TOMSK POLYTECH

JOINT INSTITUTE FOR NUCLEAR RESEARCH

EFFICIENCY ASSESSMENT OF IRT-T RESEARCH REACTOR COOLING SYSTEM BY MACHINE LEARNING METHODS

Postgraduate student: Maxim K. Kublinskiy, group A3-43

Scientific supervisor: Artem G. Naymushin, PhD, Associate Professor of NFCD, Head of the Educational and Scientific Center "Research Nuclear Reactor"

30 October 2023

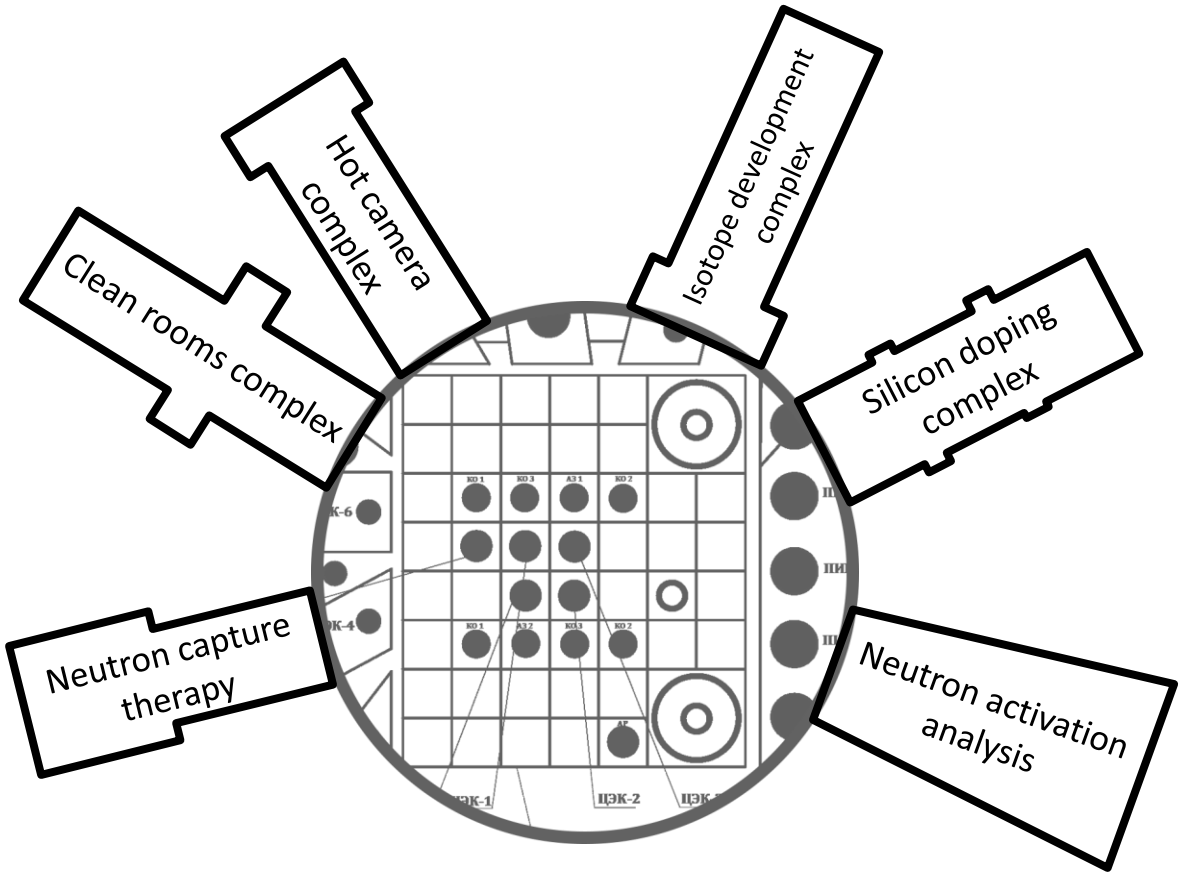
RESEARCH POSSIBILITIES

NUCLEAR MEDICINE

- Radiopharmaceuticals
- Neutron capture therapy
- $Tc^{99м}$ generation

RADIATION TECHNOLOGIES

- Modification of the optical properties of semiprecious minerals
- Element Analysis



