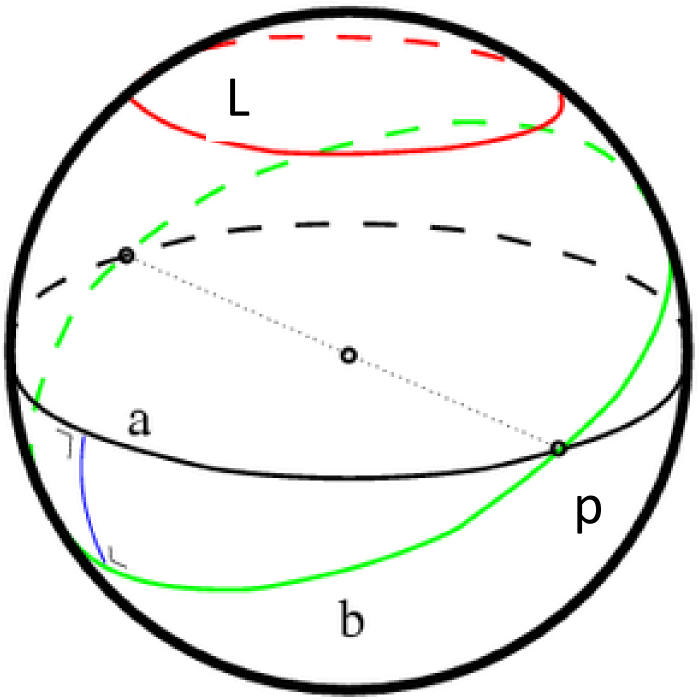
Resulting Geometries



Elliptical Geometry

(spherical geometry) Constant positive curvature everywhere

Hold on!



- The point p is not on the line L
- Lines a and b, both passing through p, don't intersect L

what is going on?

Hyperbolic Geometry

A saddle at every point in space
(constant negative curvature everywhere)

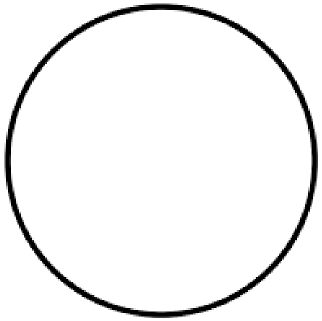
difficult to visualize directly…

why cant we just use a hyperboloid?

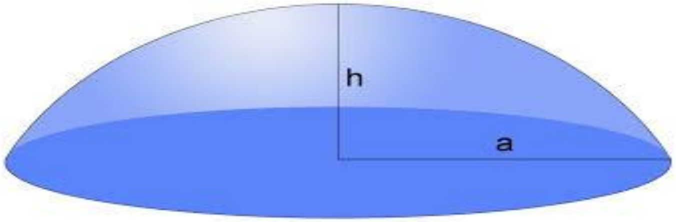
Advantages of Such Geometries

Consider a unit ball in 2-dim

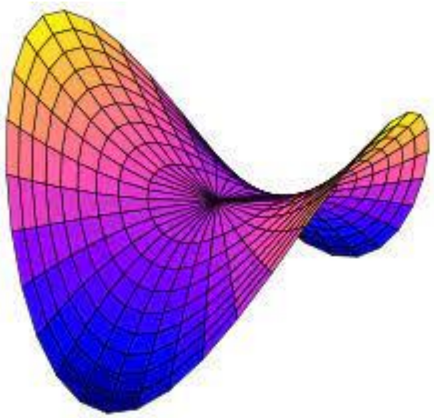
- Euclidean space



- Spherical space



- Hyperbolic space



what is the circumference?

Growth Rate of the Circumference

For a full and complete binary tree, number of leaves at level *r* :

$$2^r$$

For a Euclidean ball of radius r in R^d, the circumference is:

