Ministry of Science and Higher Education of the Russian Federation
NATIONAL RESEARCH
TOMSK STATE UNIVERSITY (NI TSU)

Institute of Applied Mathematics and Computer Science

APPROVE:
Director
_____ A. V. Zamyatin
«15» _____ 2023

Work program of the discipline

Neural networks - I

in the major of training

01.04.02 Applied mathematics and informatics

Orientation (profile) of training:
Big Data and Data Science

Form of study
full-time

Qualification
master

Year of admission
2023

Code of discipline in the curriculum: B1.O.03

AGREED:
Head of OP
_____ A.V. Zamyatin

Chairman of the CMD
_____ S.P. Sushchenko

Tomsk - 2023

1. Purpose and planned results of mastering the discipline

The purpose of mastering the discipline is the formation of the following competencies:

- UK-1 - the ability to carry out a critical analysis of problem situations based on a systematic approach, to develop an action strategy;
- GPC-3 - the ability to develop mathematical models and analyze them when solving problems in the field of professional activity;
- PC-6 - the ability to manage the receipt, storage, transmission, processing of big data.

The results of mastering the discipline are the following indicators of the achievement of competencies:

IUK-1.1 Identifies a problem situation, on the basis of a systematic approach, carries out its multifactorial analysis and diagnostics.

IUK-1.2 Carries out the search, selection and systematization of information to determine alternative options for strategic solutions in a problem situation.

IUK-1.3 Suggests and justifies the strategy of action, taking into account the limitations, risks and possible consequences.

IOPC-3.3 Develops and analyzes new mathematical models for solving applied problems of professional activity in the field of applied mathematics and informatics.

IPK-6.1 Monitors and evaluates the performance of big data processing.

IPK-6.2 Uses methods and tools for receiving, storing, transmitting, processing big data.

IPK-6.3 Develops proposals to improve the performance of big data processing.

2. Tasks of mastering the discipline

– Master the apparatus for creating various topologies of artificial neural networks , as well as the technology for developing programs that use the main stages of the work of artificial neural networks (creating structures for representing neural networks, creating and evaluating a network model, training them).

– Learn to apply the conceptual apparatus of neural networks to solve practical problems of professional activity.

3. The place of discipline in the structure of the educational program

Discipline belongs to the mandatory part of the educational program.

4. Semester of mastering and form of intermediate certification in the discipline

First semester, exam.

5. Entrance requirements for mastering the discipline

For the successful mastering of the discipline, training outcomes are required in the following disciplines: no.

6. Implementation language

English.

7. Scope of discipline

The total labor intensity of the discipline is 6 credits, 216 hours, of which:

- lectures: 20 hours
- laboratory: 44 hours

including practical training: 0 h.

The volume of independent work of the student is determined by the curriculum.

8. The content of the discipline, structured by topics

Topic 1. Introduction to neural networks

Historical trends in neural networks. Neural networks: different names and changeable fortune. Increasing the size of the data set. Increasing the size of models. Increasing accuracy and complexity and expanding the range of tasks.

Topic 2. Linear algebra for neural networks.

Scalars, vectors, matrices and tensors. Multiplication of matrices and vectors. Identity and inverse matrix. Linear dependency and linear envelope. Norms. Special types of matrices and vectors. Spectral decomposition of a matrix. Singular decomposition. Pseudo-inverse Moore–Penrose matrix. trace operator. Determinant. Example: Principal Component Method.

Topic 3. Probability theory and information theory for neural networks.

Why do we need probability? random variables. Probability distributions. Discrete random variables and probability functions. Continuous random variables and probability density functions. Marginal probability distribution. Conditional Probability. Chain rule. Independence and conditional independence. Mathematical expectation, variance and covariance. Frequently occurring probability distributions. Bernoulli distribution. categorical distribution. Normal distribution. Exponential distribution and Laplace distribution. Dirac distribution and empirical distribution. Mixtures of distributions. Useful properties of commonly used functions. Bayes' rule. Technical details of continuous quantities. Information theory. Structural probabilistic models.

Topic 4. Numerical methods for neural networks.

Overflow and underflow. Bad conditioning. Gradient optimization. Not just the gradient: Jacobi and Hesse matrices. Constrained optimization. Example: Linear Least Squares.

