Original article
DOI: https://doi.org/10.18721/JPM.16415

SIMULATION OF SEMI-INCLUSIVE DEEP INELASTIC LEPTON SCATTERING ON A PROTON AT ENERGIES OF 20–100 GEV ON THE BASIS OF A GENERATIVE-ADVERSARIAL NEURAL NETWORK

A. A. Lobanov✉, *Ya. A. Berdnikov*

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ lobanov2.aa@edu.spbstu.ru

Abstract. This paper continues a series of articles devoted to developing the capabilities of a deep inelastic lepton-proton scattering event generator based on the generative adversarial network (GAN). The investigation has focused on semi-inclusive reactions of deep inelastic scattering and, particularly, on hadron registration. The results confirmed that GAN could accurately generate distributions of physical properties of leptons and hadrons. It worked for different types of leptons and hadrons in the range of initial energies from 20 to 100 GeV in the center-of-mass system. The GAN demonstrated to preserve the inherent correlation between the characteristics of leptons and protons.

Keywords: semi-inclusive deep inelastic scattering, machine learning, neural network, generative-adversarial network

Citation: Lobanov A. A., Berdnikov Ya. A., Simulation of semi-inclusive deep inelastic lepton scattering on a proton at energies of 20–100 GeV on the basis of a generative-adversarial neural network, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 16 (4) (2023) 189–197. DOI: https://doi.org/10.18721/JPM.16415

Научная статья
УДК 539.12
DOI: https://doi.org/10.18721/JPM.16415

МОДЕЛИРОВАНИЕ ПОЛУИНКЛЮЗИВНОГО ГЛУБОКО НЕУПРУГОГО РАССЕЯНИЯ ЛЕПТОНА НА ПРОТОНЕ ПРИ ЭНЕРГИЯХ 20 – 100 ГЭВ НА ОСНОВЕ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНОЙ НЕЙРОННОЙ СЕТИ

А. А. Лобанов✉, *Я. А. Бердников*

[1] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉ lobanov2.aa@edu.spbstu.ru

Аннотация. Данная работа продолжает цикл статей, посвященных развитию возможностей генератора событий глубоко неупругого лептон-протонного рассеяния на основе генеративно-состязательной сети (ГСС). Здесь рассмотрены полуинклюзивные реакции глубоко неупругого рассеяния с регистрацией адрона. Показано, что ГСС позволяет с высокой точностью генерировать распределения физических характеристик конечных лептона и адрона в диапазоне начальных энергий 100 – 20 ГэВ в системе центра масс.

Ключевые слова: полуинклюзивное глубоко неупругое рассеяние, машинное обучение, нейронная сеть, генеративно-состязательная сеть

Ссылка для цитирования: Лобанов А. А., Бердников Я. А. Моделирование полуинклюзивного, глубоко неупругого рассеяния лептона на протоне при энергиях 20 –100 ГэВ на основе генеративно-состязательной нейронной сети // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 16. № 4. С. 189–197. DOI: https://doi.org/10.18721/JPM.16415

Introduction

Modern experimental research in high energy physics deals with increasingly large datasets [1], collected from large-scale experiments or simulation results. Processing these data requires high computational costs and much time.

Machine learning methods offer an approach to solving the above-mentioned problems [2], allowing to construct computer simulation software (called event generators) with the following new capabilities:

using the experimental results on interactions of particles and nuclei at discrete points to predict the characteristics of secondary particles at any energies in the given range based on interpolation (and possibly extrapolation) quickly and without high computational costs;

the above-mentioned software can be developed even without experimental results, using the simulation results for interactions of particles and nuclei obtained by the Monte Carlo method [3].

A generative-adversarial network (GAN) was described in [4] to create a generator for inclusive deep inelastic lepton–proton scattering.

This paper continues the research in this direction, extending the capabilities of the given event generator [4] to semi-inclusive deep inelastic scattering with hadron production.

The goal of the study was to build a generator that can be trained on experimental data (or those obtained by computer simulation), allowing to collect intermediate data based on interpolation and extrapolation, since the experiment cannot be carried out at arbitrary initial energies.