$$\sim O(r^{d-1})$$

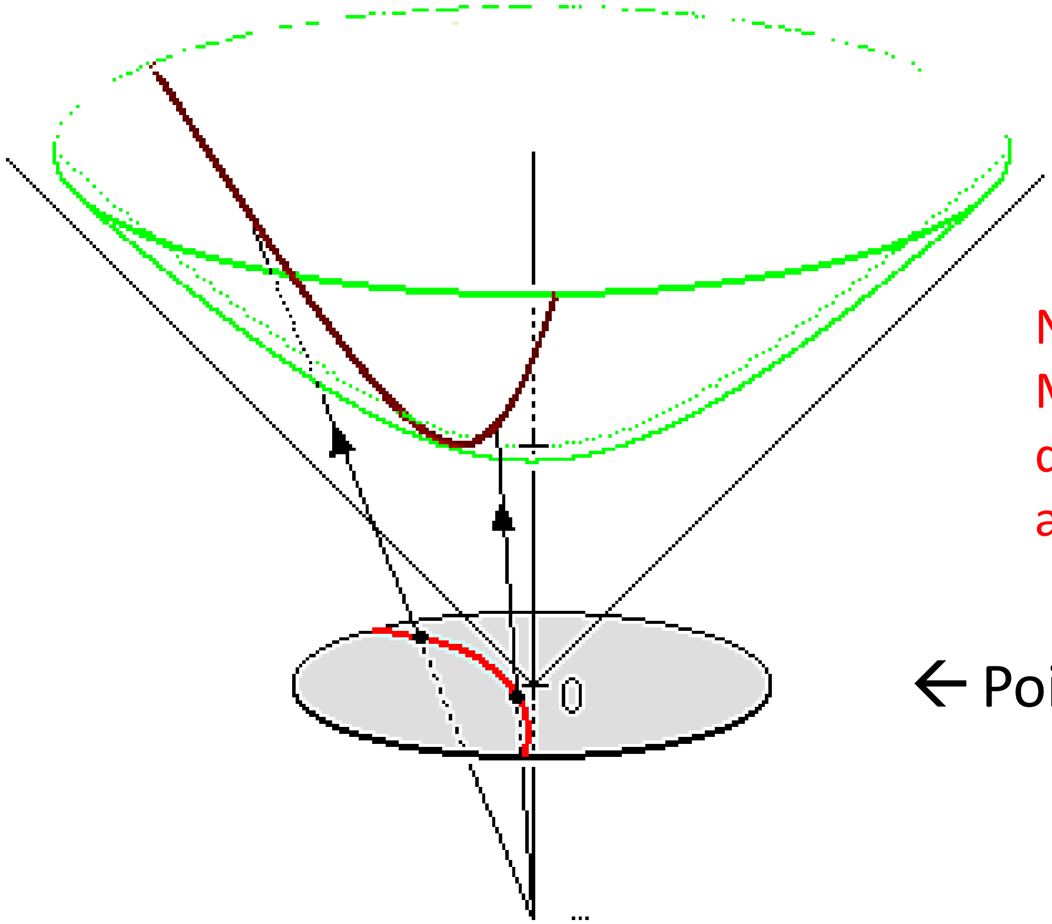
after certain number of levels, we will run out of space to isometrically embed the nodes at level r!

For a 2-dimension Hyperbolic ball of radius *r*, the circumference is:

$$\sim \sinh(r) = (e^r - e^{-r}) / 2$$

The space grows exponentially so it has the potential to accommodate trees!