Topic 5. Fundamentals of machine learning for neural networks.

Learning algorithms. Problem T. Quality measure P. Experience E. Example: linear regression. Capacity, overfitting and underfitting. The No Free Lunch Theorem. Regularization. Hyperparameters and control sets. Cross checking. Estimates, bias and variance. Point assessment. Bias. Variance and standard error. Finding a trade-off between bias and variance to minimize the mean square error. Consistency.

Topic 6. Fundamentals of machine learning for neural networks (continued).

Maximum Likelihood Estimation. Conditional log-likelihood and standard error. Maximum likelihood properties. Bayesian statistics. Estimated post hoc maximum (MAP).

Topic 7. Fundamentals of machine learning for neural networks (continued).

Supervised learning algorithms. Probabilistic supervised learning. Support vector machine. Other simple supervised learning algorithms. Learning algorithms without a teacher. Principal component method. Clustering by k means method. Stochastic gradient descent. Building a machine learning algorithm. Problems requiring deep learning. The Curse of Dimension . Regularization to achieve local consistency and smoothness. Manifold learning.

Topic 8. Neural networks of direct distribution.

Example: XOR training. Training by gradient methods. cost functions. output blocks. hidden blocks. Blocks of linear rectification and their generalizations. Logistic sigmoid and hyperbolic tangent. Other hidden blocks. Architecture design. Universal Approximation Properties and Depth . Other architectural approaches.

Topic 9. Neural networks of direct propagation (continued).

Back propagation and other differentiation algorithms. Graphs of calculations. The rule of differentiation of a complex function. Recursively applying the rule of differentiation of a complex function to obtain a backpropagation algorithm. Computing Backpropagation in a Fully Connected MSP. Symbolic-symbolic derivatives. General backpropagation algorithm. Example: Applying Back Propagation to SME Learning. Complications. Differentiation beyond the deep learning community. Derivatives of higher order . Historical remarks.

Topic 10. Regularization for neural networks.

Fines according to the norm of parameters. Regularization of parameters in the L2 norm. L1 regularization. Norm penalty as constrained optimization. Regularization and underdetermined problems. Replenishment of the data set. Robustness relative to noise. Adding noise to output labels. Teaching with the partial involvement of a teacher. Multitasking learning. Early stop. Linking and sharing parameters. Convolutional neural networks. Sparse Views. Bagging and other ensemble methods. thinning. Competitive learning. Tangent Distance, Tangent Propagation Algorithm, and Manifold Tangent Classifier.

9. Ongoing evaluation

The ongoing evaluation of the discipline is carried out by monitoring attendance, performing laboratory work, tests on lecture material, doing homework and is recorded in the form of a control point at least once a semester.

Laboratory work No. 1 "Solving problems in the design of neural networks."

The aim of the work is to design a multilayer neural network using stochastic gradient descent with a given learning rate and activation function. It is necessary to iterate the training of the network with a given loss function and calculate the value of the synaptic coefficient after its adjustment; to design a neural network with stochastic neurons, the output of each stochastic neuron is a discrete random variable described by the distribution, it is necessary to find the mathematical expectation of the network output when a given value is applied to the input.

Laboratory work No. 2 "Designing a fully connected neural network for a multiclass classifier"

The purpose of the work is to write a program in Python and R that builds and trains fully connected feed-forward neural networks that solves multiclass classification problems (samples received from the teacher), it is required to select a non-redundant network architecture that works with an acceptable error level and visualize the process of model training. The results of the work should be included in the report.

10. The procedure for conducting and criteria for evaluating the intermediate certification

The basis of training is a course of lectures read by the teacher, as well as laboratory work performed by the student.

Intermediate certification and credit is carried out on the basis of an interview, subject to the successful completion of earlier laboratory work.

The results of the test are determined by the marks "excellent", "good", "satisfactory", "unsatisfactory".

The final assessment of the knowledge of the student in the discipline is carried out according to a 5-point system and includes:

- 60% of the result obtained in the test;
- 40% of the results of the current academic performance.

The formula for calculating the final grade:

$$F = 0{,}4\,\frac{P_1+P_2}{2} + 0{,}6T \quad (1)$$

where, P1, P2 are the digital equivalents of the first and second control points, respectively; T - the digital equivalent of the assessment in the test.

