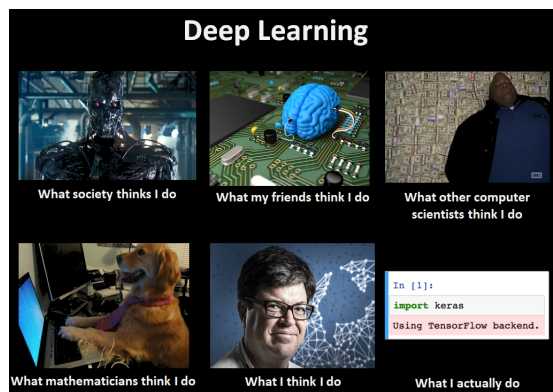


# Sieci Neuronowe 2018: Zagadnienia do Egzaminu



Odpowiadające danemu blokowi ćwiczenia i wykłady są zaznaczone w tytule danego bloku.

## 1 Podstawowe pojęcia (ćw. 1-2; wykł. 1)

- Modele uczenia maszynowego: nadzorowane, nienadzorowane oraz seminadzorowane.
- Regresja a klasyfikacja: definicje problemu; klasyfikacja binarna / wieloklasowa (multi-class) / wielo-etykietowa (multi-label).
- Modele liniowe. Funkcje kosztu dla problemów regresji/klasyfikacji. Liniowa separowalność w klasyfikacji. Wypukłość rozwiązania modelu liniowego, podstawowa postać rozwiązania.

## 2 PyTorch

- Ogólna architektura pakietu PYTORCH. Podstawowe moduły.
- Rozróżnienie między klasami VARIABLE a TENSOR.
- Struktura typowej (jak na ćwiczeniach) pętli trenowania.
- Implementacja wstecznej propagacji błędów w pakiecie PYTORCH.

### 3 Optymalizacja (ćw. 4-5, 7, 10; wykł. 2, 3, 4)

#### 3.1 Podstawowe pojęcia

- Uczenie gradientowe, wsteczna propagacja, reguła delta.
- Tryby uczenia gradientowego: batch, mini-batch, stochastic, rozróżnienie pojęć, wpływ na uczenie.
- Wsteczna propagacja błędów. W szczególności rozpisanie różniczkowania (narysowanie grafu obliczeń) dla modelu binarnej regresji logistycznej.

#### 3.2 Optymalizatory

- Algorytmy Stochastic Gradient Descent (SGD) i Gradient Descent (GD). Wzory. Zachowanie SGD oraz GD na problemie o wypukłej funkcji kosztu.
- Metoda momentum. Interpretacja oraz wpływ na uczenie sieci neuronowych.
- Metody adaptacyjne na przykładzie algorytmu RMSProp.
- Wpływ kroku uczenia i rozmiaru próbki na SGD (m.in. szybkość uczenia, generalizację oraz szum dodany do estymacji gradientu).

#### 3.3 Zanikający i eksplodujący gradient (wykł. 2, 7)

- Problem zanikającego oraz eksplodującego gradientu. Przykłady.
- Przykładowe sposoby obejścia zanikającego gradientu.

### 4 Sieci neuronowe - podstawy (ćw. 6; wykł. 2, 3, 4)

- Podstawowe funkcje kosztu: entropia krzyżowa oraz błąd średniokwadratowy.
- Budowa sieci neuronowych. Definicja warstw oraz neuronów.
- Funkcje aktywacji dla neuronów warstw wyjściowych i ukrytych. Ich możliwy wpływ na przebieg uczenia. Funkcje softmax, sigmoidalne oraz ReLU.
- Podstawowe aspekty inicjalizacji sieci. Czemu inicjalizujemy sieci neuronowe liczbami losowanymi z danego rozkładu. Wzór Glorota na inicjalizację sieci.

### 5 Regularyzacja (ćw 8; wykł. 5)

- Dropout oraz Batch Normalization (BN). Motywacja. Różnica względem pełnej normalizacji. Jak różni się działanie BN oraz Dropoutu w trakcie uczenia i w trakcie inferencji (testowania)?
- Praktyka stosowania Dropoutu i Batch Normalization w sieciach neuronowych. Wpływ na szybkość uczenia oraz generalizację.

- Regularyzacja uczenia przez czynnik kary w funkcji kosztu. Regularyzacja L2 i L1 i co każda z nich daje w rozwiązaniu? Różnica pomiędzy regularyzacją L2, a “weight decay”.
- Augmentacja danych. Przykłady. Typowe sposoby augmentacji obrazów.

## 6 Sieci konwolucyjne (ćw. 9, wykł. 6)

- Podstawy sieci konwolucyjnych. Wzory definiujące wyjście warstwy konwolucyjnej i warstwy pooling. Kiedy należy stosować warstwy konwolucyjne w sieciach neuronowych?
- Praktyka używania warstw konwolucyjnych w sieciach neuronowych, np. jak używać warstw konwolucyjnych w połączeniu z Batch Normalization.
- Wpływ translacji i obrotu obrazu na wynik warstwy konwolucyjnej.

## 7 Sieci rekurencyjne (ćw. 10; wykł. 7)

- Podstawy sieci rekurencyjnych. Podstawowy model (taki jak na ćwiczeniach). Przykłady zastosowań sieci rekurencyjnych.
- Algorytm trenowania sieci rekurencyjnych Back Propagation Through Time oraz Truncated Back Propagation Through Time.
- Modele Long Short Term Memory (LSTM) oraz Gated Recurrent Unit (GRU). Wzory. Motywacja stojący za tymi modelami. Interpretacja funkcji bramek.

## 8 Uczenie nienadzorowane (ćw. 11; wykł. 8, 9)

- Pojęcie oraz przykłady zastosowań uczenia nienadzorowanego.
- Model Self-Organising Map (SOM). Algorytm trenowania. Funkcja kosztu.
- Pojęcie oraz zastosowanie pretrainingu w Deep Learningu.
- Modele autoenkoder (AE) oraz Denoising Autoencoder (DAE). Wzory. Zastosowania autoenkoderów. Związek modelu autoenkoder z modelem analizy głównych składowych (Principal Component Analysis).
- Sieci autoasocjacyjne na przykładzie Sieci Hopfielda. Reguła Hebba. Wzór na energię oraz ewolucję stanu sieci. Czym jest ”pojemność” sieci w modelu Hopfielda? Przykład wykorzystania modelu Hopfielda dla problemu Travelling Salesman.