There are several reasons for the interest in semi-inclusive processes.

First, the production of an additional hadron allows to learn more about the structure of the proton. Thus, the type of hadron produced by lepton–proton interaction depends on the flavor of the quark in the proton that the virtual photon emitted by the charged lepton interacted with [5].

Secondly, the characteristics of an additional hadron can carry information about the processes of parton hadronization [5].

Thirdly, various spin and azimuthal asymmetries can be measured during semi-inclusive processes, allowing to gain an understanding on the spin structure of the proton [6].

Methodology

The characteristics of the final state of the charged lepton (e^+, e^-, μ^+, μ^-) and hadron (π^0, π^+, π^-, K^+, K^-) are their four-momenta $p_l = (E_l,\mathbf{p}_l)$ and $p_h= (E_h,\mathbf{p}_h)$ respectively, where E_l is the total energy of the scattered lepton; p_l, $\mathbf{p}_l$ are the four- and three-dimensional momentum vectors of the lepton, the latter determined in terms of its components p_{xl}, p_{yl}, p_{zl}; E_h is the total energy of the hadron, p_h, $\mathbf{p}_h$ are four- and three–dimensional momentum vectors of the hadron, and also the components of the latter, p_{xh}, p_{yh}, p_{zh}.

For GAN to predict the four-momenta of various hadrons (π^0, π^+, π^-, K^+, K^-), their types (as well as the types of lepton) are fed to the input of the GAN as additional parameters along with the initial energy E_0, defined as $E_0 \approx \sqrt{s_{lN}}/2$, where $\sqrt{s_{lN}}$ is the initial energy in the lepton–proton center of mass frame [4].

Since it is currently impossible to experimentally obtain the characteristics of final-state leptons and hadrons (due to the lack of experiments), the finite states of leptons and hadrons were obtained using the PYTHIA8 software package [7].

For each type of lepton (e^+, e^-, μ^+, μ^-) and hadron (π^0, π^+, π^-, K^+, K^-), 100,000 events were generated at initial energies $\sqrt{s_{lN}}$ = 20, 40, 60, 80 and 100 GeV. The four-momentum values of the final-state lepton and hadron were obtained from each event (real data).

Following the approach in [4], we solved the problems associated with irregularities in the distributions of the quantities E_l, E_h and p_{zl} by generating, instead of the actual quantities E_l, E_h, p_{zl}, the quantities obtained by their transformation (transformed quantities):

$$T(p_{zl}) = \log[(E_0 - p_{zl})/(1\ \text{GeV}/c)],$$

$$T(E_l) = \log[(E_0 - E_l)/(1\ \text{GeV}/c)],$$

$$T(E_h) = \log[(E_h)/(1\ \text{GeV}/c)].$$

As established in [4], the distribution over the transformed quantities becomes smoother, preventing predictions of unphysical values.

Also similarly to [4], the event generator in this study is based on GAN with a least square loss function [8].

The generator consists of 5 layers of 512 neurons each with a Leaky ReLU activation function and a dropout of 0.2 [9]. A 128-dimensional noise vector (a vector of values obtained from a Gaussian distribution with the mean equal to 0 and the variance equal to 1), energy E_0, lepton type and hadron type are fed to the generator input. The generator outputs 8 characteristics:

$$p_{xl}, p_{yl}, T(p_{zl}), T(E_l), p_{xh}, p_{yh}, p_{zh} \text{ and } T(E_h),$$

corresponding to lepton and hadron.

Based on these characteristics, the model calculates additional values used to increase the accuracy of GAN predictions [4]:

$p_{Tl} = \sqrt{p_{xl}^2 + p_{yl}^2}$, $p_{Th} = \sqrt{p_{xh}^2 + p_{yh}^2}$ are the lepton and hadron transverse momenta, respectively;

$\varphi_l = \arctan(p_{zl}/p_{Tl})$, $\varphi_h = \arctan(p_{zh}/p_{Th})$ are the lepton and hadron azimuthal angles, respectively;

$\theta_l = \arctan(p_{yl}/p_{xl})$, $\theta_h = \arctan(p_{yh}/p_{xh})$ are the lepton and hadron polar angles, respectively.

