

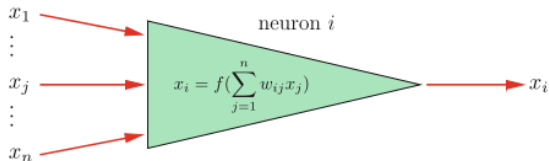
Przegląd autoenkoderów stosowanych w  
nienadzorowanym uczeniu maszynowym  
**Praca magisterska**

Alicja Hołowiecka

*pod kierunkiem  
dra Dariusza Majerka*

Lublin 2022

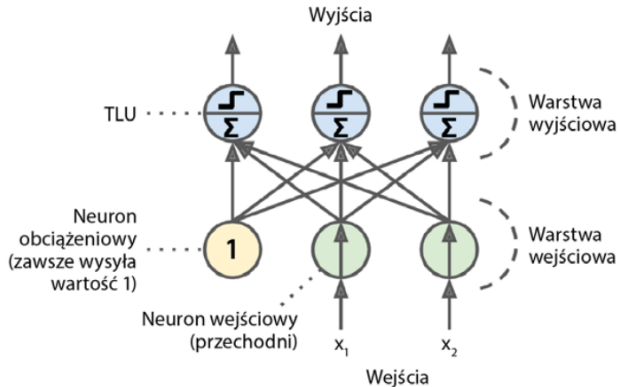
# Sztuczny neuron



**Rysunek:** Struktura sztucznego neuronu, który stosuje funkcję skokową  $f$  na ważonej sumie sygnałów wejściowych

Źródło: [2]

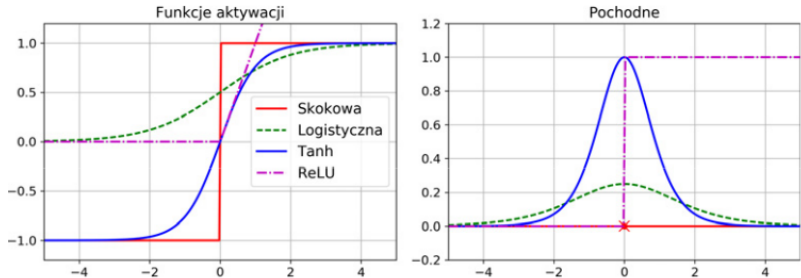
# Warstwa gęsta



**Rysunek:** Perceptron z trzema neuronami wejściowymi i trzema wyjściami

Źródło: [3]

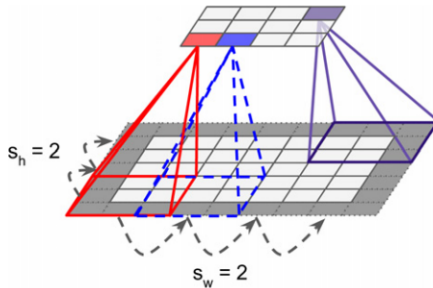
# Funkcje aktywacji



Rysunek: Przykładowe funkcje aktywacji wraz z pochodnymi

Źródło: [3]

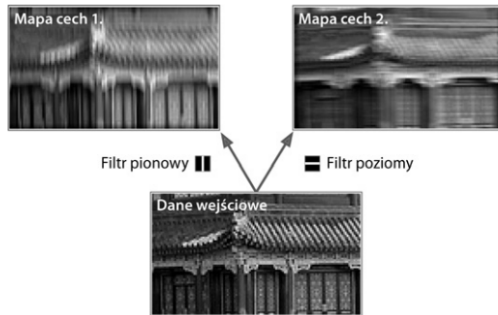
# Warstwa splotowa



Rysunek: Warstwa splotowa z krokiem o długości 2

Źródło: [3]

