МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Одеський національний технологічний університет
Університет Інформатики і прикладних знань, м.Лодзь, Польща
Національний технічний університет України «Київський політехнічний інститут»
Навчально-науковий інститут комп'ютерних систем і технологій «Індустрія 4.0» ім. П.М. Платонова

XXIII Всеукраїнська науково-технічна конференція молодих вчених, аспірантів та студентів

«СТАН, ДОСЯГНЕННЯ ТА ПЕРСПЕКТИВИ ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ»

Матеріали конференції

Одеса

20-21 квітня 2023 р.

Стан, досягнення та перспективи інформаційних систем і технологій / Матеріали XXIII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів. Одеса, 20-21 квітня 2023 р. - Одеса, Видавництво ОНТУ, 2023 р. – 449 с.

Збірник включає матеріали доповідей учасників конференції, які об'єднані за тематичними напрямками конференції.

Збірник буде корисним як для фахівців і працівників фірм, зайнятих в області ІТ, так і для викладачів, магістрів і студентів вищих навчальних закладів, які навчаються за напрямами і спеціальностями програмного забезпечення, обчислювальної техніки і автоматизованих систем, прикладної математики та обробки інформації, буде корисним професіоналам з комп'ютерного моделювання та розробки комп'ютерних ігор.

Результати досліджень у збірнику представляють собою своєрідний зріз сучасного стану справ в перерахованих галузях знань, який може допомогти як фахівцям, так і студентам університетів скласти загальну картину розвитку інформаційних технологій та пов'язаних з ними питань.

Наукові праці згруповані за напрямками роботи конференції та наведені в алфавітному порядку прізвищ авторів.

Матеріали (тези доповідей) друкуються в авторській редакції. Відповідальність за якість та зміст публікацій несе автор.

Матеріали подано українською та англійською мовами.

Редактор збірника Котлик С.В.

Experiments. The workability and effectiveness of the developed method and algorithms in the task of divergence ships from own ships with ships targets according to the criterion of minimum risk were checked on the Simulation and modeling stand created on the basis of Navi Trainer 5000 of the training laboratory of the Kherson State Maritime Academy.

Conclusion. The method of optimal divergence in the field of risks has been developed, which allows, in comparison with traditional methods, to minimize the length of the trajectory of divergence, provided that the given collision risk is not exceeded. The obtained result is based on the use of the automatic divergence module integrated into the on-board computer, which builds risk fields, calculates at each step of the calculation the gradient of the risk field at the point of the vessel and the direction of the vessel's movement, the perpendicular direction of the gradient tangent to the ellipse of equal risk at the point of the vessel. Forms controls that ensure the movement of the vessel along the given risk ellipse during the divergence process. The developed method can be applied in automatic divergence modules of automated vessel traffic control systems.

References

[1] P. Mamenko. Minimization of ships' passing path in the field of risks. *Technology Audit and Production Reserves, 2*(70). (2023).https://doi.org/10.15587/2706-5448.2023.276419

[2] P. Mamenko, S. Zinchenko, V. Kobets, P. Nosov and I. Popovych, «Solution of the Problem of Optimizing Route with Using the Risk Criterion,»*In: Babichev, S., Lytvynenko, V. (eds) Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2021. Lecture Notes on Data Engineering and Communications Technologies, Springer, Cham*, no. 77, pp. 252-265,2021.https://doi.org/10.1007/978-3-030-82014-5_17

[3] S. Zinchenko, O. Tovstokoryi, P. Nosov, I. Popovych, V. Kobets and G. Abramov, «Mathematical support of the vessel information and risk control systems,» *CEUR Workshop Proceedings*,no. 2805, pp. 335-354, 2020.http://ceur-ws.org/Vol-2805/paper25.pdfM. Abdelaal, M. Franzle and A. Hahn, «Nonlinear Model Predictive Control for trajectory tracking and collision avoidance of underactuated vessels with disturbances,»*Ocean Engineering*,no. 160, pp. 168–180, 2018.doi: 10.1016/j.oceaneng.2018.04.026