ISOTOPE DESIGN

- Isotopes for medical and technological purposes ^{32}P, ^{90}Y, ^{99}Mo, ^{177}Lu, ^{153}Sm
- Neutron-transmutation doping of semiconductor materials

NUCLEAR EDUCATION

- World-class educational programs
- Virtual simulators and simulators

RELEVANCE

Maintaining proper operating conditions for research nuclear facilities is **an important factor** in ensuring the safety and implementation of high-tech research.

Correct evaluation of secondary cooling circuit operation is one of these operating conditions.

PURPOSE AND TASKS

JOINT INSTITUTE FOR NUCLEAR RESEARCH

TOMSK POLYTECH

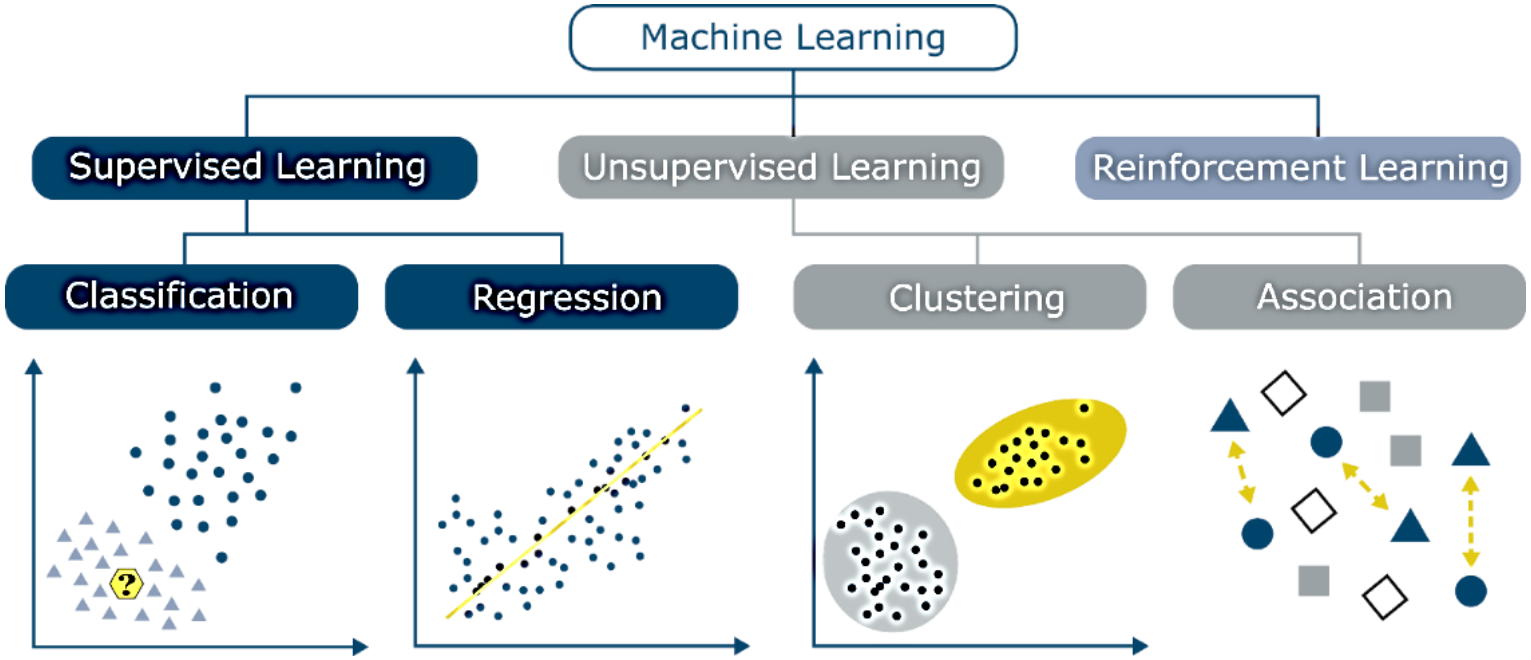
The purpose:

to develop the machine learning model of the IRT-T cooling system to evaluate and predict temperature difference on secondary circuit.

Tasks:

- to develop the unique preprocessing software to evaluate SCADA System data;
- to evaluate preprocessing transformers influence on training accuracy;
- to estimate the best machine learning model on Time Series and Default (80/20) Folds.

MACHINE LEARNING



Learning tracking algorithms study the relationship between predictive variables and the results of selected training data sets, and learning formulas are used to prepare models to classify new data.

SUPERVISED LEARNING

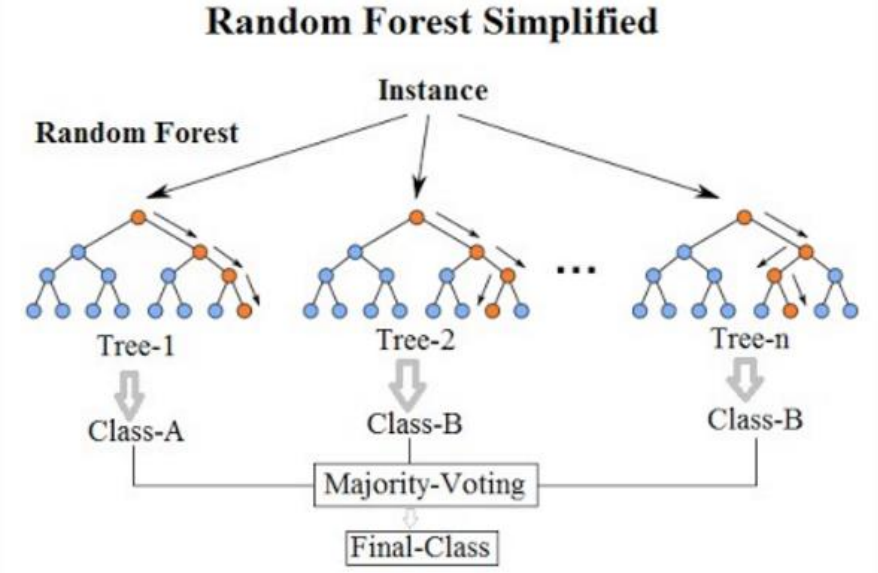
Supervised learning

is an approach to creating artificial intelligence (AI), where a computer algorithm is trained on input data that has been labeled for a particular output.

Regression

Gradient Boost Regression (GBR);
Gaussian Process Regression (GPR);
Multilayer Perception (MLP);
Stochastic Gradient Descent with Warm Restarts (SGDR).