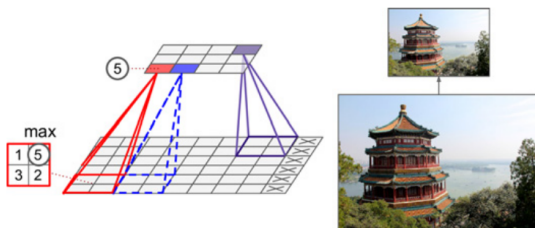
## Filtry i mapy cech



**Rysunek:** Uzyskiwanie dwóch map cech za pomocą dwóch różnych filtrów

Źródło: [3]

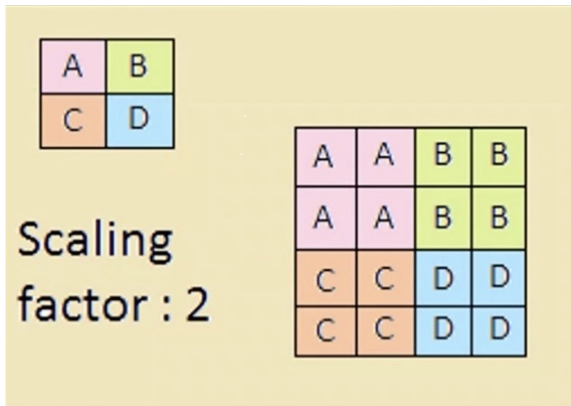
## Warstwa łącząca



**Rysunek:** Maksymalizująca warstwa łącząca (jądro łączące:  $2 \times 2$ , krok: 2, brak uzupełniania zerami)

Źródło: [3]

## Warstwy ekspansji

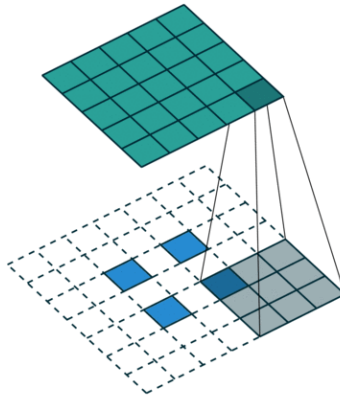


Rysunek: Działanie warstwy ekspansji (nadpróbkiwania)

Źródło: [1]



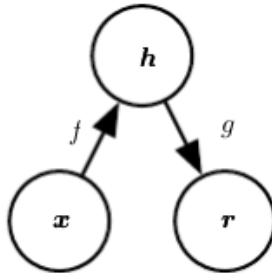
# Warstwy splotowe transponowane



**Rysunek:** Działanie warstwy splotowej transponowanej

Źródło: [5]

## Definicja autoenkodera



**Rysunek:** Struktura autoenkodera odwzorowującego wejście  $x$  na wyjście  $r$  (nazywane rekonstrukcją) poprzez reprezentację ukrytą (kodowanie)  $h$ . Autoenkoder składa się z dwóch składników: kodera  $f$  (odwzorowującego  $x$  na  $h$ ) i dekodera  $g$  (odwzorowującego  $h$  na  $r$ )

Źródło: [4]

## Rodzaje autoenkoderów

- autoenkodery niedopełnione (ang. *undercomplete*), w których wyjście musi mieć mniejszy wymiar niż wejście
- autoenkodery z regularyzacją (ang. *regularized*), w których wyjście ma taki sam lub większy (ang. *overcomplete*) wymiar niż wejście, ale używają specjalnie dopasowanych funkcji straty. Wśród autoenkoderów z regularyzacją można rozróżnić na przykład:
  - autoenkodery rzadkie (ang. *sparse*), które dążą do rzadkiej reprezentacji ukrytej
  - autoenkodery odszumiające (ang. *denoising*), które na wejściu dostają zniekształcone dane, a starają się odzyskać pierwotne, niezaszumione informacje
  - autoenkodery kurczliwe (ang. *contractive*) dążące do małego rozmiaru pochodnej