All additional quantities are then fed to the discriminator input during training.

The discriminator also consists of 5 layers of 512 neurons each with a Leaky ReLU activation function and a dropout of 0.2 [9]. A dropout layer with a rate of 10% [11] is applied to each of the layers to prevent overfitting of the discriminator, randomly dropping 10% of the layer weights.

Spectral normalization is additionally applied to all layers for more stable training [12]. The output layer consists of a single neuron with a linear activation function. The higher the value obtained, the more confident the discriminator is in identifying the given values as realistic.

The model was trained for 400 epochs. RMSProp was used for gradient descent optimization, with $\rho = 0.9$ [13], $1\cdot10^{-4}$ training steps for the generator and $5\cdot10^{-5}$ for the discriminator. Using different training steps contributes to better training convergence, as shown in [14].

The Kullback–Leibler (KL) divergence was used as a measure of the divergence between the real data and those generated by GAN [15]. This measure was used to compare the histograms of the obtained distributions. In this case, the Kullback–Leibler divergence D_{KL} is defined as follows [15]:

$$D_{KL}(P \| Q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i},$$

where P, Q are the distributions of the real and generated data, respectively; p_i, q_i are the probabilities of the ith bins of histograms for real and generated data; n is the number of bins.

Simulation results

Since there is a wide range of scattering scenarios (different types of leptons and hadrons as well as different initial energies E_0), only some of the individual cases are given below to illustrate GAN's predictive capabilities.

Fig. 1 shows the distributions of p_T, θ, φ for the positron e^+ and the negative kaon K^-, obtained using GAN and PYTHIA8. Multiplicity is understood (in Fig. 1 and below) as the number of events in the bin normalized by the total number of events. Evidently, the model generates quantities whose distributions are almost identical, as indicated by the values of the Kullback–Leibler divergence shown in the graphs as well as the logarithmic ratios of GAN to PYTHIA8 predictions given for each graph.