Classification

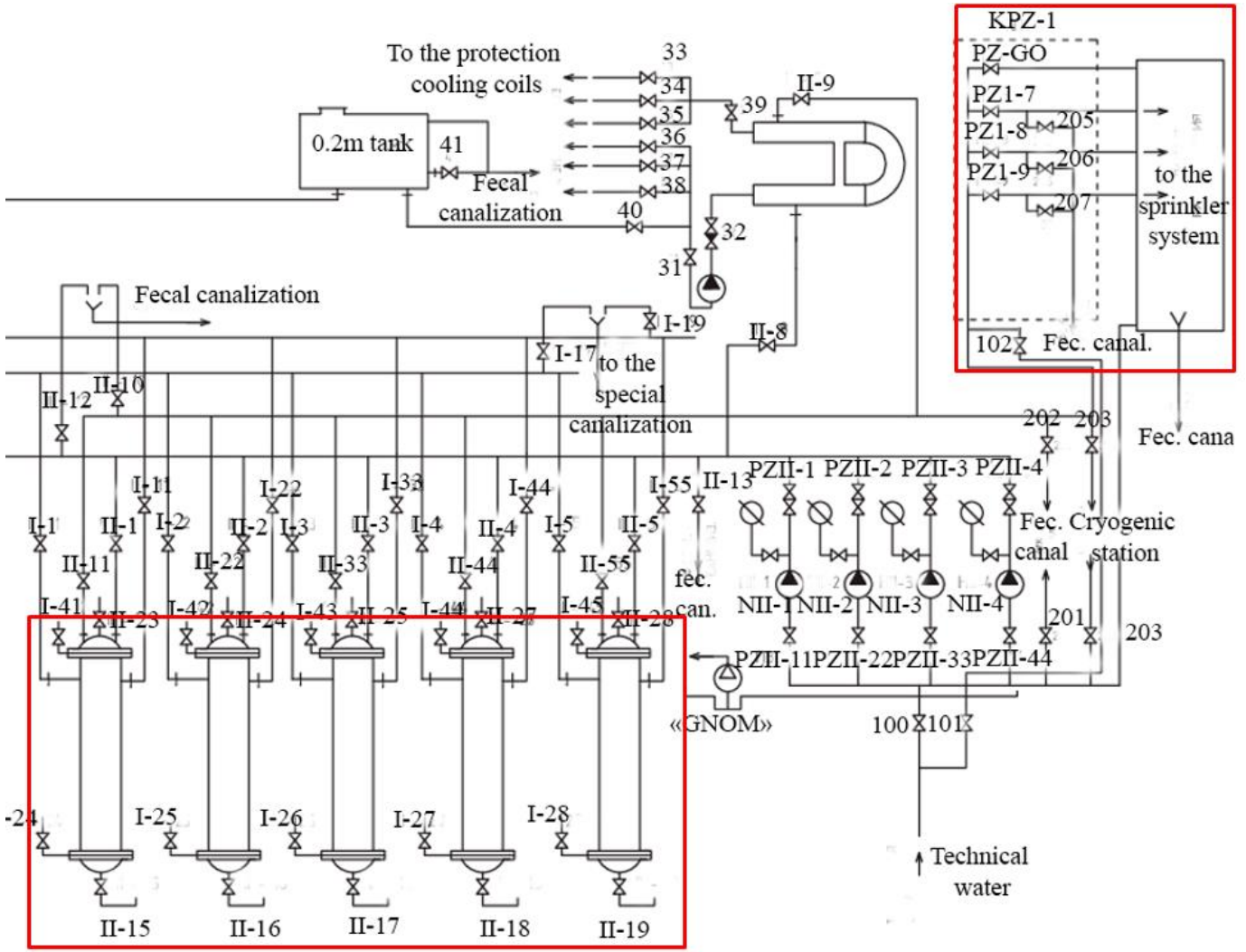


SECONDARY COOLING CIRCUIT OF IRT-T REACTOR

JOINT INSTITUTE FOR NUCLEAR RESEARCH

TOMSK POLYTECH

← to primary circuit



SIMPLE SCADA SOFTWARE COMPLEX

Simple SCADA software complex was developed in 2019 by service department of IRT-T for simplification of management and informational support for the personnel who is responsible for reactor operation.

