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

Institute of Applied Mathematics and Computer Science

APPROVE:
Director

_____ A. V. Zamyatin

Work program of the discipline

**Data Science for Industrial Data Analytics**

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.P.V.DV.02.02

AGREED:
Head of EP
_____ A.V. Zamyatin

Chairman of the EMC
_____ 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:
- PC-5 - the ability to choose methods, draw up terms of reference and develop algorithms for solving problems of industrial data analysis.
The results of mastering the discipline are the following indicators of the achievement of competencies:
IPK-5.1 Uses modern information processing technologies, computer technology in solving problems of industrial data analysis.
IPK-5.2 Knows how to collect industrial data, knows the specifics of such data.
IPK-5.3 Draws up the terms of reference for the task of the professional area.

## 2. Tasks of mastering the discipline

− To gain knowledge for the design and development of industrial data analysis systems;

− To get acquainted with the methods of analysis and technologies for collecting and processing industrial data.

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

The discipline refers to the part of the educational program formed by the participants in educational relations, Elective disciplines 2.

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

Third semester, exam.

## 5. Entrance requirements for mastering the discipline

Successful mastering of the discipline requires competencies formed in the course of mastering educational programs of the previous level of education.

## 6. Implementation language

English.

## 7. Scope of discipline

The total labor intensity of the discipline is 3 credits, 108 hours, of which:
- lectures: 10 hours
- laboratory: 20 hours
  including practical training: 16 h.
The volume of independent work of the student is determined by the curriculum.

## 8. The content of the discipline, structured by topics

**Section 1.** Basics of Industry 4.0
Introduction to Industry 4.0
Digital twins
Industrial Internet of Things
Predictive analytics systems
Classification and description of technological data
Study of educational material, preparation for practical exercises
Current progress monitoring
**Section 2.** Industrial Data Analysis
Pre-processing of process signals
Extraction of informative features from technological signals

Detection of anomalies in process data
Algorithms for Classification and Clustering in the Problem of Anomaly Detection.

### 9. Ongoing evaluation

Current discipline control is carried out by monitoring attendance, conducting tests, tests on lecture material, business games on topics, doing homework and is recorded in the form of a control point at least once a semester..

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

Mastering the course takes place by listening to a course of theoretical lectures and performing practical work. Each student implements an individual or group project as a sequence of practical work. Project topics have the following template:

1. Implement an algorithm for analyzing technological data.

2. Propose and implement technologies for improving the performance of calculations performed by the algorithm.

### 11. Educational and methodological support

a) Electronic training course on the discipline at the electronic university "Moodle"

b) Assessment materials of the ongoing evaluation and intermediate certification in the discipline.

### 12. List of educational literature and Internet resources

a) basic literature
1. Zamyatin A.V. Introduction to data mining. State University Publishing House, 2016
2. Klaus Schwab. Fourth industrial revolution. Moscow: Eksmo, 2016

b) additional literature
3. Denisenko V.V. Computer control of technological process, expert, equipment. Moscow, Hot Line Telecom, 2009
4. Frank E. Grubbs Procedures for detecting outlying observations in samples. Technometrics, 11(1), doi: 10.1080/00401706.1969.10490657, 1969, 1–21
5. Varun Chandola, Arindam Banerjee, and Vipin Kumar Anomaly detection: A survey ACM Computing Surveys, 41(3), doi: 10.1145/1541880.1541882 2009, 1–72

c) Internet resources:
– open online courses
- "Expert" magazine - http://www.expert.ru
– Official website of the Federal State Statistics Service of the Russian Federation - www.gsk.ru
– Official website of the World Bank - www.worldbank.org
– All-Russian Network ConsultantPlus Reference Legal System. http://www.consultant.ru

**13. List of information technologies**

a) licensed and freely distributed software:

Python programming tools and environments.

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

– EBS Lan – http://e.lanbook.com/

– EBS Student Advisor – http://www.studentlibrary.ru/

– Urayt educational platform – https://urait.ru/

– EBS ZNANIUM.com – https://znanium.com/

- EBS IPRbooks - http://www.iprbookshop.ru/

**14. Logistics**

Halls for lectures.

Classrooms for seminars, individual and group work, ongoing evaluation and intermediate certification.

Classrooms for independent work, equipped with computer technology and access to the Internet, to the electronic information and educational environment and to information reference systems.

Halls for lectures and seminars, individual and group consultations, ongoing evaluation and intermediate certification in a mixed format ("Aktru").

**15. Authors information**

Murzagulov Damir Albertovich, Assistant of the Department of Theoretical Foundations of Informatics, TSU.