

Analiza porównawcza zastosowań tradycyjnych oraz wariacyjnych autoenkoderów

Filip Ręka

Uniwersytet Marii Curie Skłodowskiej

24 stycznia 2022

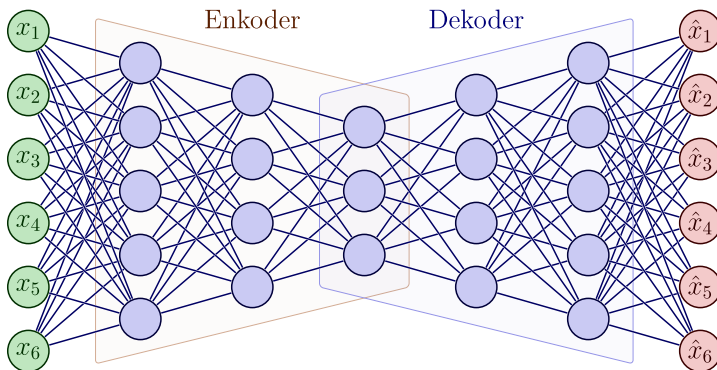
- Analiza możliwości i zastosowań autoenkoderów
- Możliwości generacji nowych danych przy użyciu autoenkodera
- Charakterystyka architektury autoenkodera wariacyjnego
- Analiza możliwości generacji danych przez autoenkoder wariacyjny
- Ocena możliwości zastosowania autoenkoderów wariacyjnych do generacji obrazów

Autoenkoder składa się z dwóch części: enkodera oraz dekodera. Obie z nich są w pełni połączonymi sieciami neuronowymi, które również są połączone pomiędzy sobą. Zadaniem enkodera jest zakodowanie jego wejścia do kodu o stałej długości, co jest osiągnięte przez architekturę warstw, w której każda następna ma mniej neuronów niż poprzednia. Dekoder jest częścią, która z kodu, próbuje odwzorować zakodowane dane. Cały autoenkoder dostaje na warstwę wejściową oraz wyjściową te same dane, przez co jego sposób uczenia jest nazywany częściowo nadzorowanym.

Architektura autoenkodera

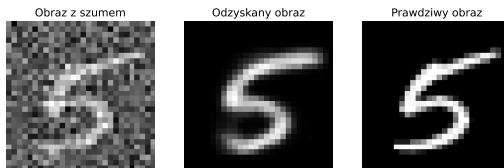
wejście

wyjście



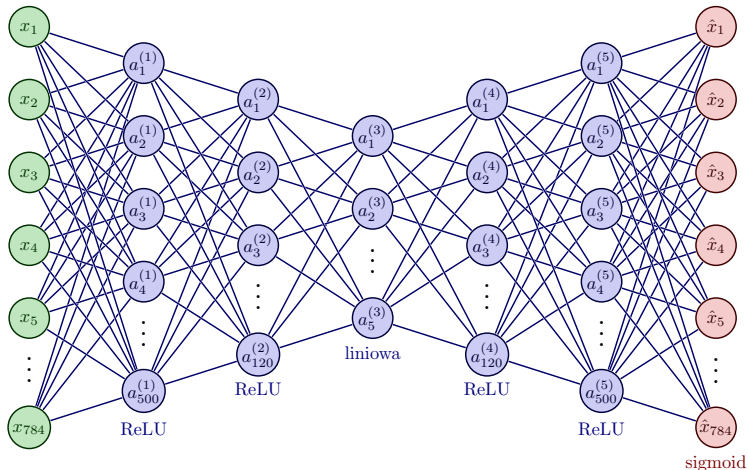
Zastosowania autoenkodera

Głównym zadaniem autoenkodera jest kompresja. Funkcję tą, można zastosować do redukcji wymiarów czy odsumiania danych, co zawdzięczamy temu, że kompresja dokonywana przez model jest stratna. W ten sposób autoenkoder, musi zachować w kodzie jak najwięcej istotnych elementów.



Rysunek: Rekonstrukcja obrazu przy pomocy autoenkodera

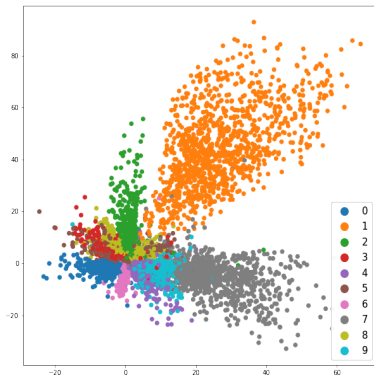
Zaletą używania autoenkoderów w celu redukcji wymiarów w przeciwieństwie do metody analizy składowych głównych (PCA) jest ich możliwość nauki nieliniowych zależności pomiędzy wymiarami.