UDC 003.26

OPTIMIZATION PROBLEMS IN MACHINE LEARNING: GRADIENT DESCENT MODIFICATIONS

Y. FEDIAIEVA(yevheniia.fediaieva@stud.onu.edu.ua), **A. STEHUN**(angela.stehun@gmail.com)
Odesa I. I. Mechnikov National University

The following modifications of gradient descent were considered: SGD (Mini-Batch SGD), Learning Rate Control, Momentum SGD, Nesterov Momentum, Adagrad, Adadelta, RMSprop, Adam. The main idea of the second-order methods (Newton Momentum) and the problems of their usage were considered. Using them, a simple neural network was created and trained for the binary classification task, a comparison of the obtained results and the effectiveness of these methods was carried out.

To train many machine learning models, iterative processes are used, the goal of which is to reduce some error of a given model on the data. In the class of linear models, the ideas of Gradient Descent are widely popular for minimizing the cost function of the model, which is the average of the values of the error function on each dataset object. But when we try to transfer this idea to training neural networks, quite big problems arise:

Neural networks are trained on huge datasets, and one iteration of gradient descent over such a dataset takes a very long time. Also, given the Gradient Descent convergence issues, this approach is very suboptimal. Optimizing a large neural network, it turns out that the cost function is a very complex composition of other functions, there can be an exponentially large number of different extrema depending on the number of neurons.

In reality, this means that we can not know if we are really at the global minimum, or if it is somewhere else. Moreover, due to the possibility of overfitting, we can not even be sure if we really want to be in the global minimum.

But, although it turns out to be impossible to guarantee the exact solution of the optimization problem, we can try to find the best option from those that are available. Heuristic optimization methods are designed for this task, the main one being Gradient Descent. Almost all modern methods for optimizing non-convex functions are one or another of its modifications, some of which are considered in research.

In search of improvements, we will start from the basic idea of GD:

$$w_i = w_{i-1} - \alpha \nabla E(w),$$

where w – a vector of parameters, $E(w)$ – cost function, α – learning rate.

To solve the problem about the size of the dataset, let's try to do a GD step not over the entire dataset, but for each example (stochastic gradient descent). Obviously, we will get a strong acceleration of learning (the error is calculated quickly and the weights change immediately). There may even be a situation where the SGD has converged even before we have walked through the entire training set. It also works "more randomly": this leads to the fact that the probability of not getting stuck in small lows increases.

But in practice, updating the model weights after each example is too expensive, so you can try to find a compromise: let's step on some fixed size subsample (Mini-Batch SGD). Since this is usually a very small part of the training set, the speed and local minima selectivity advantages of the SGD are retained. This approach can and should be used with all GD modifications.

Carry on. We can see that the learning rate is only one parameter that should be tuned. But instead of choosing a specific value, we can dynamically change this parameter. The intuitive solution suggests that the speed should first be large in order to arrive at the correct area of the function landscape as quickly as possible, and then decrease in order to search for a minimum. There can be many strategies for controlling the learning rate, such as linear

$$\alpha = \alpha_0 \left(1 - \frac{t}{T}\right)$$

or exponential decay

$$\alpha = \alpha_0 e^{\frac{-t}{T}}$$

where t – the time elapsed since the beginning of training (the number of mini-batches or epochs), and T – a parameter that determines how quickly α will decrease. This approach greatly improves the capacity of the entire optimization algorithm.

Next, consider the following idea: we can use not only information about the gradient at the current point, but also about some previous gradients, to step with an idea of the function surface view in some area, and not just at the one point. This is implemented using an exponential moving average like that:

$$v_i = \beta v_{i-1} + (1 - \beta) g_i\ ,$$

where v –a vector with the information we need, which will then be used for updating the weights, g – gradients, β – hyperparameter that should be tuned.