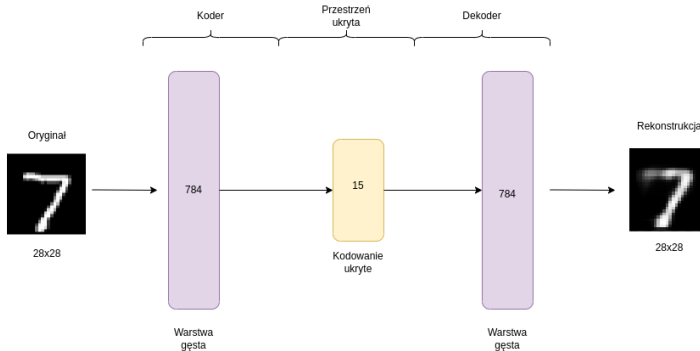
## Rodzaje autoenkoderów

- autoenkodery stosowe (ang. *stacked*) nazywane również głębokimi (ang. *deep*)
- autoenkodery splotowe (ang. *convolutional*)
- autoenkodery rekurencyjne (ang. *recurrent*)
- autoenkodery wariacyjne (ang. *variational*)
- autoenkodery przeciwstawne (ang. *adversarial*)

## Zastosowania autoenkoderów

- 1 Zmniejszenie wymiarowości (ang. *Dimensionality Reduction* )
- 2 Wyodrębnianie cech (ang. *Feature Extraction*)
- 3 Odszumianie obrazu (ang. *Image Denoising* )
- 4 Kompresja obrazu (ang. *Image Compression*)
- 5 Wyszukiwanie obrazu (ang. *Image Search* )
- 6 Wykrywanie anomalii (ang. *Anomaly Detection*)
- 7 Uzupełnianie brakujących danych (ang. *Missing Value Imputation* )

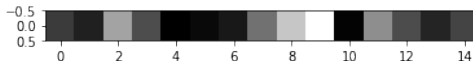
# Struktura prostego autoenkodera



**Rysunek:** Struktura autoenkodera prostego (niedopełnionego)

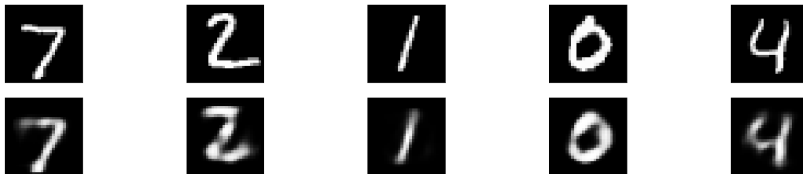
*Źródło:* Opracowanie własne

# Reprezentacja ukryta i rekonstrukcje



**Rysunek:** Reprezentacja 15-wymiarowa dla pierwszej obserwacji ze zbioru testowego MNIST

*Źródło:* Opracowanie własne



**Rysunek:** Pięć pierwszych obserwacji ze zbioru testowego MNIST: w górnym rzędzie oryginalne obrazy, w dolnym rekonstrukcje z autoenkodera niedopełnionego

*Źródło:* Opracowanie własne

# Struktura autoenkodera splotowego

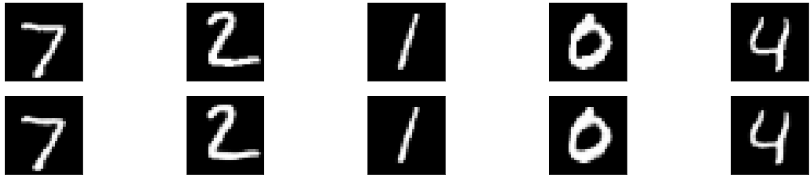


Rysunek: Struktura autoenkodera splotowego

Źródło: Opracowanie własne



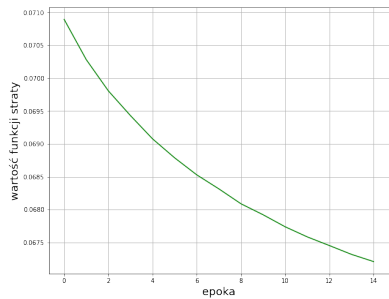
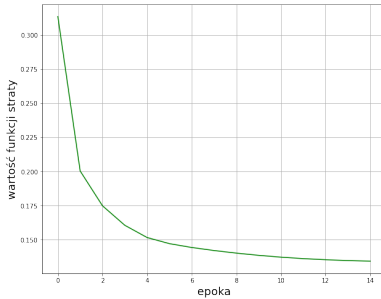
## Rekonstrukcje z autoenkodera splotowego



