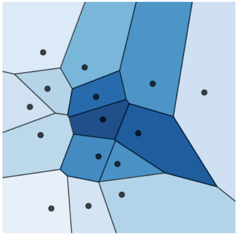
Active Nearest Neighbor Regression Through Delaunay Refinement

Alexander Kravberg*, Giovanni Luca Marchetti*, Vladislav Polianskii*,
Anastasiia Varava, Florian T. Pokorny, Danica Kragic

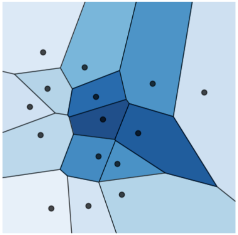
Nearest Neighbor Regressor

The Nearest Neighbor Regressor (NNR) approximates an unknown function f by the value at the closest datapoint. It is locally constant on Voronoi cells.



Nearest Neighbor Regressor

The Nearest Neighbor Regressor (NNR) approximates an unknown function f by the value at the closest datapoint. It is locally constant on Voronoi cells.



We upgrade the NNR to an active regressor by querying a novel datapoint p_{t+1} at step t based on the current dataset P_t.

Active Nearest Neighbor Regressor

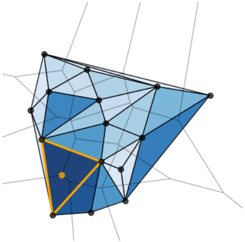
Our querying strategy looks at the geometry of the graph of f discretized via the dual Delaunay triangulation Del_{P_t}

Active Nearest Neighbor Regressor

Our querying strategy looks at the geometry of the graph of f discretized via the dual Delaunay triangulation Del_{P_t} :

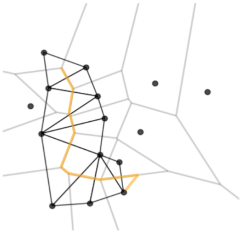
$$p_{t+1} = \mathrm{Circ}(\overline{\sigma}), \qquad \overline{\sigma} = \mathrm{argmax}_{\sigma \in \mathrm{Del}_{P_t}} \mathrm{Vol}(\hat{\sigma}).$$

Here $\hat{\sigma}$ is the lifting of σ to the graph of λf and Circ is the circumcenter. The hyperparameter λ controls an 'exploration-exploitation' tradeoff.



Computation and Theory

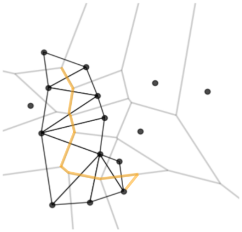
We compute volumes of simplices via the Cayley-Menger determinant and approximate the Delaunay triangulation via a random walk over the Voronoi boundaries [PP20].



[PP20] Polianskii and Pokorny, Voronoi Graph Traversal in High Dimensions, KDD 2020.

Computation and Theory

We compute volumes of simplices via the Cayley-Menger determinant and approximate the Delaunay triangulation via a random walk over the Voronoi boundaries [PP20].



We prove that the ANNR is guaranteed to halt:

$$\underline{\lim}_{t\to\infty} \max_{\sigma\in \mathrm{Del}_{P_t}} \mathrm{Vol}(\hat{\sigma}) = 0.$$

We give a continuous geometric interpretation of the querying:

$$\log \mathrm{Vol}(\Gamma_f) \geq$$
$$\geq C\lambda^2 \left\| f - \fint_\Omega f \right\|_2^2 + \log \mathrm{Vol}(\Omega) + \cdots$$

[PP20] Polianskii and Pokorny, Voronoi Graph Traversal in High Dimensions, KDD 2020.

Empirical Results

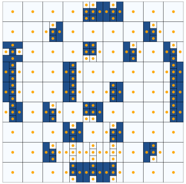
We compare the ANNR with DEFER [BDCE21] – an active function approximator based on rectangular partitioning.

[BDCE21] Bodin et al., Black-Box Density Function Estimation Using Recursive Partitioning, ICML 2021.

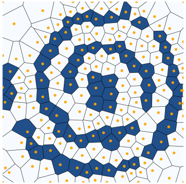
Empirical Results

We compare the ANNR with DEFER [BDCE21] – an active function approximator based on rectangular partitioning.

Approximation of (characteristic functions of) manifolds:



DEFER



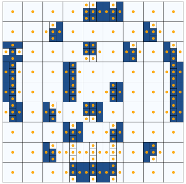
ANNR

[BDCE21] Bodin et al., Black-Box Density Function Estimation Using Recursive Partitioning, ICML 2021.

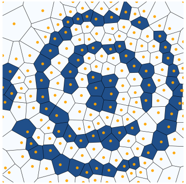
Empirical Results

We compare the ANNR with DEFER [BDCE21] – an active function approximator based on rectangular partitioning.

Approximation of (characteristic functions of) manifolds:



DEFER



ANNR

Approximation of articulated densities:

Table: Mean Test Error

Density	ANNR	DEFER
Gravitational Waves	0.47	0.53
Latent Volume Density	47.05	50.50

[BDCE21] Bodin et al., *Black-Box Density Function Estimation Using Recursive Partitioning*, ICML 2021.

Thank You!

References I

Erik Bodin, Zhenwen Dai, Neill Campbell, and Carl Henrik Ek.
Black-box density function estimation using recursive partitioning.
In *ICML*, 2021.

Vladislav Polianskii and Florian T Pokorny.
Voronoi graph traversal in high dimensions with applications to topological data analysis and piecewise linear interpolation.
In *KDD*, 2020.