

Book of Abstracts

Draft

The technology of long-term forecasting of water inflow into reservoirs using a multi-parameter neural network

N.V. Abasov, Melentiev Energy Systems Institute SB RAS, Russia

E.N. Osipchuk, Melentiev Energy Systems Institute SB RAS, Russia

V.M. Berdnikov, Melentiev Energy Systems Institute SB RAS, Russia

Short presentation (15 min)

The technology of long-term forecasting of water inflow into reservoirs from a month to several years is considered. Using created by the authors a multi-parameter neural network (MNN), a forecasting technology has been developed with stages: 1) search for potential predictors in the form of areas correlated with water inflow by the vorticity indices of the selected atmospheric layer; 2) neural network model is synthesized with different parameters (prediction accuracy in the form of division into a finite number of intervals, a list of predictors, the number of hidden layers, the number of neurons of different types in each layer). In most cases, for a different set of predictors learning occurs successfully on the training samples. Significant errors occur on verification samples. To reduce them an algorithm has been developed for automatic search and selection of predictors, varying the structural parameters of MNN with the selection of models that form the minimum errors on the verification samples. Models with a maximum deviation error of no more than one interval are selected. The final forecast decision is determined based on the creation of probability distributions over a set of selected models. In further studies, it is proposed to use indices that take into account the power of cosmic rays, solar activity, the influence of various planets, and other factors that can exclude or reduce errors in the verification samples.

Legacy of Tunka-Rex software and data.

Pavel Bezyazeev, API ISU, Russia

Tunka-Rex is a digital antenna array for measuring the radio emission from air-showers, induced by high-energy cosmic rays. The array started operation in 2012 with 18 antennas and had significantly developed over the years, finishing measurements in 2019 with 63 antennas and upgraded data acquisition. Analysis and processing of the collected data is a complex procedure which contains a number of steps (monitoring of state of the array, low-level filtration, quality cuts, reconstruction of air-shower parameters etc.). We give an overview of software developed for these tasks and our experience gained during the work with Tunka-Rex data. The legacy of software and data is discussed in the frame of the FAIR (Findability – Accessibility – Interoperability – Reuse) concepts.

Short presentation (15 min)

Equivariant Gaussian Processes as Limiting Convolutional Networks with Infinite Number of Channels

A.Demichev, SINP MSU, Russia

Short presentation (15 min)

The topic within which this work was carried out is related to the establishment of relationships between various methods of machine learning (ML). The ultimate goal of establishing such interrelations is to achieve a better theoretical understanding of these methods and their improvement. In particular, a correspondence has recently been established between the appropriate asymptotics of deep neural networks (DNNs), including convolutional ones (CNNs), and the ML method based on Gaussian processes. Since Gaussian processes are mathematically equivalent to free (Euclidean) quantum field theory (QFT), one of the intriguing consequences of these relationships is the potential for using a broad range of QFT methods for analyzing DNNs. There are evidences (including experimental) that non-asymptotic (that is, implementable in practice) DNNs correspond to QFT with interactions. An important feature of convolutional networks is their equivariance (consistency) with respect to the symmetry transformations of the input data. In this work, we establish a relationship between the many-channel limit of equivariant CNNs and the corresponding equivariant Gaussian process (GP), and hence the QFT with the appropriate symmetry. The approach used provides explicit equivariance at each stage of the derivation of the relationship.

Modeling images of proton events for the TAIGA project using a generative adversarial network: features of the network architecture and the learning process

Ju.Dubenskaya, SINP MSU, Russia

A.P.Kryukov, SINP MSU, Russia

Short presentation (15 min)

High-energy particles interacting with the Earth atmosphere give rise to extensive air showers emitting Cherenkov light. This light can be detected on the ground by imaging atmospheric Cherenkov telescopes (IACTs). One of the main problems solved during primary processing of experimental data is the separation of signal events (gamma-quanta) against the hadronic background, the bulk of which is made up of proton events. To ensure correct gamma event/proton event separation under real conditions, a large amount of experimental data, including model data, is required. Thus, although proton events are considered as background, their images are also necessary for accurate registration of gamma quanta. We applied a machine learning method - generative adversarial networks to generate images of proton events for the TAIGA project. This approach allowed us to significantly increase the speed of image generation. At the same time testing the results using third-party software showed that over 90% of the generated images were correct. In this article we provide an example of a GAN architecture suitable for generating images of proton events similar to those obtained from IACTs of the TAIGA project. The features of the training process are also discussed, including the number of learning epochs and selecting appropriate network parameters.