Models for Understanding Hyperbolic Spaces

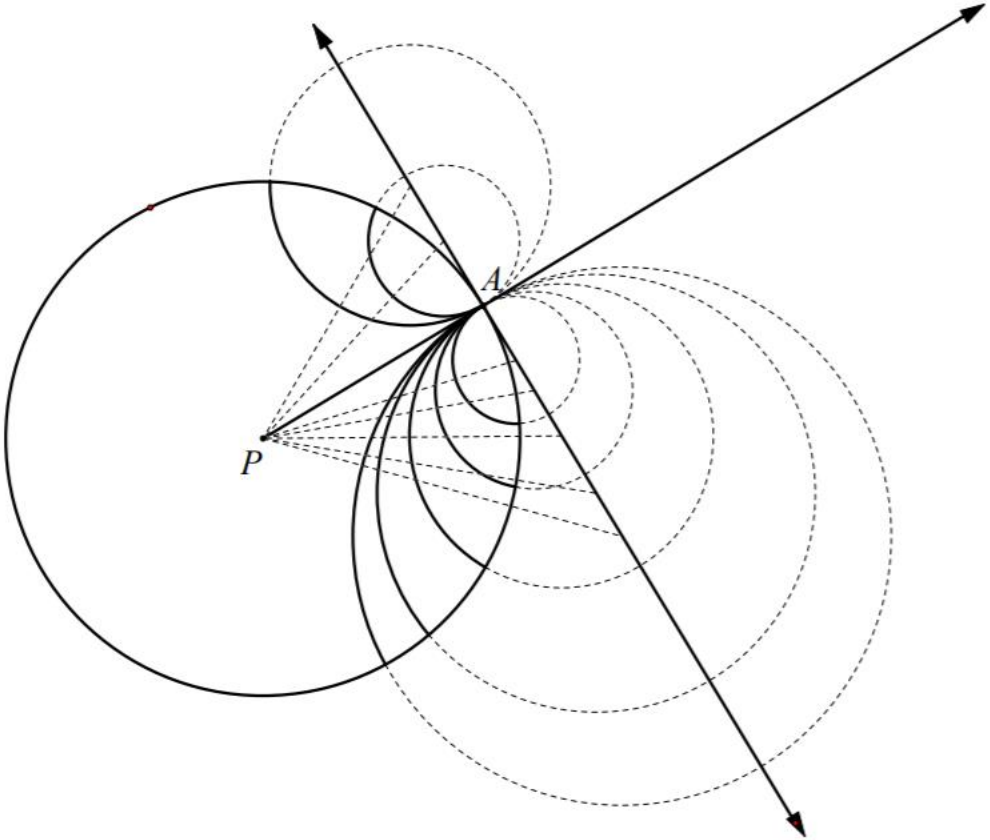


← Hyperboloid model
(using the positive sheet)

Need to be careful, we use the Minkowski metric to compute distances and hence the geodesics are not what you expect!

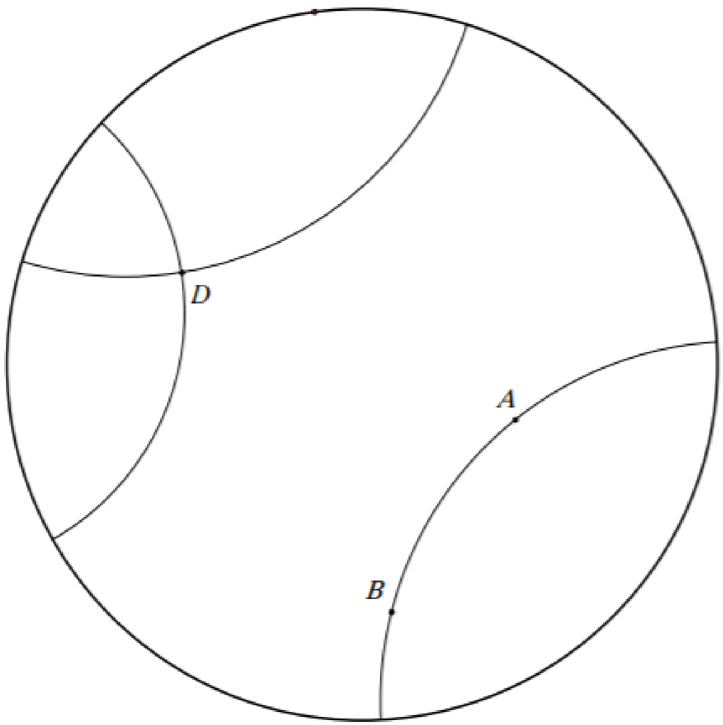
← Poincare disk model

Poincare Disk Model



curves (that are part of a circle) that intersect perpendicularly with the disk boundary are straight lines in the hyperbolic space

Poincare Disk Model



A depiction of parallel postulate for hyperbolic space

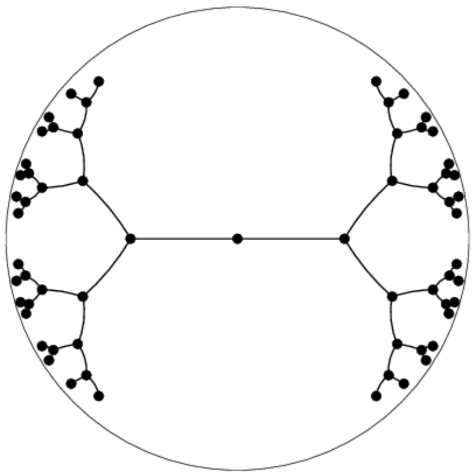
pairwise distance in this model:

$$d_H(x, y) = \text{acosh}\left(1 + 2\frac{\|x - y\|^2}{(1 - \|x\|^2)(1 - \|y\|^2)}\right)$$

(distinct) points close to the boundary are **very** far apart from each other, and from the origin