**Rysunek:** Pięć pierwszych obserwacji ze zbioru testowego MNIST: w górnym rzędzie oryginalne obrazy, w dolnym rekonstrukcje z autoenkodera splotowego

*Źródło:* Opracowanie własne

# Porównanie funkcji strat z autoenkodera prostego i splotowego



**Rysunek:** Porównanie straty z autoenkodera prostego (po lewej) i splotowego (po prawej)

**Źródło:** Opracowanie własne

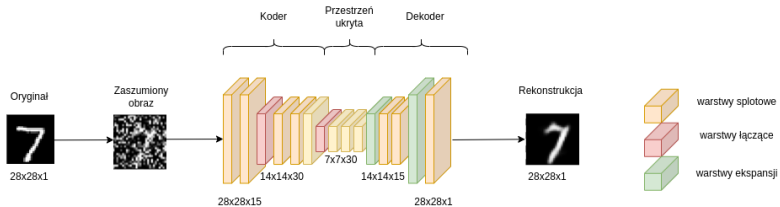
## Zaszumione obrazy



**Rysunek:** Zaszumione obrazy ze zbioru testowego MNIST

*Źródło:* Opracowanie własne

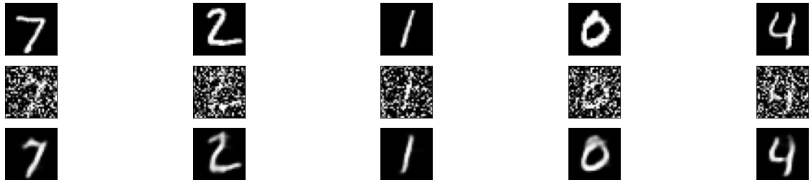
# Struktura autoenkodera odszumiającego



Rysunek: Struktura autoenkodera odszumiającego

Źródło: Opracowanie własne

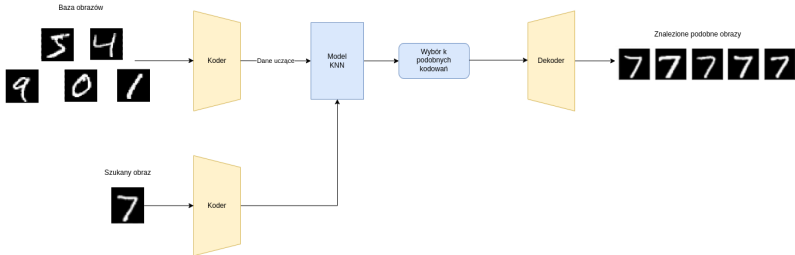
## Odszumione rekonstrukcje



**Rysunek:** Pięć pierwszych obserwacji ze zbioru testowego MNIST: w górnym rzędzie oryginalne obrazy, w środkowym zaszumione, w dolnym rekonstrukcje z autoenkodera odsumiającego

