

Implement k-Means Clustering

Implement k-Means Clustering

Implement k-Means using the TensorFlow [k-Means](https://www.tensorflow.org/api_docs/python/tf/compat/v1/estimator/experimental/KMeans)

(https://www.tensorflow.org/api_docs/python/tf/compat/v1/estimator/experimental/KMeans) API.

The TensorFlow API lets you scale k-means to large datasets by providing the following functionality:

- Clustering using mini-batches instead of the full dataset.
- Choosing more optimal initial clusters using [k-means++](https://wikipedia.org/wiki/K-means%2B%2B) (<https://wikipedia.org/wiki/K-means%2B%2B>), which results in faster convergence.

The TensorFlow k-Means API lets you choose either Euclidean distance or cosine distance as your [similarity measure](/machine-learning/clustering/similarity/measuring-similarity) (</machine-learning/clustering/similarity/measuring-similarity>).

[Previous](#)



[k-means Advantages and Disadvantages](#)

(</machine-learning/clustering/algorithm/advantages-disadvantages>)

[Next](#)

[Clustering Programming Exercise](#)



(</machine-learning/clustering/programming-exercise>)

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see the [Google Developers Site Policies](https://developers.google.com/site-policies) (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2020-02-10 UTC.