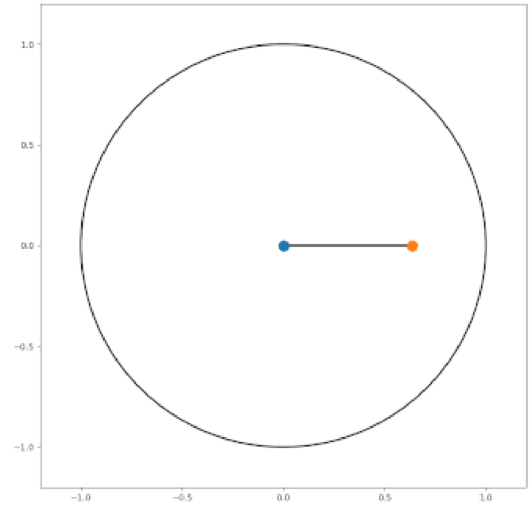
(Approximate) Isometric Embeddings of Trees

Sarkar's Construction [2011]



Surprisingly simple construction:

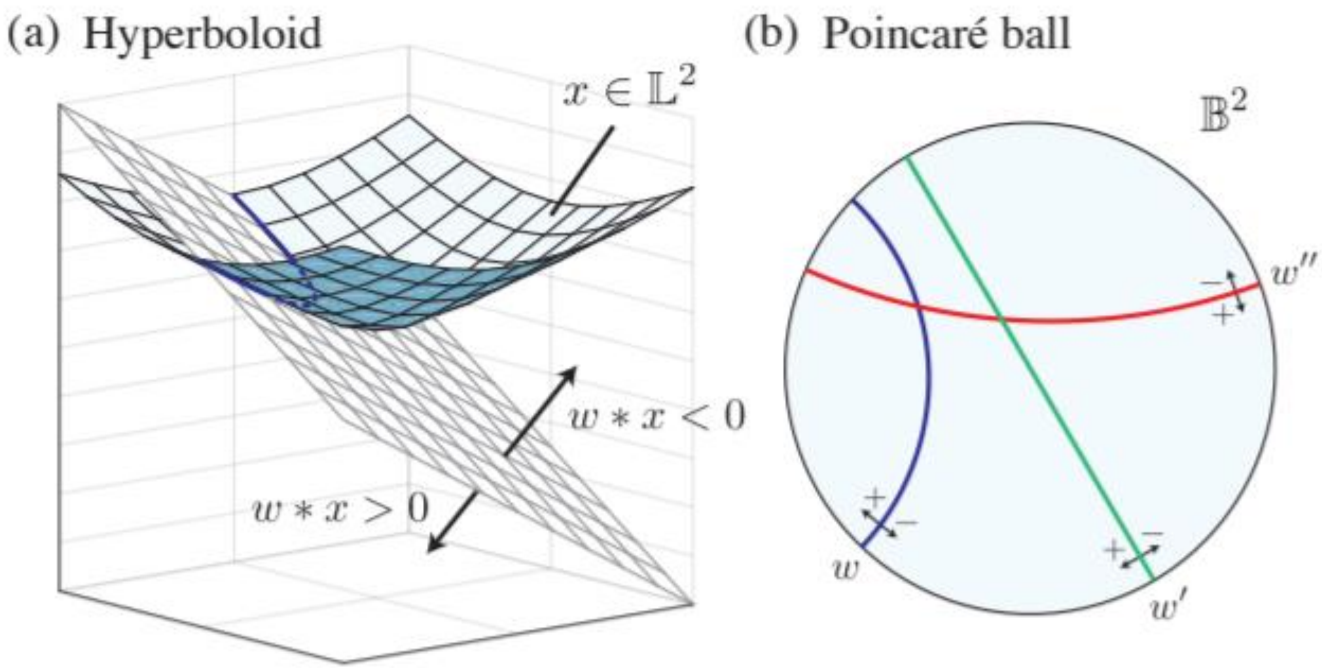
- Suppose nodes *a* and *b* are already embedded (where *b* is a parent of *a*). Let *f(a)* and *f(b)* be the corresponding embedding. Let $c_1,\ldots,c_a$ be the children of *a*
- *f(a)* and *f(b)* are reflected across a geodesic s.t. *f(a)* is mapped to origin and *f(b)* is mapped to some point *z*
- Place the children $c_1,\ldots,c_a$ equally spaced around a circle around origin maximally away from *z*
- Reflect all points back across the geodesic!