*Źródło:* Opracowanie własne

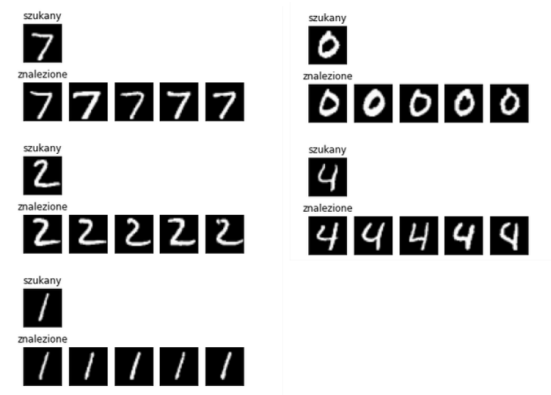
## Wyszukiwanie obrazu



**Rysunek:** Schemat działania wyszukiwania obrazów przy użyciu autoenkodera

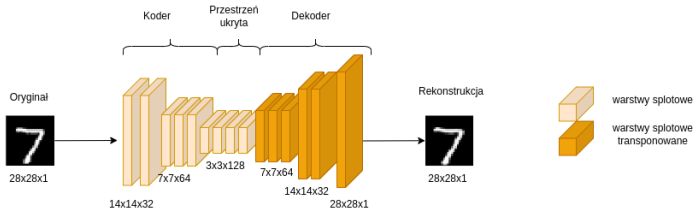
**Źródło:** Opracowanie własne

## Wyszukiwanie obrazu



**Rysunek:** Wyniki zastosowania prostego autoenkodera przy wyszukiwaniu obrazu

# Struktura autoenkodera użytego do wykrywania anomalii

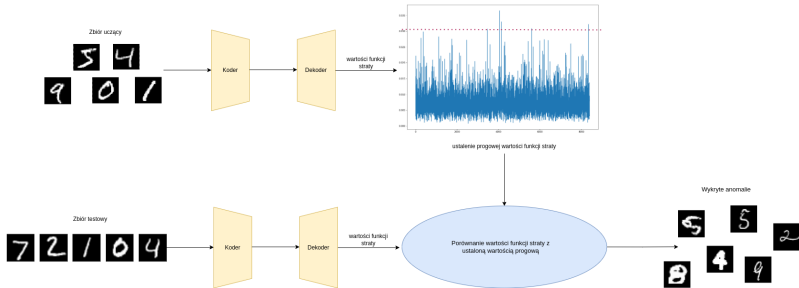


**Rysunek:** Struktura autoenkodera splotowego użytego do wykrywania anomalii

*Źródło:* Opracowanie własne



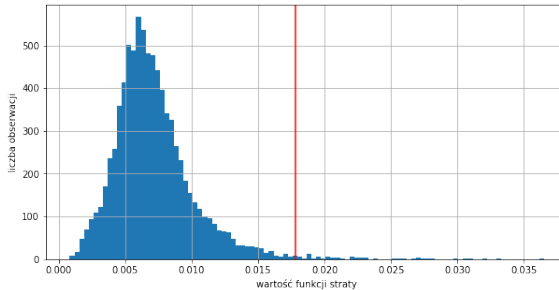
# Wykrywanie anomalii



**Rysunek:** Schemat wykrywania anomalii przy użyciu autoenkodera

**Źródło:** Opracowanie własne

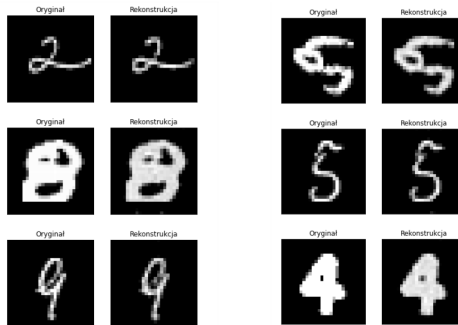
# Graniczna wartość straty



Rysunek: Histogram funkcji straty

Źródło: Opracowanie własne

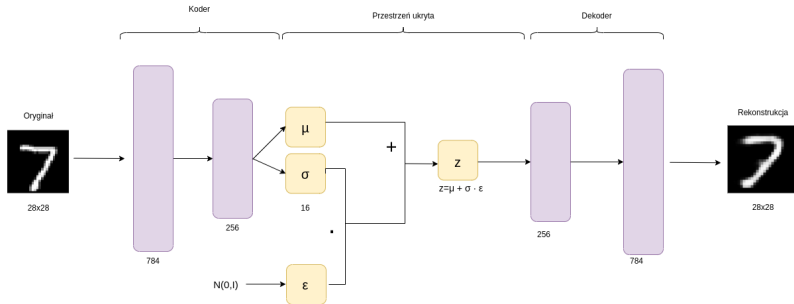
## Przykłady wykrytych anomalii



**Rysunek:** Przykłady anomalii w zbiorze MNIST wykrytych za pomocą autoenkodera

*Źródło:* Opracowanie własne

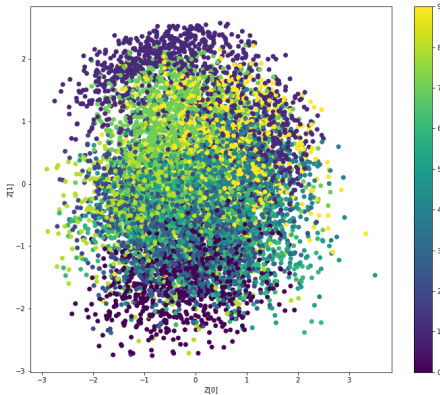
# Struktura autoenkodera wariacyjnego



Rysunek: Struktura autoenkodera wariacyjnego

Źródło: Opracowanie własne

# Przestrzeń ukryta autoenkodera wariacyjnego



Rysunek: Przestrzeń ukryta autoenkodera wariacyjnego