Assessment according to the traditional system	Criterion
Excellent	The theoretical content of the course was mastered completely, without gaps the necessary practical skills of working with the mastered material were formed, all the training tasks provided for by the training program were completed, the quality of their implementation was estimated by a number of points close to the maximum.
Good	The theoretical content of the course has been mastered completely, without gaps, some practical skills in working with the mastered material are not sufficiently formed, all the training tasks provided for by the training program have been completed, the quality of none of them has been assessed with a minimum number of points, some types of tasks have been completed with errors.
Satisfactorily	The theoretical content of the course has been partially mastered, but the gaps are not significant, the necessary practical skills for working with the mastered material are basically formed, most of the training tasks provided for by the training program have been completed, some of the completed tasks may contain errors
unsatisfactory	The theoretical content of the course has not been mastered, the necessary practical work skills have not been formed, the completed training tasks contain gross errors, additional independent work on the course material will not lead to a significant improvement in the quality of the training tasks.

11. Educational and methodological support

a) An electronic training course in the discipline at the electronic university "Moodle" https :// moodle . ido . tsu . en / course / view . php ? id =1396& section =3

b) Estimated materials of the current control and intermediate certification in the discipline.

c) Plan of seminars / practical classes in the discipline.

d) Guidelines for laboratory work.

e) Guidelines for the organization of independent work of students .

12. List of educational literature and Internet resources

a) basic literature:

1. Joel Grace. Data Science: Data science from scratch. 2nd edition. ISBN 978-5-9775-6731-2, St. Petersburg: BHV-Petersburg, 2021

2. Sebastian Raska, Vahid Mirjalili. Python and machine learning. ISBN 978-5-907203-57-0, Moscow: Dialectics, 2020

3. Ameet V. Joshi. Machine Learning and Artificial Intelligence. ISBN 978-3-030-26621-9. Springer Nature Switzerland AG, 2020

4. Denis Rothman. Artificial Intelligence by Example. second edition. ISBN 978-1-83921-153-9. Packt Publishing, 2020

5. Stuart Russel, Peter Norvig. artificial intelligence. A Modern Approach. 4th edition. ISBN : 978-0-13-461099-3. Hoboken : Pearson , 2021
6. Ian Goodfellow, Joshua Bengio, Aaron Courville. Deep learning. Second color edition, revised. ISBN 978-5-97060-618-6. M .: DMK Press , 2018
7. Roman Shirkin. artificial intelligence. The Complete Beginners' Guide to Artificial Intelligence. ISBN: 9798609154415 . Amazon KDP Printing and Publishing, 2020
8. François Cholet. Deep learning in Python. ISBN 978-5-4461-0770-4. St. Petersburg: Peter, 2018

b) additional literature:

1. Andrew Glassner. Deep learning without math. Volume 1. Basics. ISBN 978-5-97060-701-5. Moscow: DMK Press, 2020
2. Andrew Glassner. Deep learning without math. Volume 2. Practice ISBN 978-5-97060-767-1. Moscow: DMK Press, 2020

c) Internet resources:

www.MachineLearning.ru is a professional wiki resource dedicated to machine learning and data mining

MMRO - Mathematical Methods of Pattern Recognition

Konstantin Vorontsov. The course "machine learning" of the data analysis school of Yandex.

Igor Kuralenok. Machine learning course Lectorium.

Roman Shamin. Course "Machine learning and artificial intelligence in mathematics and applications". REC of the Mathematical Institute. V. A. Steklov RAS

13. List of information technologies

a) licensed and freely distributed software:

- For the acquisition of practical skills - freely distributed environments with open source Python (https :// www . python . org /) and RStudio (https :// www . rstudio . com /);

- For project -group and remote work - Russian software Mind (https://mind.com/) .

– publicly available cloud technologies (Google Docs, Yandex disk, etc.).

b) information reference systems:

– Electronic catalog of the TSU Scientific Library – http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system

– TSU electronic library (repository) – http://vital.lib.tsu.ru/vital/access/manager/Index

14. Logistics

When mastering the discipline, computer classes of the IPMKN TSU are used with access to the resources of the Scientific Library of TSU, including domestic and foreign periodicals and the Internet.

15. Authors information

Baklanova Olga Evgenievna – Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of the Department of Theoretical Foundations of Informatics, TSU.