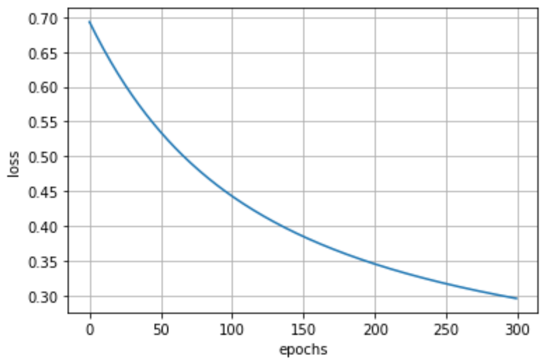
These are basic ideas that can be applied in different combinations with different variables. Many of the GD modifications considered in the research are based on them. For example, one of most popular and powerful algorithm Adam uses an exponential moving average of gradients and squared gradients for updating parameters like that:

$$m_i = \beta_1 m_{i-1} + (1 - \beta_1) g_i, v_i = \beta_2 v_{i-1} + (1 - \beta_2) {g_i}^2,$$

$$w = w - \alpha \frac{m_i}{\sqrt{v_i + \varepsilon}},$$

where α, β_1, β_2 – hyperparameters, ε –very small number (10^{-8}) to avoid division by zero.

In the research, various methods of varying complexity and logic were considered with subsequent comparisons of the obtained results. Depending on the specific task, these methods may show better or worse results relative to each other. For example, methods that use the idea of an exponential moving average perform better on sparse data. But all of them clearly outperform classical GD in terms of speed and the final value of the loss function.



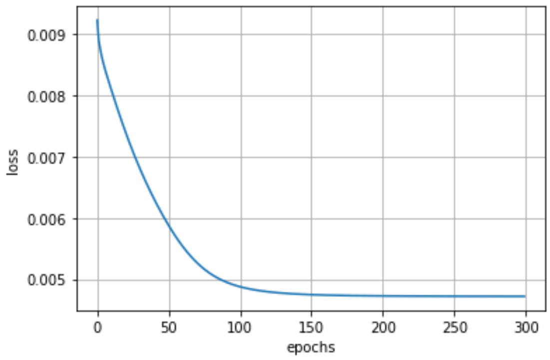


Figure 1 - Gradient Descent, Figure 2 - combination of Adam and Exponential Learning Rate Decay

LIST OF SOURCES USED

1. I. Goodfellow, Y. Bengio, A. Courville. “Deep Learning”. MIT Press, 2016, P. 271-309. Resource access mode: https://www.deeplearningbook.org/.

2. S. Nikolenko, A. Kadurin, E. Arkhangelskaya. “Deep Learning. Dive into the world of neural networks”. 2018, P. 69-81, 164-175

UDC 621.565.83

USE OF PELTIER ELEMENTS AS A HEAT PUMP FOR CONDENSATION DRYING OF FRUIT RAW MATERIALS

YAKUBASH I.V. (jakubash@ukr.net)
Odesa National University of Technology, Odesa

Over the last decade, the average annual increase in the global production of thermoelectric cooling modules is 15-25%. The reason for this high and sustained growth rate is the absolute ecological purity of this method of energy conversion, as opposed to traditional methods [1].

The method of drying most materials by convective heat exchange is currently the most common. In order to increase the intensity of the drying process, the drying agent is forced into motion, which accelerates the heat-mass transfer processes between the wet material and the environment. As a result of this interaction, the material is heated and the moisture in it evaporates. During the drying process, the heated air is saturated with vaporized moisture from the surface of the material and is removed from the dried object. This process will be completed when the moisture values of the material and drying agent have reached equilibrium.

In the face of ever-increasing energy costs, the need to conserve energy, and especially in the drying processes, is very relevant. Increasing the efficiency of different types of drying plants is inextricably linked to the intensification of thermal and mass exchange characteristics through