**Course project MLT 2025**

**"The study of the comparative efficiency of machine learning algorithms in solving specific problem"**

(result - **report**)

1. Select a task:

Choose a topic according to the theme of your Master research in the University (preferable) or select from the list below:

https://www.kaggle.com/datasets/hojjatk/mnist-dataset

https://www.kaggle.com/datasets/janiobachmann/bank-marketing-dataset

https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting/data

https://www.kaggle.com/c/rossmann-store-sales/data

2. Build a mathematical formalization of the problem (loss function, optimization algorithm etc.)

3. Describe the quality metric(s) by which the solution of the problem can be assessed.

4. Select at least 3 methods / algorithms of machine learning to study the comparative effectiveness of the methods of solving the task.

5. Compare the quality of algorithms with respect to the chosen metric(s).

6. Make a conclusion about which algorithm gives the best results for the task (and why).

**7. Report form:**

Report in English 5-10 pages (introduction, related works, problem statement, description of methods, data processing, experiments, conclusion).

Template is on the following pages.

**Ministry of Science and Higher education**

**of the Russian Federation**

**ITMO University**

Faculty of Digital Transformations

Subject area (major) 01.04.02. Applied mathematics and informatics

REPORT

(report title)

Student: Ivanov Ivan C42111c

Supervisor: Gladilin P.E. – or your research supervisor

Date 23.03.2024

St. Petersburg

2024

TABLE OF CONTENT

Оглавление пустое, так как стили абзацев, выбранные для отображения в оглавлении, не используются в документе.

## 1. Introduction

XXX

## 2. Literature review

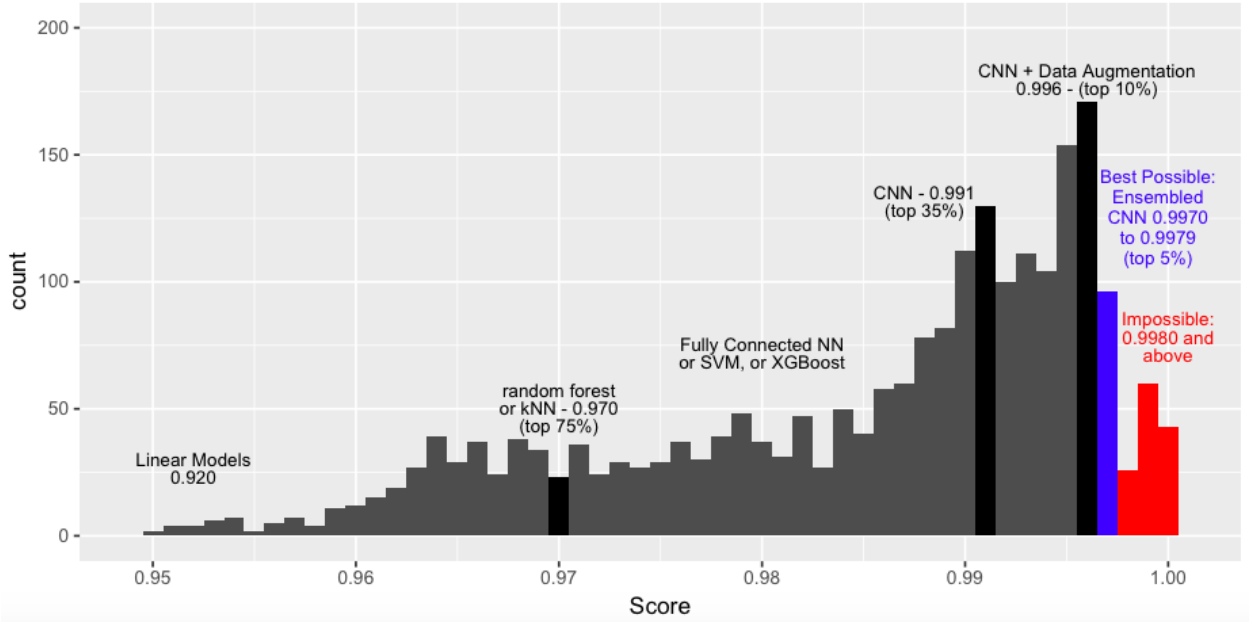


Figure 1 – (example of the figure caption) Kaggle MNIST public leaderboards scores, July 15 2018

## 3. Models, Algorithms and Datasets

## 4. Experimental research

XXX

Table 1 – *(example of the figure caption)* CNN-5 results

| Submission and description | Private | Public |
|---|---|---|
| cnn-5-baseline | 0.075 | 0.089 |
| cnn5_balanced | 0.006 | 0.009 |
| cnn5_balanced_20ep | 0.010 | 0.010 |
| cnn5_balanced_20ep_undersampling (16x16) | 0.010 | 0.012 |
| cnn5_64x64_balaced_20ep_undersampling | 0.046 | 0.050 |

## 5. Conclusions

XXX

**References (example)**


1. Tan M., Le Q. V. EfficientNet: Rethinking model scaling for convolutional neural networks // 36th Int. Conf. Mach. Learn. ICML 2019. 2019. Vol. 2019-June. P. 10691–10700.
2. Huang G. et al. Densely connected convolutional networks // Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017. 2017. Vol. 2017-January. P. 2261–2269.
3. Yun S. et al. CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features. 2019.
4. Zhang H. et al. MixUp: Beyond empirical risk minimization // 6th Int. Conf. Learn. Represent. ICLR 2018 - Conf. Track Proc. 2018. P. 1–13.
5. Zoph B. et al. Learning Transferable Architectures for Scalable Image Recognition // Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2018. P. 8697–8710.