SIMPLE SCADA

SOFTWARE BASIS FOR RESEARCH WORK

JOINT INSTITUTE FOR NUCLEAR RESEARCH

TOMSK POLYTECH

Pandas

SCLearn

NumPY

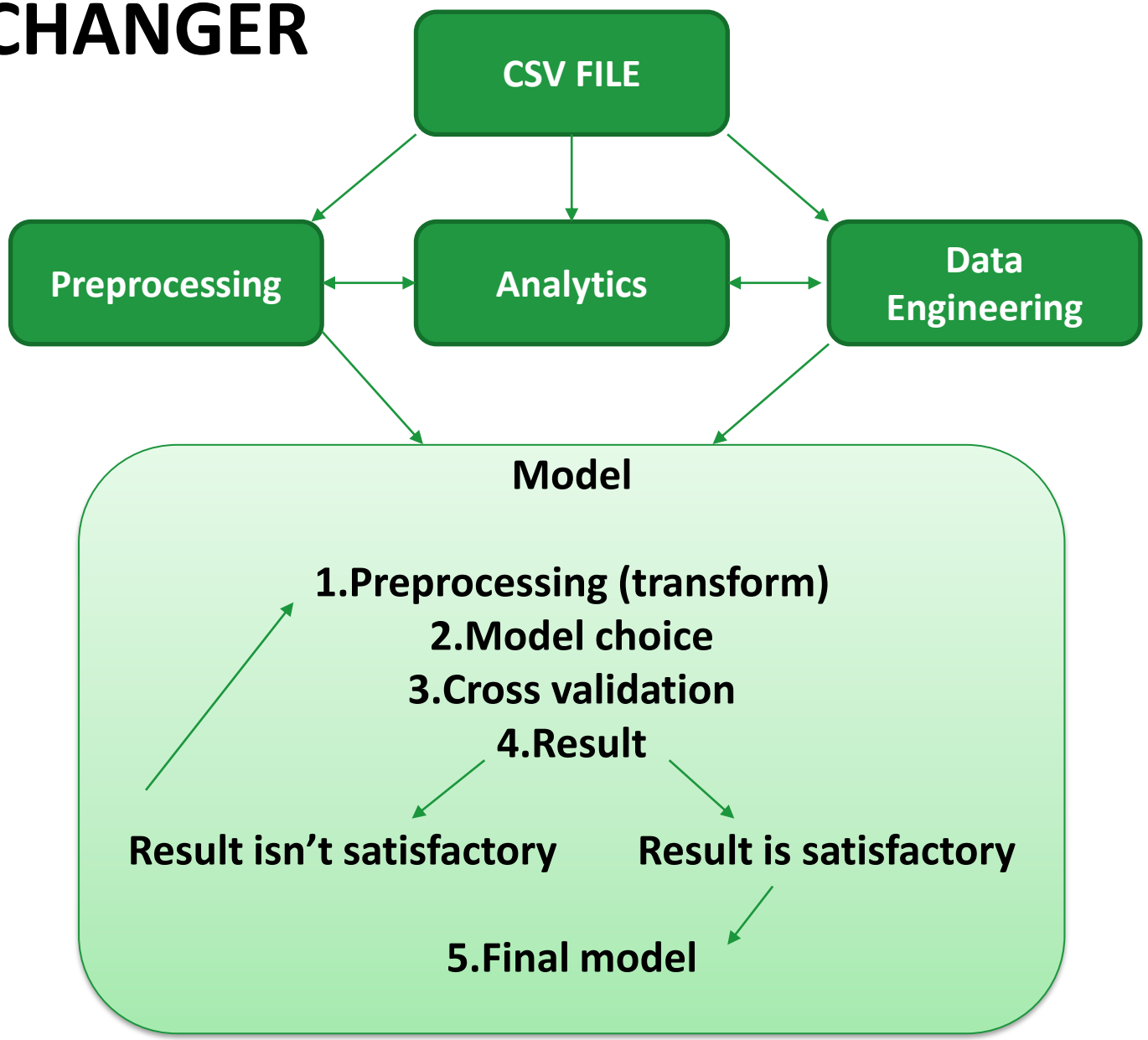
Jupyter

is a web-based interactive computing platform that combines live code, equations, narrative text, visualizations.

IAPWS

is a library for calculation of water and steam properties for Python.