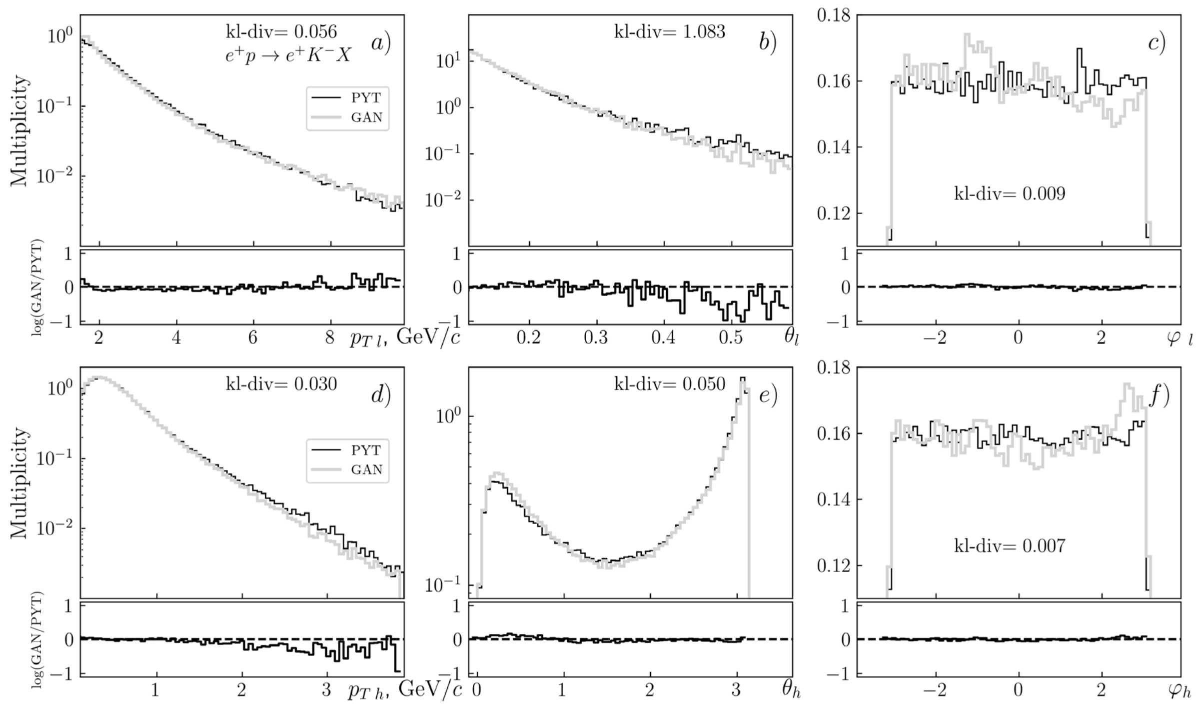

Fig. 1. Distributions of quantities p_T, θ, φ for positrons e^+ (a, b, c) and negative kaons K^- (d, e, f) at initial energy E_0 = 50 GeV.
The data were obtained using GAN (gray curves) and PYTHIA8 (black).
The corresponding values of KL divergence (kl-div) and graphs of the logarithmic ratio of GAN to PYTHIA8 (GAN/PYT) predictions are given for each distribution.

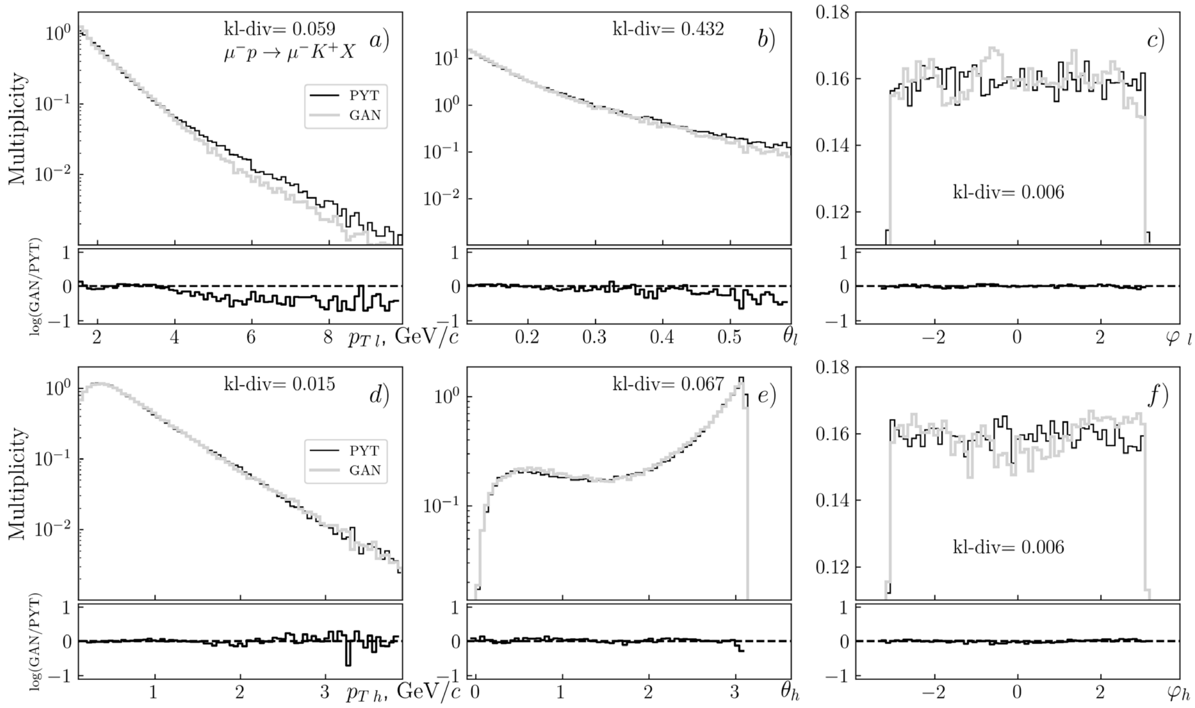

Fig. 2. Graphs similar to those shown in Fig. 1, but for muons μ^- (*a*, *b*, *c*) and positive kaons K^+ (*d*, *e*, *f*) at the initial energy E_0 = 20 GeV