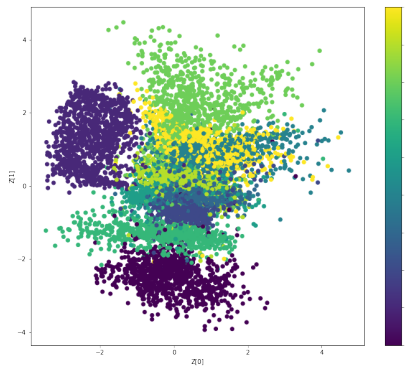
Źródło: Opracowanie własne

## Przykłady wygenerowanych obrazów



**Rysunek:** Przykłady obrazów cyfr wygenerowanych przez autoenkoder wariacyjny

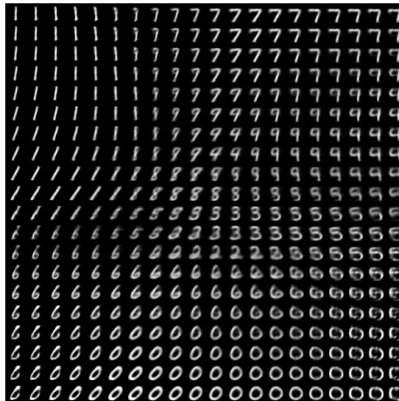
# Autoenkoder wariacyjny



**Rysunek:** Przestrzeń ukryta autoenkodera wariacyjnego z wymiarem kodowania ukrytego równym dwa

*Źródło:* Opracowanie własne

# Autoenkoder wariacyjny



**Rysunek:** Przestrzeń ukryta w postaci obrazów cyfr

Źródło: Opracowanie własne





Artificial Intelligence in Plain English, *Convolutional Autoencoders (CAE) with Tensorflow*

<https://ai.plainenglish.io/>

[convolutional-autoencoders-cae-with-tensorflow-97e8d8859](https://ai.plainenglish.io/convolutional-autoencoders-cae-with-tensorflow-97e8d8859)

(dostęp: 15.05.2022)



Ertel W. (2017) *Introduction to Artificial Intelligence. Second Edition*, Springer International Publishing



Géron A. (2020) *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow. Wydanie II*, Helion SA



Goodfellow I., Bengio Y., Courville A. (2018), *Deep Learning. Systemy uczące się*, PWN, Warszawa



Pröve P. L., *An Introduction to different Types of Convolutions in Deep Learning*, <https://towardsdatascience.com/>

types-of-convolutions-in-deep-learning-717013397f4d  
(dostęp: 15.05.2022)