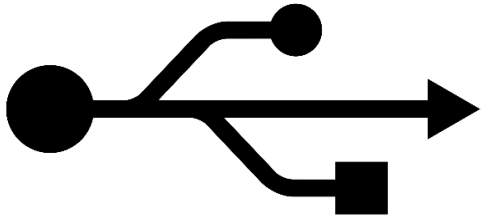
DECISION TREE FOR THE MODEL OF HEAT EXCHANGER



The following structure formats the decision tree. Preprocessing, analytics and data engineering go together in one file.

Model construction has its internal steps that have to be done before the final evaluation.

CSV-FILE PREPROCESSING



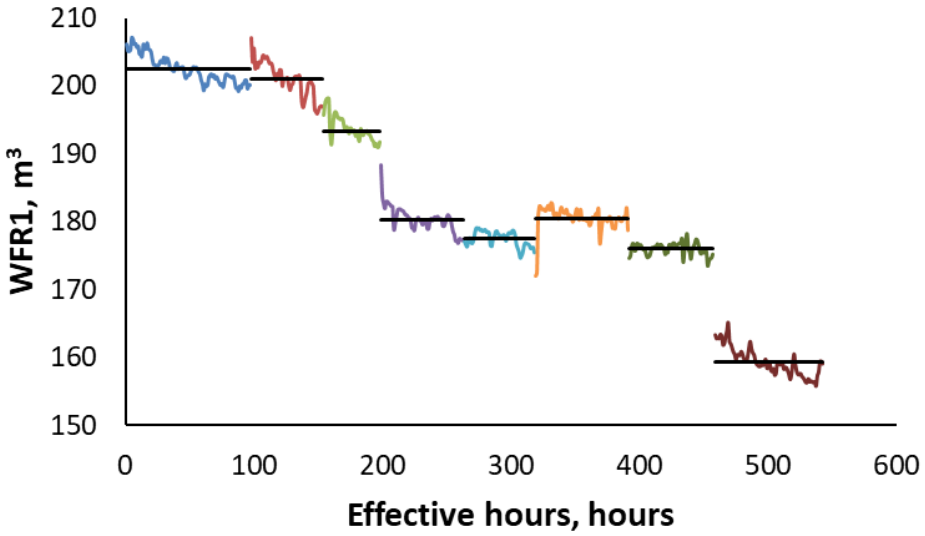
Downloading the data from SimpleSCADA

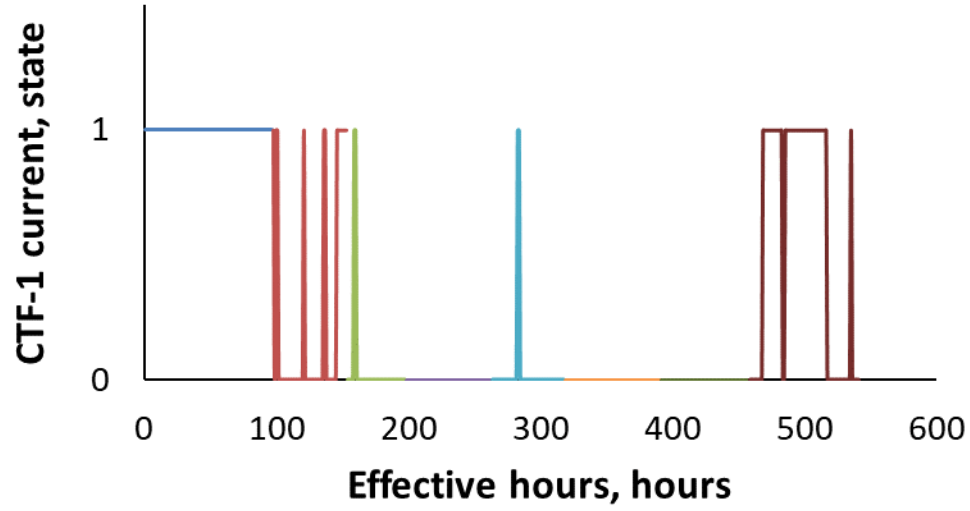
Initial preprocessing through CsvRefactorer function