Machine Learning in Hyperbolic Spaces

- Hyperbolic SVM

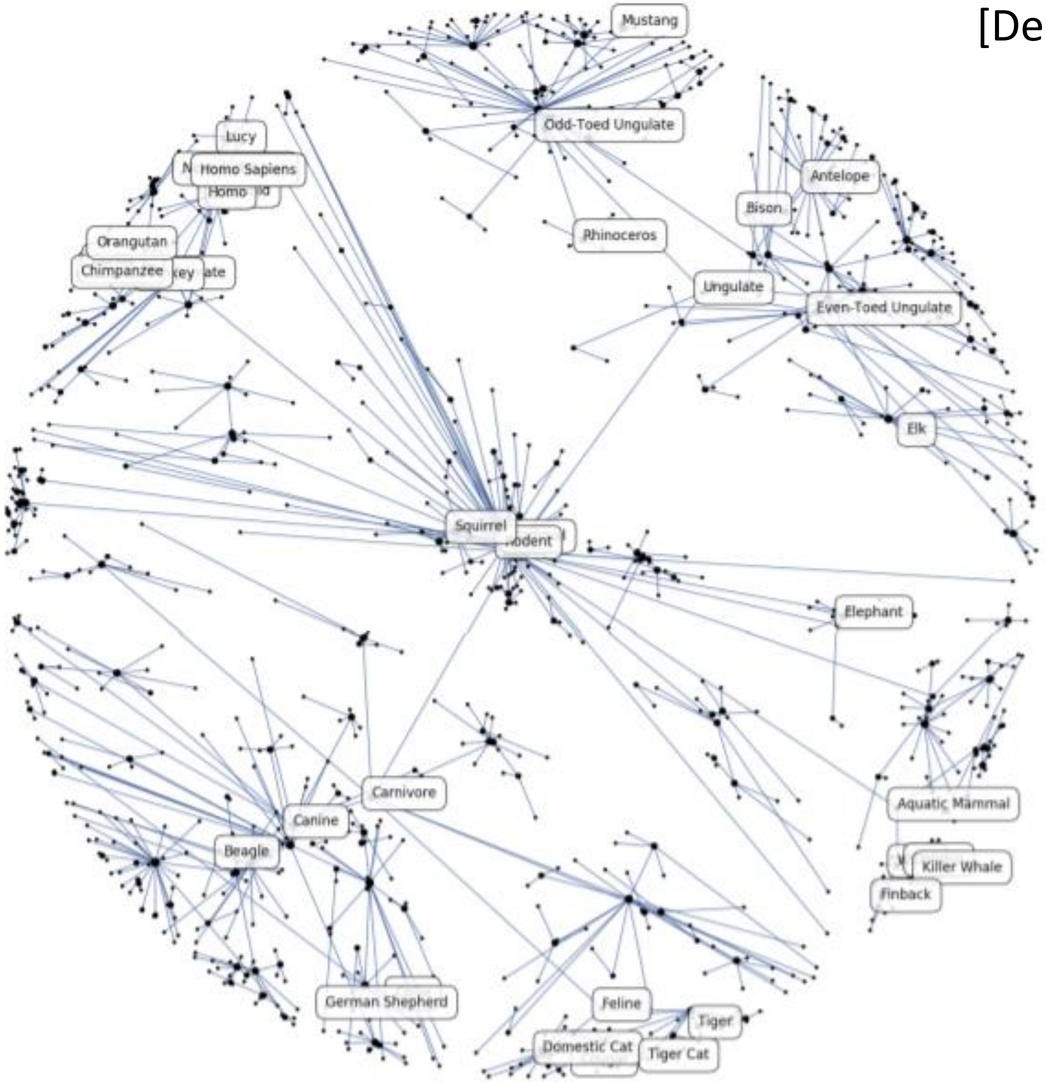
[Cho, DeMeo, Peng, Berger ’19]