Use of conditional generative adversarial networks to improve representativity of data in optical spectroscopy

A.O.Efitorov, SINP MSU, Russia

S.A.Burikov, Department of Physics Lomonosov Moscow State University Moscow, Russia

T.A.Dolenko, Department of Physics Lomonosov Moscow State University Moscow, Russia

K.A.Laptinskiy, SINP MSU, Russia

S.A.Dolenko, SINP MSU, Russia

Short presentation (15 min)

The report considers the approach of improving the results of solving the inverse problem of spectroscopy of water-ethanol solutions by generating an additional array of patterns by a generative adversarial neural network (GAN). To solve this problem, 40710 Raman spectra (low-frequency region + region of the valence band of water) of ethanol solutions containing impurities (methanol, ethyl acetate, fusel oil) in various concentrations or without impurities were recorded in laboratory conditions. More than 8 thousand examples were extracted from the dataset as a test set, which was used only to assess the performance of the trained classifier network. The considered problem was that of detecting presence of each of the three possible impurities. In order to increase the number of patterns, a conditioned GAN (1D deep convolutional network) was trained, the "conditional" parameter was a vector with binary encoding of the presence of different components in the water-ethanol solution. Due to the significant differences in the structure of the spectra of the low-frequency spectral region (a number of narrow high-intensity peaks) and the region of the valence band of water (no narrow peaks, a smooth shape of the valence band), for each spectral region its own generator-discriminator pair was trained. After training cGAN, more than 40 thousand examples were generated. Then an additional neural network was trained to solve the classification problem. Various combinations of patterns in the classification training set were considered: real spectra only, generated spectra only, merged (generated + real) dataset. A comparative analysis of the results of these approaches on a test set of real spectra demonstrated that the joint use of generated and real data increased the accuracy of solving the classification problem. The study was supported by the Russian Foundation for Basic Research, project no. 19-01-00738. No presentation

A convolutional hierarchical neural network classifier

I.M. Gadzhiev, SINP MSU, Russia

S.A. Dolenko, SINP MSU, Russia

Short presentation (15 min)

The report presents an algorithm for constructing a convolutional hierarchical neural network classifier, which is a modification of the algorithm for constructing hierarchical neural network classifiers suggested before. The original algorithm was designed to exploit intrinsic class hierarchy to build a class tree with a neural network in each node classifying groups of initial classes (in a non-terminal node) or a subset of original classes (in a terminal node). The convolutional modification utilizes convolutional neural networks instead of regular fully connected networks in order to apply the model to image classification tasks. Use of class hierarchy for image classification should reduce the number of adjusted neural network parameters compared to deep convolutional neural networks, and therefore it should reduce training and inference time. In this context the algorithm may be

compared with some pruning techniques. The convolutional hierarchical neural network classifier inherits some hyperparameters of a conventional hierarchical neural network classifier, like the activation threshold and the threshold by the share of voting patterns. The goal of this study was to explore different strategies of choosing these hyperparameters. To test these strategies, we used the CIFAR-10 dataset. Also, for demonstration purposes we apply the convolutional hierarchical neural network classifier to the CIFAR-100 dataset. This study has been funded by the SINP MSU state budget topic 6.1 (01201255512).

