

Data Wrangling & Visualization: First Checkpoint

Interactive Visualization of UMAP Algorithm on User Data

Nikita Zagainov, Dmitry Tetkin, Nikita Tsukanov

March 2025

1 Results Of Data Collection & Preprocessing

As it was stated in project proposal, our project has little data processing, this is why our main focus during this checkpoint was on algorithm implementation.

2 Our Work So Far

During the first spring, our team adapted the original [UMAP](#) implementation to be able to stream intermediate results of the descent. Our contribution can be found in [this repository](#).

3 Demo

Figure [1](#) shows the visualization of the UMAP algorithm applied to [MNIST](#) dataset. The code to reproduce this visualization can be found in our repository, in the `notebooks/visualization_example.ipynb` notebook.

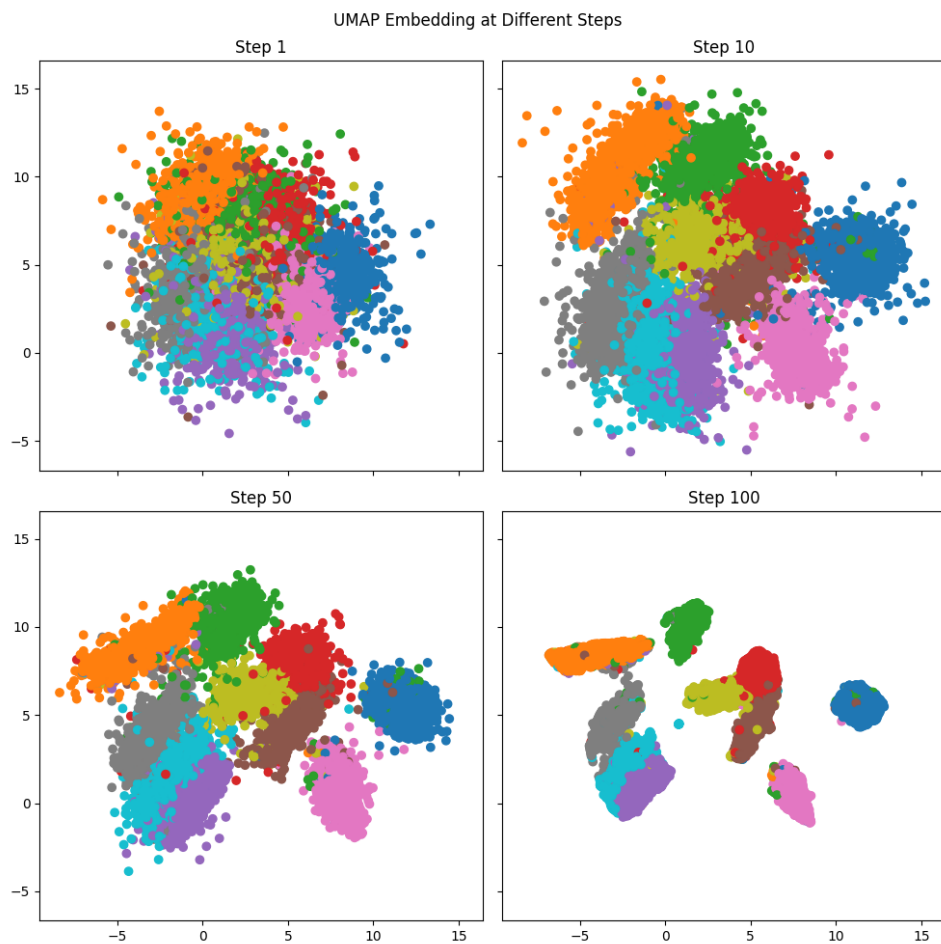


Figure 1: UMAP Steps Visualization on MNIST dataset. Different colors correspond to different labels (digits), algorithm had no access to true labels during learning embedding