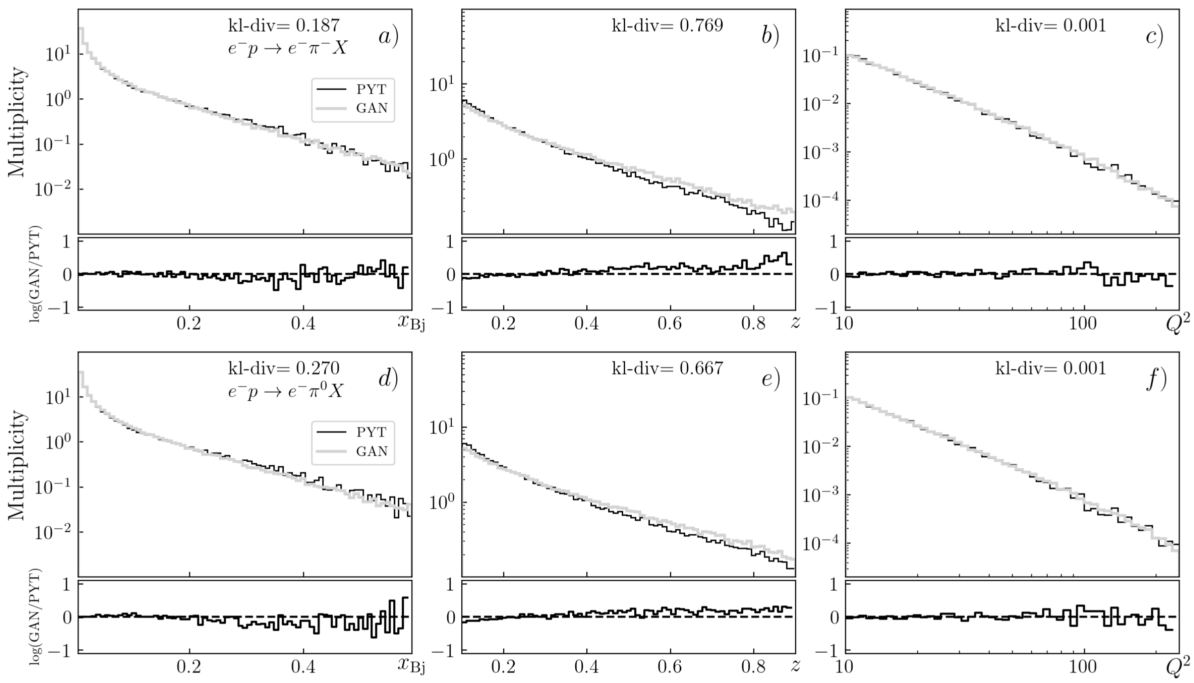

Fig. 3. Distributions of quantities x_{Bj}, z, Q^2 for the reactions $e^- p \to e^- \pi^- X$ (*a*, *b*, *c*) and $e^- p \to e^- \pi^0 X$ (*d*, *e*, *f*), respectively, at initial energy E_0 = 40 GeV.
The corresponding values of KL divergence (kl-div) and graphs of the logarithmic ratio of GAN to PYTHIA8 (GAN/PYT) predictions are given for each distribution.

Fig. 2 shows the distributions of the quantities p_T, θ, φ for the muon μ^- and the positive kaon K^+, obtained using GCC and PYTHIA8. These data demonstrate that the model can operate just as accurately with different leptons and hadrons at different initial energies.

Fig. 3 shows the distributions of squared momentum transfer $Q^2 = -q^2$ (q is the momentum of the virtual photon), as well as the Bjorken variable $x_{\rm Bj} = Q^2/2Pq$ (P is the momentum of the incident proton) and the fraction of the energy of the virtual photon transferred to the hadron, $z = P{\cdot}P_h/P{\cdot}q$ (P_h is the momentum of the proton) for nuclear reactions $e^-p \to e^-\pi^-X$ and $e^-p \to e^-\pi^0X$, where X denotes all other reaction products.