The preliminary results on analysis of TAIGA-IACT images using Convolutional Neural Networks

Elizaveta Gres, ISU, Irkutsk, Russia

Alexander Kryukov, SINP MSU, Moscow, Russia

Short presentation (15 min)

The imaging Cherenkov telescopes TAIGA-IACT, located in The Tunka valley of the republic Buryatia, accumulate a lot of data in a short period of time which must be qualitatively and quickly analyzed. One of the methods of such analysis is the machine learning, which has proven its effectiveness in many technological and scientific fields in recent years. The aim of the work is to study the possibility of the machine learning application to solve the tasks set for TAIGA-IACT. In the work the method of Convolutional Neural Networks was applied to process and analyze Monte-Carlo events simulated with CORSIKA. Also various CNN architectures for the processing were considered. It has been demonstrated that this method gives good results of the determination the type of primary particles of Extensive Air Shower (EAS) and the recovery of gamma-rays energy. The results are significantly improved in the case of stereoscopic observations.

Neural network solution of inverse problems of geological prospecting with discrete output

Igor Isaev, SINP MSU, Moscow, Russia, Kotelnikov Institute of Radio Engineering and Electronics, RAS, Moscow, Russia

Ivan Osbornev, SINP MSU, Moscow, Russia

Eugeny Osbornev, S.Ordjonikidze Russian State Geological Prospecting University, Moscow, Russia

Eugeny Rodionov, S.Ordjonikidze Russian State Geological Prospecting University, Moscow, Russia

Mikhail Shimelevich, S. Ordjonikidze Russian State Geological Prospecting University, Moscow, Russia

Sergey Dolenko, SINP MSU, Moscow, Russia

Short presentation (15 min)

The inverse problems of exploration geophysics are to reconstruct the spatial distribution of the properties of the medium in the Earth's thickness from the geophysical fields measured on its surface. In particular, this paper deals with the problems of gravimetry, magnetometry, and magnetotelluric sounding, as well as their integration, i.e., the simultaneous use of several geophysical fields to restore the desired distribution. To implement the integration, a 4-layer 2D model was used, where the inverse problem was to determine the lower boundary of the layers, and each layer was characterized by variable values of the depth of the lower boundary along the section and fixed

values of density, magnetization, and resistivity, both for the layer and for the entire data set. To implement the neural network solution of the inverse problem, a data set was generated by solving the direct problem, where for each pattern, the distribution of layer depth values was set randomly in a given range and with a given step, i.e. it took discrete values from a certain set. In this paper, we consider an approach involving the use of neural networks to solve the problem of multiclass classification, where class labels correspond to discrete values of the determined layer depths. The results of the solution are compared with the results of the solution of the same inverse problem in the formulation of the regression problem, in terms of the error in determining the depth of the layers. This study has been performed at the expense of the grant of the Russian Science Foundation (project no. 19-11-00333).

Graph Neural Networks and application for Cosmic-Ray Analysis

Paras Koundal; IAP , KIT Karlsruhe, Germany

Long presentation (30 min)

Deep Learning has emerged as one of the most promising areas of computational research for pattern learning, inference drawing, and decision-making, with wide-ranging applications across various scientific disciplines. This has also made it possible for faster and more precise analysis in astroparticle physics, enabling new insights to be drawn from massive volumes of input data. Graph Neural Networks have naturally developed as a critical implementation method among the numerous deep-learning architectures over the last few years because of the unique ability they provide to represent complex input data from a wide range of problems in its most natural form. Described using nodes and edges, graphs allow us to represent relational data easily and learn hidden representations of input data for obtaining better model accuracies. At IceCube Neutrino Observatory, a multi-component detector, doing traditional likelihood-based analysis on a per-event basis to reconstruct cosmic-ray air shower parameters is time-consuming and computationally costly. Using advanced and flexible models based on Graph Neural Networks allows us to reduce the time and computing cost of performing such analysis while boosting sensitivity. An outline of Graph Neural Networks and a possible application using such methods at IceCube will be discussed.

