Deep Learning

project

Stanisław Kaźmierczak, Maciej Żelaszczyk, Adam Żychowski

Faculty of Mathematics and Information Science,
Warsaw University of Technology

2022

Overview

1. Assessment
2. Project I - convolutional neural networks
3. Project II - recurrent neural networks
4. Project III - generative adversarial networks

Overview

- **3 projects**
- Project I (35 points), Project II (35 points), Project III (30 points)
- groups of 2 people
- first week of delay: -5 points; second week of delay: additional -10 points (-15 points in total); exceeding 2 weeks of delay results in 0 points for the project
- grades: [0;50] $\rightarrow$ 2, [51;60] $\rightarrow$ 3, [61;70] $\rightarrow$ 3.5, [71;80] $\rightarrow$ 4, [81;90] $\rightarrow$ 4.5, [91;100] $\rightarrow$ 5,

Timetable

1.	Mar. 1	**Introduction to Project 1**
2.	Mar. 8	Project 1 plan (presence not mandatory)
3.	Mar. 15	Tutorial hours
4.	Mar. 22	Tutorial hours
5.	Mar. 29	Initial presentation of Project 1
6.	Apr. 5	**Project 1 deadline, introduction to Project 2**
7.	Apr. 12	Project 2 plan (presence not mandatory)
8.	Apr. 26	Tutorial hours
9.	May 4	Initial presentation of Project 2
10.	May 10	**Project 2 deadline, introduction to Project 3**
11.	May 17	Project 3 plan (presence not mandatory)
12.	May 24	Tutorial hours
13.	May 31	Tutorial hours
14.	Jun. 7	Initial presentation of Project 3
15.	Jun. 14	**Project 3 deadline**

Rules

- tutorial hours are optional
- other classes are obligatory
- absence from mandatory classes results in -2 points
- let me know a day before (no later than 6p.m.) if you plan to attend tutorial hours
- source code and report should be sent before the beginning of class
- you can utilize code from external sources (books, articles, blogs) provided that:
  - reference is cited in the report
  - some modifications to the original solution are applied
- violation of the above or any other kind of plagiarism results in a failing grade

Project guidelines

General

- take care of reproducibility by initializing a random number generator with a constant seed
- training models on other datasets than mentioned in the task description is not allowed
- application of pre-trained models (AlexNet, VGG, etc.) is permitted (and even recommended as a part of the experiments)
- full discretion in terms of libraries
- crucial elements taken into account while assessment:
  - report quality
  - achieved accuracy registered on Kaggle
  - the originality of the solution
- present your experiments and results during a 10-minute slide show

Reports

The report should include:

- description of the research problem, understandable to the person who did not see the content of the task
- instruction of the application (containing information on how to reproduce results)
- theoretical introduction
- description of the conducted experiments
- statistically processed results (presented clearly)
- conclusions, presumed reasons for successes/failures and further research proposals

Reports

Some additional remarks:

- if the experiment is not described in the report it is regarded as not conducted
- the report is an official document, so please keep it formal (table of contents, bibliography, captions under figures, tables, etc.)
- results should be commented
- to obtain statistically significant results, each experiment ought to be repeated multiple times (when possible)
- in addition to the mean, standard deviation should also be calculated (in some scenarios worse mean with low variance may be a more desirable result than a better mean with high variance)

Project I - convolutional neural networks

Topic: **Image classification with convolutional neural networks**

Dataset: CIFAR-10
`https://www.kaggle.com/c/cifar-10/`

Project I - convolutional neural networks

- Test and compare different network architecture (at least one should be a convolutional neural network)
- Investigate influence of the following hyper-parameter change on obtained results:
  - At least 2 hyper-parameters related to training process
  - At least 2 hyper-parameters related to regularization
- Investigate influence of at least X data augmentation techniques from the following groups:
  - Standard operations (where x=3)
  - More advanced data augmentation techniques like mixup, cutmix, cutout (where x=1)
- Consider application of ensemble (hard/soft voting, stacking)

Assessment: source code, report, presentation

Project I - convolutional neural networks

Useful resources:

- Google Colab or GPU to speed up training
- https://machinelearningmastery.com/voting-ensembles-with-python/
- https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/
- https://benchmarks.ai/cifar-10
- https://www.robots.ox.ac.uk/ vgg/practicals/cnn/index.html
- https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner’s-Guide-To-Understanding-Convolutional-Neural-Networks/
- https://medium.com/kaggle-blog/profiling-top-kagglers-bestfitting-currently-1-in-the-world-58cc0e187b

Project II - recurrent neural networks

Topic: **Speech commands classification with recurrent neural networks**

Dataset: Speech Commands Dataset
`https://www.kaggle.com/c/tensorflow-speech-recognition-challenge/data`

Project II - recurrent neural networks

- test and compare different network architectures (at least one of them should be Long short-term memory (LSTM))
- investigate influence of parameters change on the obtained results
- present confusion matrix (with appropriate discussion)
- in case of accuracy or efficiency problem a subset of classes can be selected and tested (e.g. only "yes" and "no" commands)
- please pay special attention on "silence" and "unknown" classes - test different approaches (e.g. separate network for their recognition)

Assessment: source code, report, presentation

Project II - recurrent neural networks

Useful resources:

- https://www.kaggle.com/davids1992/speech-representation-and-data-exploration
- https://towardsdatascience.com/recognizing-speech-commands-using-recurrent-neural-networks-with-attention-c2b2ba17c837
- https://www.coursera.org/lecture/nlp-sequence-models/recurrent-neural-network-model- ftkzt
- https://pathmind.com/wiki/lstm

Project III - generative adversarial networks

Topic: **Image generation with generative adversarial networks**

Dataset: 10% sample of the LSUN Bedrooms Dataset
`https://www.kaggle.com/jhoward/lsun_bedroom`

The whole dataset is $\sim$ 4.7GB in total.

Project III - generative adversarial networks

- test and compare different network architectures (at least one of them should converge to generate satisfactory images)
- potential architectures (can use different ones): vanilla GAN with CNNs, DC-GAN, Wasserstein GAN, WGAN-GP, SN-GAN, StyleGAN1, StyleGAN2 (with or without data augmentations), StyleGAN3
- calculate the Fréchet Inception Distance (FID) for your generated images and compare it to results from literature
- assess your results qualitatively
- investigate the influence of hyperparameters on obtained results

Project III - generative adversarial networks

- discuss sets of hyperparameters which help in overcoming training collapse and mode collapse
- select two of your generated images together with their latent vectors; interpolate linearly between the two latent vectors to generate 8 additional latent vectors; use these 8 vectors to generate images from your model; present the 10 generated images (8 newly generated and 2 generated previously) and discuss the importance of the results
- discuss any additional findings
- no required language or libraries; suggested language: Python; suggested libraries: PyTorch or TensorFlow

Assessment: source code, report, presentation

Project III - generative adversarial networks

Useful resources:

- Original GAN paper: https://arxiv.org/pdf/1406.2661.pdf
- DC-GAN paper: https://arxiv.org/pdf/1511.06434.pdf
- StyleGAN3: https://arxiv.org/pdf/2106.12423.pdf
- FastAI on GANs (skip to relevant part): https://course18.fast.ai/lessons/lesson12.html
- Machine Learning Mastery (general overview): https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/
- DeepLearning.AI (for those interested; ability to access content for free): https://deeplearning.ai/program/generative-adversarial-networks-gans-specialization/

The End