It follows from the presented results that the distributions generated by the model only differ slightly, as indicated by the values of the KL divergence obtained for each distribution.

Fig. 4 shows the distributions of the quantities $x_{\rm Bj}$, z, Q^2 for the reactions $e^+p \to e^+\pi^+X$ and $e^+p \to e^+K^-X$. Analyzing the obtained data, we can conclude that the accuracy of GAN predictions is preserved relative to real data from PYTHIA8 for different types of leptons and hadrons and different initial energies.

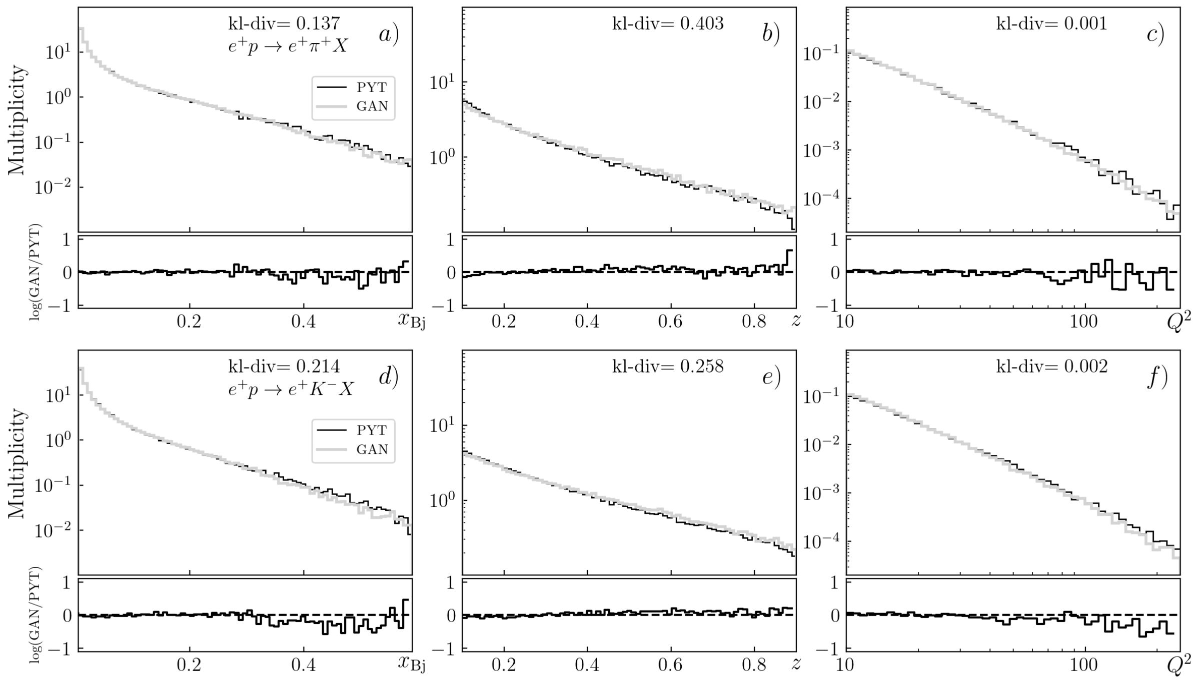

Fig. 4. Distributions of quantities $x_{\rm Bj}$, z, Q^2 for reactions $e^-p \to e^-\pi^-X$ (a, b, c) and $e^-p \to e^-\pi^0X$ (d, e, f) at initial energy E_0 = 30 GeV

The corresponding values of KL divergence (kl-div) and graphs of the logarithmic ratio of GAN to PYTHIA8 (GAN/PYT) predictions are given for each distribution.

Conclusion

We developed a generative-adversarial network model that can predict the characteristics of final-state leptons (e^+, e^-, μ^+, μ^-) and hadrons (π^0, π^+, π^-, K^+, K^-) in semi-exclusive deep inelastic lepton–proton scattering in the initial energy range of 20–100 GeV.