Identifying partial differential equations of land surface schemes in INM climate models with neural networks

Mikhail Krinitskiy, Shirshov Institute of Oceanology, RAS, Russia

Viktor Stepanenko, Research Computing Center, MSU, Russia

Ruslan Chernyshev, Research Computing Center, MSU, Russia

Long presentation (30 min)

The core of a land surface scheme in climate models is a solver for a nonlinear PDE system describing thermal conductance and water diffusion in soil. This system includes thermal conductivity and water diffusivity coefficients that are functions of the solution of the system, i.e., water vapor content W and soil temperature T . For the climate models to accurately represent the Earth system's evolution, one needs to identify the equations meaning either approximating the coefficients or estimating their

values empirically. Measuring the coefficients is a complicated in-lab experiment without a chance to cover the full range of environmental conditions. The fact that there are many soil types obstructs comprehensive studies as well. There are also known approximate parametric forms of the coefficients that lack accuracy and, in turn, need their identification w.r.t. their own parameters. In this work, we propose a data-driven approach for approximating the parameters of the PDE system, describing the evolution of soil characteristics. We formulate the coefficients as parametric functions, namely artificial neural networks with expressive power high enough to represent a wide range of nonlinear functions. In a routine supervised data-driven problem, one needs to present ground truth for a target value. In the case of a soil PDE system, one cannot afford measurements of the full range of coefficients' ground truth values. In contrast with this approach, we propose training the neural networks with the loss function computed as a discrepancy between the PDE system solution and the measured characteristics W and T . We also propose a scheme inherited from the backpropagation method for calculating the gradients of the loss function w.r.t. network parameters. In contrast with recently developed physics-informed neural networks (PINN) methods, our approach is not meant to approximate a PDE solution directly. Instead, we state an inverse problem and propose its solution using artificial neural networks and the method for its optimization. As a very first step, we assessed the capabilities of our approach in four scenarios: a nonlinear thermal diffusion equation, a nonlinear water vapor W diffusion equation, Richards equation, and the system of thermal conductance equation and Richards equation. We generated realistic initial conditions and simulated synthetic evolutions of W and T that we used as measurements in the networks' training procedure in all these scenarios. We exploited batch normalization, learning rate schedule, and scheduling of additive noise rate to improve the convergence. We also added a few regularization terms to the loss function that penalize negative output values, non-zero output values in the origin, and negative gradients of the network w.r.t. its input. The results of our study show that our approach provides an opportunity for reconstructing the PDE coefficients of different forms accurately without actual knowledge of their ground truth values.

TAIGA: status, results and perspectives

L.Kuzmichev (SINP MSU) for the TAIGA collaboration

Invited presentation (45 min)

TAIGA (Tunka Advanced Instrument for cosmic ray physics and Gamma Astronomy) Astrophysical complex, located in the Tunka Valley, about 50 km from Lake Baikal, is designed for the study of gamma rays and charged cosmic rays in the energy range 10^{13} eV - 10^{18} eV. The deployment of the first stage installation consists of 120 optical station of HiSCORE array and 3 IACTs will be finished in autumn of 2021. In this report we present the main experimental and MC results in the field of high energy gamma-rays and high-energy cosmic-rays. The future plan for the installation upgrading will be also discussed.

The Russian language corpus and a neural network to analyse Internet tweet reports about Covid-19

Alexander Sboev, National Research Centre "Kurchatov Institute", Russia

Ivan Moloshnikov, National Research Centre "Kurchatov Institute", Russia

Alexander Naumov, National Research Centre "Kurchatov Institute", Russia

Anastasia Levochkina, National Research Centre “Kurchatov Institute”, Russia

Short presentation (15 min)