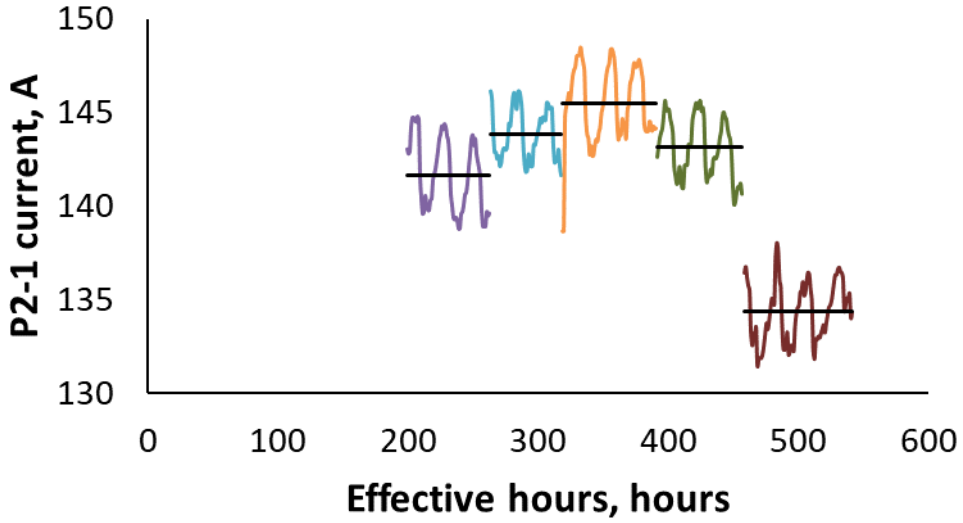
- Reading
- Export
- Time Bounds
- Grouping
- Length Check