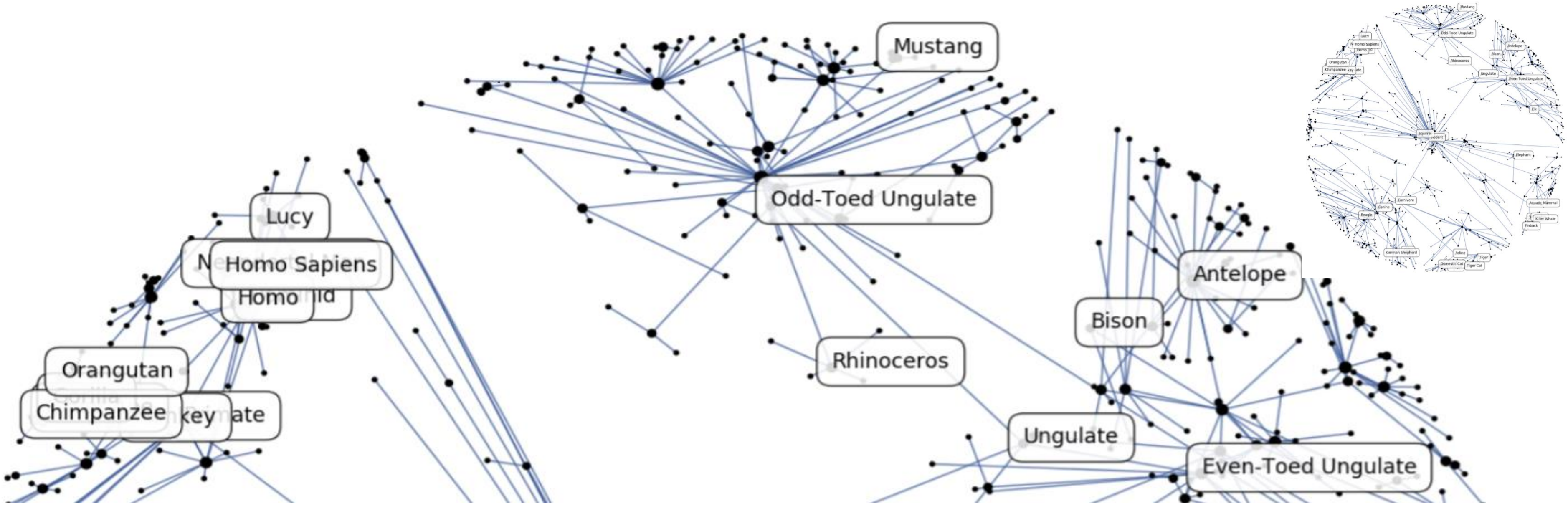
Machine Learning in Hyperbolic Spaces

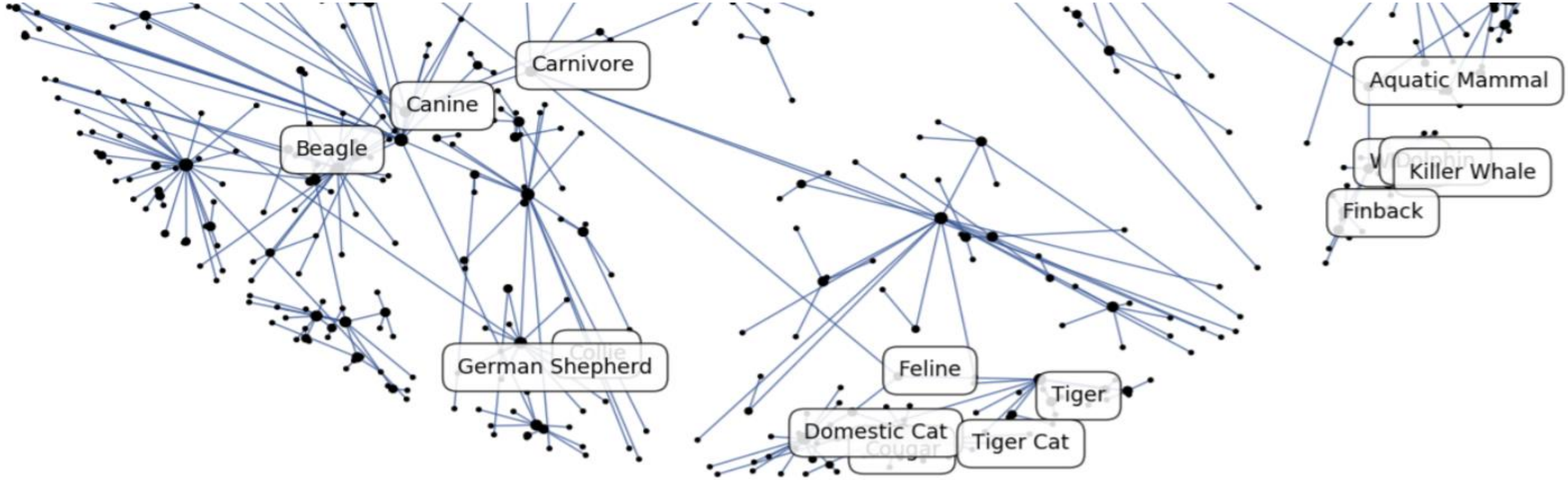
- Hierarchical Word Embeddings (WordNet)