The problem of forecasting the evolution of the Covid-19 pandemic is extremely relevant because of the need for planning hospital beds demand and containment policies. Given the limitation of the available temporal datasets of the Covid-19 evolution, the complexity of machine learning algorithms to be created for pandemic time series data must not be high, and their efficiency mainly depends on the correct selection of highly-meaningful features. We view as one such feature the number of tweets where Internet users report having Covid-19. The task to extract such tweets from the internet is complicated by the lack of a Russian tweet dataset for training the machine learning models for extracting this category of tweets. In this work we present a corpus of about 10000 tweets labelled with the following classes: the Ill class when the authors declare that they are ill; the Recovered class when the authors declare that they have been ill and recovered, and the Others class of all other cases. Using this corpus, a Data-Driven model based on XLM-RoBERTa language model has been trained. It demonstrates the F1-macro accuracy of 0.85 for binary task (class 1 – Ill / Recovered, class 2 – Others), and 0.60 for the five-class task (Ill with high/low confidence, Recovered with high/low confidence, Others). These results outperform the RuDR-BERT model by 5 to 7%. The XLM-RoBERTa model thus obtained has been applied to the binary classification task for the unlabeled data of 486 000 tweets. Based on the model's predictions, a curve has been plotted of the COVID cases in Moscow from January 1, 2020 to March 1, 2021. This plot has been compared to the official statistic on the confirmed cases, and correlation analysis of these two curves, shifted from one another by 1 to 5 days, has been performed. This analysis shows the highest correlation to be when the true curve is shifted 4 days to the future relative to the predicted one. Therefore, the data from the corpus collected are 4 days ahead of the official statistic. Thus, the numbers of tweets collected have proven to be a helpful input feature to use within pandemic forecasting models.

Performance of convolutional neural networks processing simulated IACT images in the TAIGA experiment

Stanislav Polyakov, SINP MSU, Russia

Alexander Kryukov, SINP MSU, Russia

Evgeny Postnikov SINP MSU, Russia

Short presentation (15 min)

Extensive air showers created by high-energy particles interacting with the Earth atmosphere can be detected using imaging atmospheric Cherenkov telescopes (IACTs). The IACT images can be analyzed to distinguish between the events caused by gamma rays and by hadrons and to infer the parameters of the event such as the energy of the primary particle. We use convolutional neural networks (CNNs) to analyze Monte Carlo-simulated images of the telescopes of the TAIGA experiment. The analysis includes selection of the images corresponding to the showers caused by gamma rays and estimates of the energy of the gamma rays. We compare performance of the CNNs using images from a single telescope and the CNNs using images from two telescopes as inputs.

Keywords: deep learning; convolutional neural networks; gamma astronomy; extensive air shower; IACT; stereoscopic mode; TAIGA

EVALUATION OF MACHINE LEARNING METHODS FOR RELATION EXTRACTION BETWEEN DRUG ADVERSE EFFECTS AND MEDICATIONS IN RUSSIAN TEXTS OF INTERNET USER REVIEWS

A.G. Sboev, NRC «Kurchatov Institute», NRNU MEPhI, Russia

A.A. Selivanov, NRC «Kurchatov Institute», Russia

R.B. Rybka, NRC «Kurchatov Institute», Russia

I.A. Moloshnikov NRC «Kurchatov Institute», Russia

Short presentation (15 min)

The considered problem is automatic recognition of the relations between mentions of adverse drug reactions and medications in russian online drug reviews. This task solution is useful for pharmacovigilance and reprofiling of medicines. This problem hasn't been studied for the Russian language, due to the lack of corpora with relation labeling in Russian. Current research is based on a developed dataset with labeling of relations between entities from the Russian Drug Review Corpus of Russian Internet reviews on medications. Computational experiments were carried out on developed corpora using classical machine learning methods, as well as more advanced BERT topology model — RuDR-BERT. The classical machine learning methods were: support vector machine, logistic regression, Naive Bayes classifier and gradient boosting. In frame of these methods, concatenation of entity vectors obtained using TF-IDF of characters n-gram was used as a vector data representation for words, also the following hyperparameters of these method were selected based on a set of experiments: size of n-grams and limitation on the frequency of occurrence of n-grams (too rare or too frequent n-grams were excluded from the feature vector). For RuDR-BERT the input data is represented as usual for such type of models (language models based on Transformer topology). The following input types were considered during the experiments: text of target entities pair; text of target entities pair with words between them; text of target entities pair and the whole input text, the latter input type is the one that allows to achieve the highest accuracy. It is shown, that RuDR-BERT model achieve a result of 88%, according to the macro-averaged f1 metric, which is the state-of-the-art result in recognition of the relations between mentions of adverse drug reactions and medications in 0 russian online drug reviews. The Naive Bayes classifier with multivariate normal distribution achieves the best result among classical machine learning methods: 75%, according to which exceeds the result of random label generation by 21%.