Odszumianie obrazu zostało wykonane przy pomocy powyższego modelu, składającego się dwóch warstw ukrytych w enkoderze i dekoderze po 120 i 500 neuronów. Funkcją aktywacji użytą w modelu jest ReLU, poza warstwą kodu, w której jest to funkcja liniowa aby nie ograniczać wartości kodu i w ostatniej sigmoid, ponieważ dane wejściowe są w przedziale 0-1.

Aby wygenerować nowe dane, podobne do tych na których model został wytrenowany, potrzebna jest ściśle określonej dystrybucji według której punkty, reprezentujące skompresowane wejście, są rozłożone. Tradycyjna architektura nie zapewnia tego faktu, ani takich wymagać jak kompletności oraz ciągłości przestrzeni kodu. Przed treningiem nie wiadomo czy losowo wybrany punkt z przestrzeni, będzie generował odpowiednie dane.

Przestrzeń kodu AE

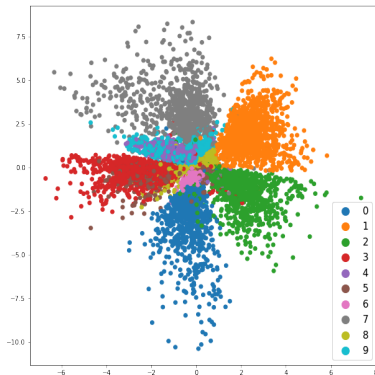


Obrazek obok przedstawia dwuwymiarową przestrzeń zmiennych ukrytych dla zbioru danych MNIST. Jak widać, nie jest ona kompletna ani ciągła. Wybierany, a następnie odkodowany punkt z miejsca pomiędzy pomarańczowymi a szarymi punktami, nie będzie podobny do obrazów ze zbioru danych.

Wariacyjny autoenkoder - wprowadzenie

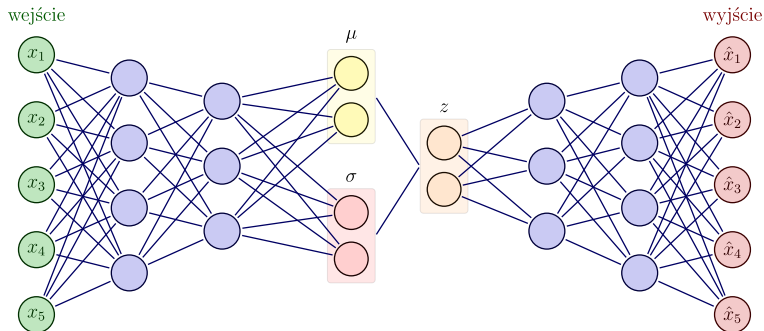
Wariacyjny autoenkoder rozwiązuje problemy z generacją nowych danych tradycyjnego modelu. Aby model był generacyjny, potrzeba ściśle ustalonej dystrybucji, z której losuje się punkty, na podstawie których generowane są dane. W przypadku modelu VAE, jest to wielowymiarowy rozkład normalny, opisywany za pomocą wektorów odchylenia standardowego σ oraz średniej μ . Następnie z tych dystrybucji losowane są punkty z czyli kod, na podstawie, którego model dokonuje rekonstrukcji wejścia. Aby wynikiem enkodera były punkty należące do wybranej dystrybucji, trzeba zmodyfikować funkcję straty. Poza błędem rekonstrukcji, obliczana jest również dywergencję Kullbacka-Leiblera, która określa rozbieżność pomiędzy dwoma rozkładami prawdopodobieństwa.

Przestrzeń kodu VAE



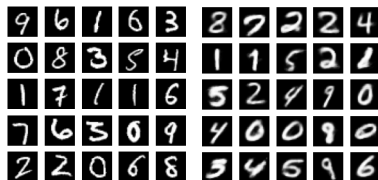
Rysunek przedstawia dwuwymiarową przestrzeń zmiennych ukrytych dla zbioru MNIST. Jest to przestrzeń wygenerowana przy pomocy modelu VAE, dlatego, jak widać, jest ona dwuwymiarowym rozkładem normalnym. Tym razem można wybrać z niej losowy punkt i mieć pewność, że zostanie on poprawnie odekodowany.