We established that the above-mentioned GAN model is capable of faithfully reproducing four-momentum components of final-state leptons and hadrons.

It was confirmed that the model constructed can calculate the distributions for particles with high accuracy based on the transverse momentum p_T of the particles, the azimuthal (φ) and polar (θ) angles, the Bjerken variable $x_{\rm Bj}$, the energy fractions z of the virtual photon and the square momentum Q^2 transferred by the lepton to the hadron. The distributions of these quantities show high accuracy relative to the real data, proving that the model is capable of preserving the internal relationships between the values.

We also established that the GAN model accurately predicts the characteristics of leptons and hadrons both for the initial energies at which the model was trained and for the interpolated energies (intermediate values).

REFERENCES

1. **Cremonesi M., Bellini C., Bian B., et al.,** Using big data technologies for HEP analysis. arXiv: 1901.07143, 2019. https://doi.org/10.48550/arXiv.1901. 07143
2. **Jordan M. I., Mitchell T. M.,** Machine learning: Trends, perspectives, and prospects, Science. 349 (6245) (2015) 255–260.
3. **Weinzierl S.,** Introduction to Monte Carlo methods. arXiv: 0006269, 2000. https://doi.org/10.48550/arXiv.hep-ph/0006269.
4. **Lobanov A. A., Berdnikov Ya. A.,** A generator of deep inelastic lepton-proton scattering based on the Generative-Adversarial Network (GAN), St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 16 (4) (2023) ...–.... (in Russian).
5. **Aschenauer E. C., Borsa I., Sassot R., Van Hulse C.,** Semi-inclusive deep-inelastic scattering, parton distributions, and fragmentation functions at a future electron-ion collider, Phys. Rev. D. 99 (9) (2019) 094004.
6. **Barone V., Boglione M., Hernandez J. O. G., Melis S.,** Phenomenological analysis of azimuthal asymmetries in unpolarized semi-inclusive deep inelastic scattering, Phys. Rev. D. 91 (7) (2015) 074019.
7. **Sjöstrand T., Mrenna S., Skands P.,** A brief introduction to PYTHIA 8.1, Comput. Phys. Commun. 178 (11) (2008) 852–867.
8. **Mao X., Li Q., Xie H., et al.,** On the effectiveness of least squares Generative Adversarial Networks, IEEE Trans. Pattern Anal. Mach. Intell. 41 (12) (2019) 2947–2960.
9. **Sharma O.,** A new activation function for deep neural network, Proc. Int. Conf. Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), IEEE, Faridabad, India, Febr. 14–16 (2019) 84–86.
10. **Hawkins D. M.,** The problem of overfitting, J. Chem. Inf. Comput. Sci. 44 (1) (2003) 1–12.
11. **Srivastava N., Hinton G., Krizhevsky A., et al.,** Dropout: A simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (2014) 1929–1958.
12. **Miyato T., Kataoka T., Koyama M., Yoshida Y.,** Spectral normalization for Generative Adversarial Networks, arXiv: 1802.05957/v1, 2018. https://doi.org/10.48550/arXiv.1802.05957.
13. **Xu D., Zhang Sh., Zhang H., Mandic D. P.,** Convergence of the RMSProp deep learning method with penalty for nonconvex optimization, Neural Netw. 139 (July) (2021) 17–23.
14. **Heusel M., Ramsauer H., Unterthiner T., et al.,** GANs trained by a two time-scale update rule converge to a local Nash equilibrium, arXiv: 1706.08500v6, 2017. https://doi.org/10.48550/arXiv.1706. 08500.
15. **Shlens J.,** Notes on Kullback–Leibler divergence and likelihood, arXiv: 1404.2000, 2014. https://doi.org/10.48550/arXiv.1404.2000

СПИСОК ЛИТЕРАТУРЫ

1. **Cremonesi M., Bellini C., Bian B., et al.** Using big data technologies for HEP analysis. arXiv: 1901.07143, 2019. https://doi.org/10.48550/ arXiv. 1901. 07143

2. **Jordan M. I., Mitchell T. M.** Machine learning: Trends, perspectives, and prospects // Science. 2015. Vol. 349. No. 6245. Pp. 255–260.

