

**Explaining Black-Box
Structured Models**

EMNLP'17

**Robustness in
Interpretability**

WHI'18

**Self-explaining
neural nets**

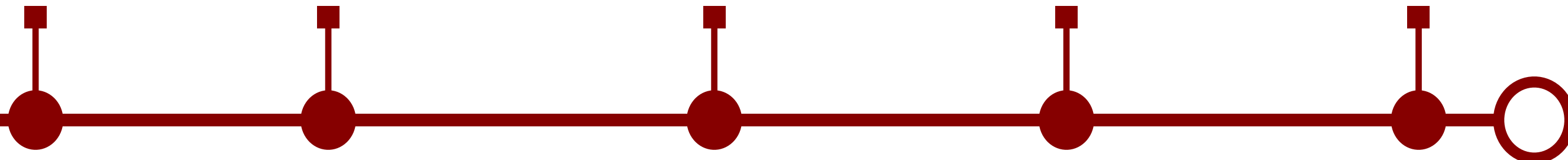
NeurIPS'18

**Robustness via
Local Linearity**

ICLR'19

**Game-theoretic
transparency**

ICML'19



Interpretability and Robustness

Structure in ML

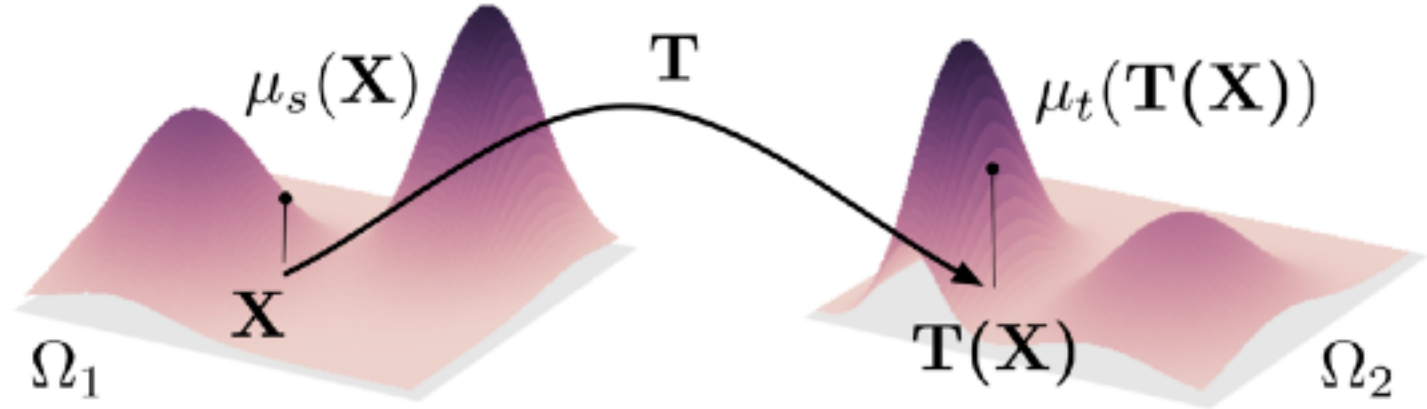


**Theoretical
framework for
Word Embeddings**
TACL'16

**Tree-structured
decoding with
Neural Networks**
ICLR'17

Generative Models





A dark blue line starts at the top left, slopes down to the left, then turns 90 degrees to the right and extends horizontally across the bottom of the image, ending in a small circle.