Architektura wariacyjnego autoenkodera



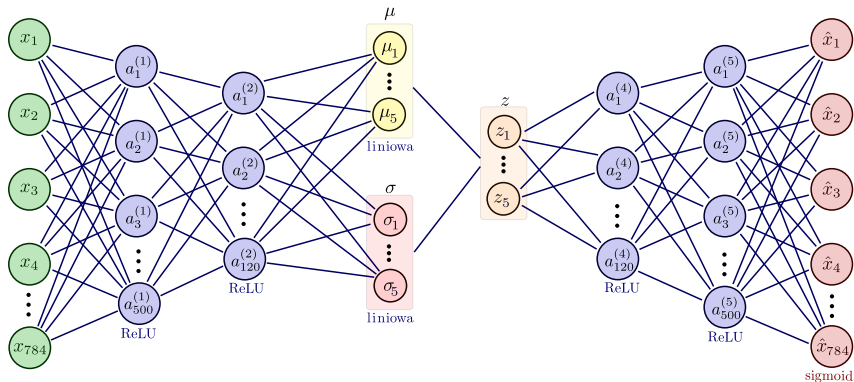
Generacja nowych danych

Posiadając enkoder, który generuje kod należący co ściśle określonej dystrybucji, można wybierając z niej losowe punkty generować dane.



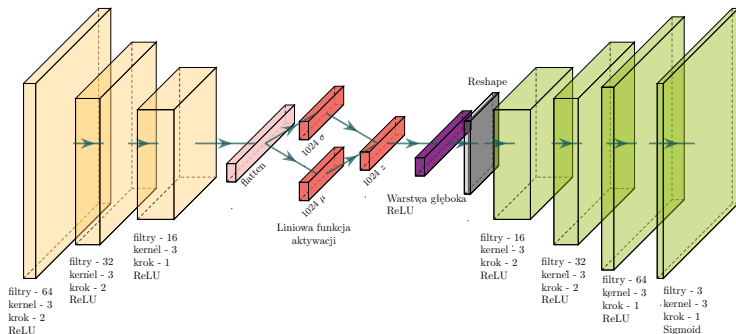
Rysunek: Po lewej stronie są dane ze zbioru, a po prawej wygenerowane

Powyższy przykład przedstawia generację na podstawie zbioru danych MNIST.



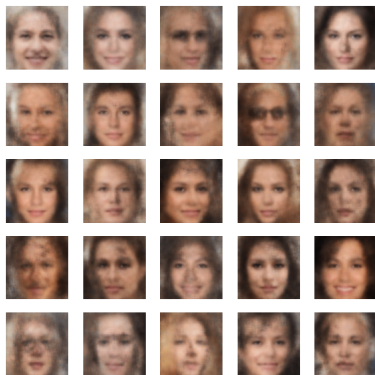
Generacja danych na podstawie zbioru MNIST odbyła się z użyciem powyższego modelu. Ilość neuronów w warstwach oraz funkcje aktywacji są takie same, jak w modelu AE przeznaczonym do odsumowania obrazów. Po treningu wystarczy na warstwę wejściową dekodera przekazać punkt losowo wybrany z pięciowymiarowej normalnej dystrybucji, przez co wynikiem modelu będzie nowy obraz.

Konwolucyjny wariacyjny autoenkoder



Sieci konwolucyjne można wykorzystać w architekturze modelu VAE, który będzie przeznaczony do pracy ze skomplikowanymi obrazami. Powyższa architektura została wytrenowana na zbiorze danych twarzy celebrytów.

Generacja twarzy



Twarze zostały wygenerowane przy użyciu modelu konwolucyjnego VAE. Jak widać na obrazie, model jest w stanie dość dobrze wygenerować podstawowe elementy twarzy jak oczy, nos, usta. W zależności od wylosowanych wartości, twarz może mieć uśmiech, zamknięte usta, różne odcienie koloru skóry lub koloru włosów.

Autoenkoder tradycyjny oraz wariacyjny są podobnymi modelami uczenia maszynowego. Składają się z dwóch części: enkodera próbującego zakodować zmienne do kodu o określonej długości oraz dekodera rekonstruującego kod do danych wejściowych. Wariacyjny autoenkoder, zamiast generować zmienne bezpośrednio, wybiera je z wielowymiarowego rozkładu normalnego. Jest to sposób na rozwiązanie problemów z generacją nowych danych tradycyjnego modelu. Porównując obie przestrzenie kodu, jasno widać, że model VAE, nauczył się generować zmienne ukryte według wybranej dystrybucji, która jest ciągła i kompletna. Dzięki temu, można wybrać z niej losowy punkt i na jego podstawie wygenerować nowe dane.

Dziękuję za uwagę!