3. **Weinzierl S.** Introduction to Monte Carlo methods. arXiv: 0006269, 2000. https://doi.org/10.48550/arXiv.hep-ph/0006269.

4. **Лобанов А. А., Бердников Я. А.** Генератор глубоко неупругого рассеяния лептонов на протоне на основе генеративно-состязательной нейронной сети // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 16. № 4. С. ...–....

5. **Aschenauer E. C., Borsa I., Sassot R., Van Hulse C.** Semi-inclusive deep-inelastic scattering, parton distributions, and fragmentation functions at a future electron-ion collider // Physical Review D. 2019. Vol. 99. No. 9. P. 094004.

6. **Barone V., Boglione M., Hernandez J. O. G., Melis S.** Phenomenological analysis of azimuthal asymmetries in unpolarized semi-inclusive deep inelastic scattering // Physical Review D. 2015. Vol. 91. No. 7. P. 074019.

7. **Sjöstrand T., Mrenna S., Skands P.** A brief introduction to PYTHIA 8.1 // Computer Physics Communications. 2008. Vol. 178. No. 11. Pp. 852–867.

8. **Mao X., Li Q., Xie H., Lau R. Y. K., Wang Zh., Smolley S. P.** On the effectiveness of least squares generative adversarial networks // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019. Vol. 41. No. 12. Pp. 2947–2960.

9. **Sharma O.** A new activation function for deep neural network // Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). IEEE, Faridabad, India, February 14–16, 2019. Pp. 84–86.

10. **Hawkins D. M.** The problem of overfitting // Journal of Chemical Information and Computer Sciences. 2003. Vol. 44. No. 1. Pp. 1–12.

11. **Srivastava N., Hinton G., Krizhevsky A., Slutskever I., Salakhutdinov R.** Dropout: A simple way to prevent neural networks from overfitting // The Journal of Machine Learning Research. 2014. Vol. 15. Pp. 1929–1958.

12. **Miyato T., Kataoka T., Koyama M., Yoshida Y.** Spectral normalization for Generative Adversarial Networks. arXiv: 1802.05957/v1, 2018. https://doi.org/10.48550/arXiv.1802.05957.

13. **Xu D., Zhang Sh., Zhang H., Mandic D. P.** Convergence of the RMSProp deep learning method with penalty for nonconvex optimization // Neural Networks. 2021. Vol. 139. July. Pp. 17–23.

14. **Heusel M., Ramsauer H., Unterthiner T., Nessler B., Hochreiter S.** GANs trained by a two time-scale update rule converge to a local Nash equilibrium. arXiv: 1706.08500v6, 2017. https://doi.org/10.48550/arXiv.1706. 08500.

15. **Shlens J.** Notes on kullback-leibler divergence and likelihood. arXiv: 1404.2000, 2014. https://doi.org/10.48550/arXiv.1404.2000.

THE AUTHORS

LOBANOV Andrey A.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
lobanov2.aa@edu.spbstu.ru
ORCID: 0000-0002-8910-4775

BERDNIKOV Yaroslav A.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
berdnikov@spbstu.ru
ORCID: 0000-0003-0309-5917

СВЕДЕНИЯ ОБ АВТОРАХ

ЛОБАНОВ Андрей Александрович – *студент* Физико-механического института *Санкт-Петербургского политехнического университета Петра Великого.*

195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
lobanov2.aa@edu.spbstu.ru
ORCID: 0000-0002-8910-4775

БЕРДНИКОВ Ярослав Александрович – *доктор физико-математических наук, профессор* Высшей школы фундаментальных физических исследований *Санкт-Петербургского политехнического университета Петра Великого.*

195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
berdnikov@spbstu.ru
ORCID: 0000-0003-0309-5917

Received 28.09.2023. Approved after reviewing 12.10.2023. Accepted 12.10.2023.
Статья поступила в редакцию 28.09.2023. Одобрена после рецензирования 12.10.2023. Принята 12.10.2023.