Optimal Transport



OT on Hyperbolic Embeddings

In Submission

Generative Models across Incomparable Spaces

ICML '19





OT with Global Invariances

AISTATS'19



Gromov-Wasserstein Embedding Alignment

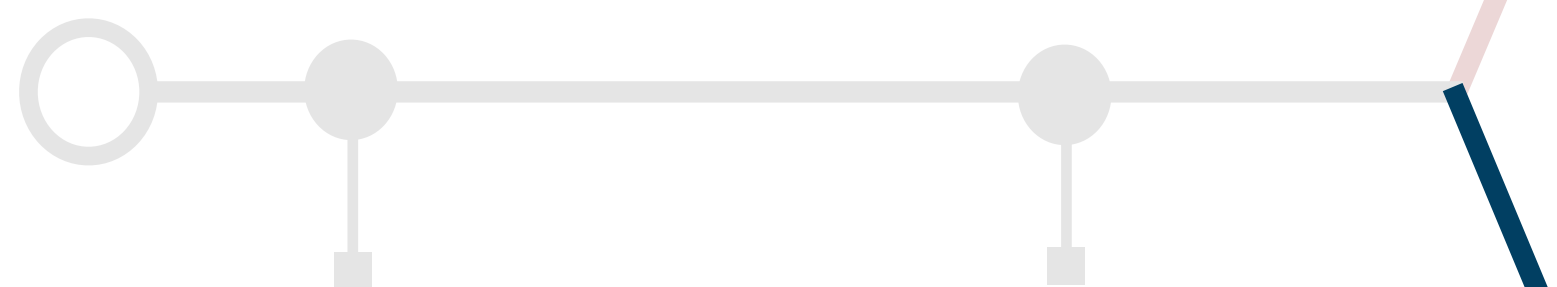
EMNLP'18



Structured OT via Submodularity

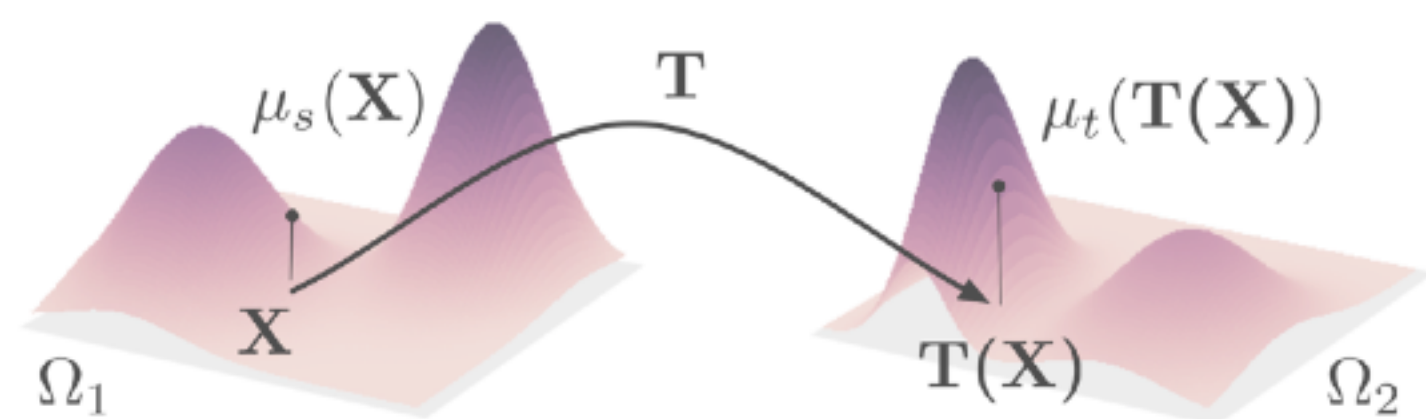
A/STATS'18

Structure in ML



Theoretical
framework for
Word Embeddings
TACL'16

Tree-structured
decoding with
Neural Networks
ICLR'17



Explaining Black-Box
Structured Models
EMNLP'17

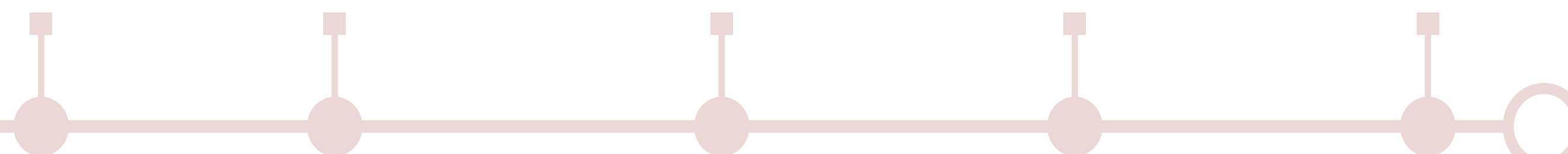
Robustness in
Interpretability
WHI'18

Self-explaining
neural nets
NeurIPS'18

Robustness via
Local Linearity
ICLR'19

Game-theoretic
transparency
ICML'19

Interpretability and Robustness



Generative Models



Optimal Transport

Structured OT
via Submodularity
AISTATS'18

Gromov-Wasserstein
Embedding Alignment
EMNLP'18

OT with Global
Invariances
AISTATS'19

Generative Models across
Incomparable Spaces
ICML'19

OT on Hyperbolic
Embeddings
In Submission