[Nickel, Kiela ’17]
[De Sa, Gu, Re, Sala ’18]



Machine Learning in Hyperbolic Spaces



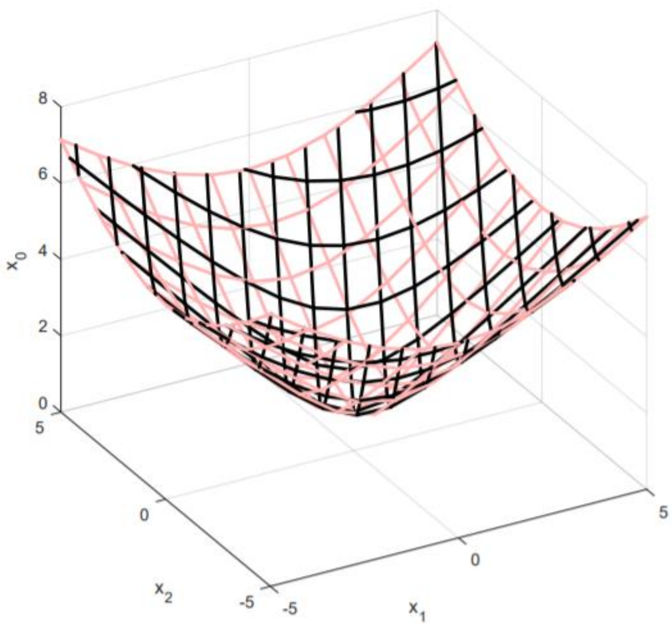


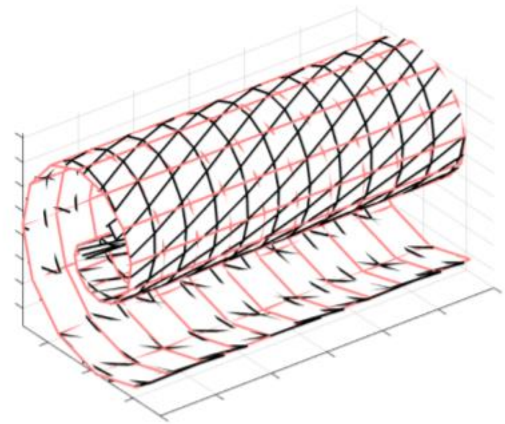
Machine Learning in Hyperbolic Spaces

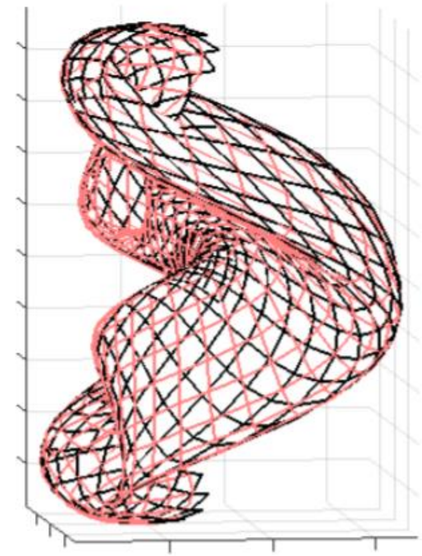
Manifold

- ~~Hyperbolic~~ Metric Learning

[Aalto & Verma '19]







k-NN accuracy:

Dataset	Euclidean	Euclidean+Metric Learn	Hyperbolic	Hyperbolic+Metric Learn
`football`	0.41 ± 0.01	0.40 ± 0.09	0.29 ± 0.09	$\mathbf{0.25 \pm 0.10}$
`polbooks`	0.24 ± 0.05	0.31 ± 0.12	0.25 ± 0.06	$\mathbf{0.23 \pm 0.06}$
`adjnoun`	0.58 ± 0.06	0.56 ± 0.07	0.55 ± 0.09	$\mathbf{0.49 \pm 0.05}$