Using modern machine learning methods on KASCADE data for science and education

Victoria Tokareva, IAP KIT, Germany

Short presentation (15 min)

Modern astroparticle physics makes wide use of machine learning methods in such problems as noise separation, image recognition, event classification, and analysis of spectrum mass composition. When using these methods, in addition to obtaining new scientific knowledge, it is important also to take advantage of their educational potential. In this work we present a demo version of the machine-learning based application we have created, which helps students and a broader audience to get

more familiar with the cosmic ray physics, and shows how machine learning methods can be used to analyze data. The work discusses the prospects for expanding the app's functionality and methodological approaches to the development of interactive outreach materials in this area.

Gamma/hadron separation for a ground based IACT (imaging atmospheric Cherenkov telescope) in experiment TAIGA using machine learning methods Random Forest

Vasyutina M.R., Moscow State University. Physical Department, Russia

Sveshnikova L.G., SINP MSU, Russia.

Short presentation (15 min)

In this report we present adaptation of the machine learning algorithm Random Forest (RF) to the gamma/hadron separation (g/h-separation) in the TAIGA experiment (Tunka Advanced Instrument for cosmic ray physics and Gamma-ray Astronomy). The first stage of TAIGA experiment will include HiSCORE array with 120 wide-angle Cherenkov detectors on the area of 1 sq.km and 5 Imaging Atmospheric Cherenkov Telescopes (IACT) on the same area. At this first stage of the analysis, only images obtained by one IACT were included in consideration. The training process occurs on samples of parameterized images obtained from Monte Carlo (MC) data for gammas and hadrons with a 'Scaled Hillas Parameters' standard technique. It was shown that the program effectively separates gamma-like showers, RF method does produce stable results and is robust with respect to input parameters and provides a simple control and setup of the procedure for extracting showers from gamma rays.

Using convolutional neural network for analysis of HiSCORE events

Vlaskina Anna, Physics Department, MSU, Russia

A. Kryukov, SINP MSU, Russia

Short presentation (15 min)

Project TAIGA is a hybrid observatory for gamma-ray astronomy at high energies in range from 10 TeV to several EeV. Project consists of instruments such as TAIGA-IACT, TAIGA-HISCORE and others. TAIGA-HISCORE, in particular, is an array of wide-angle timing Cherenkov light stations. TAIGA-HISCORE data enable air shower characteristics reconstructing, such as air shower energy, arrival direction and axis coordinates. In this report, we propose to consider the use of convolution neural networks in task of air shower characteristics determination. We use Convolutional Neural Networks (CNN) to analyze HISCORE events, treating them like images. For this, the times and amplitudes of events recorded at HiSCORE stations are used. The work discusses a simple convolutional neural network, its training. In addition, we present some preliminary results on the determination of the parameters of air showers such as the direction and position of the shower axis and the energy of the primary particle and compare them with the results obtained by the traditional method.

Application of deep learning technique to an analysis of hard scattering processes at colliders

A.Zaborenko, Faculty of Physics, M.V.Lomonosov Moscow State University, Russia

Lev Dudko, SINP MSU,Russia

Petr Volkov, SINP MSU,Russia

G. Vorotnikov, SINP MSU,Russia

Short presentation (15 min)

Deep neural networks have rightfully won the place of one of the most accurate analysis tools in high energy physics. In this talk we will cover several methods of improving the performance of a neural network in a classification task in an instance of top quark analysis. The approaches and recommendations will cover hyperparameter tuning, boosting on errors and AutoML algorithms applied to collider physics.

From:

<https://theory.sinp.msu.ru/> - **THEORY**

Permanent link:

<https://theory.sinp.msu.ru/doku.php/dlcp21/abstracts>

Last update: **22/06/2021 19:54**