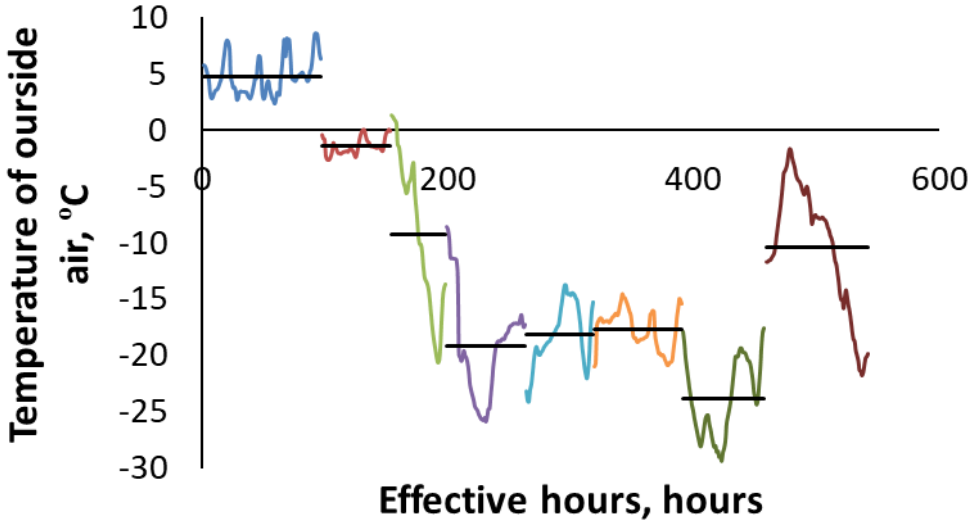
Subsequent analysis

DATA ANALYSIS









17.10 – 21.10

08.11 – 11.11

16.11 – 18.11

22.11 – 25.11

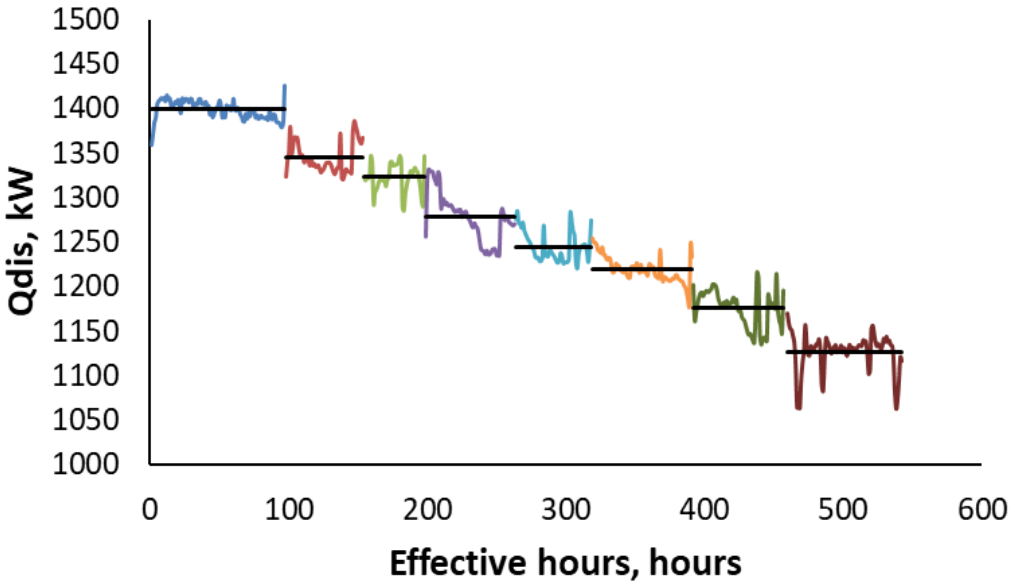
30.11 – 02.12

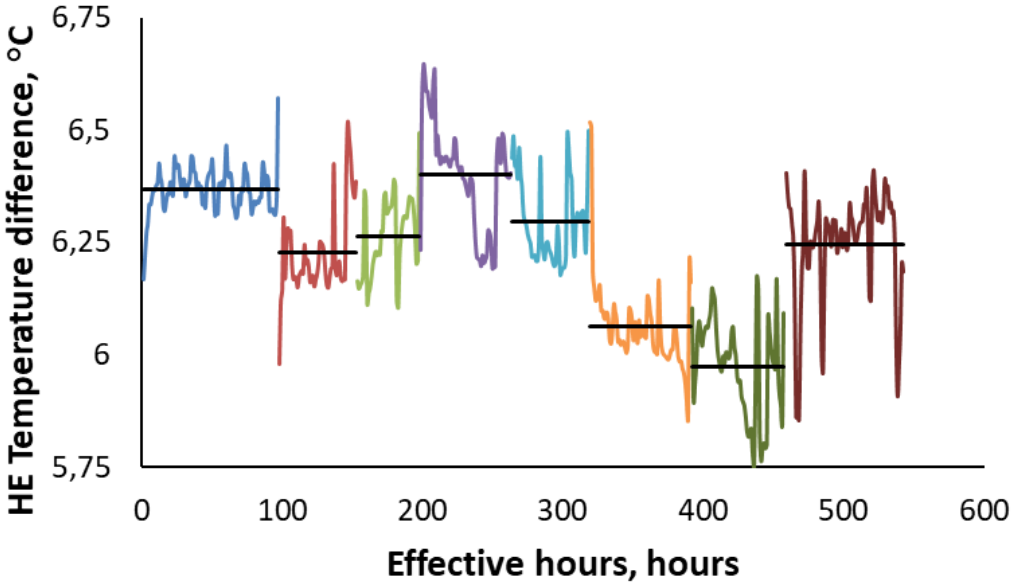
06.12 – 09.12

13.12 – 16.12

20.12 – 23.12

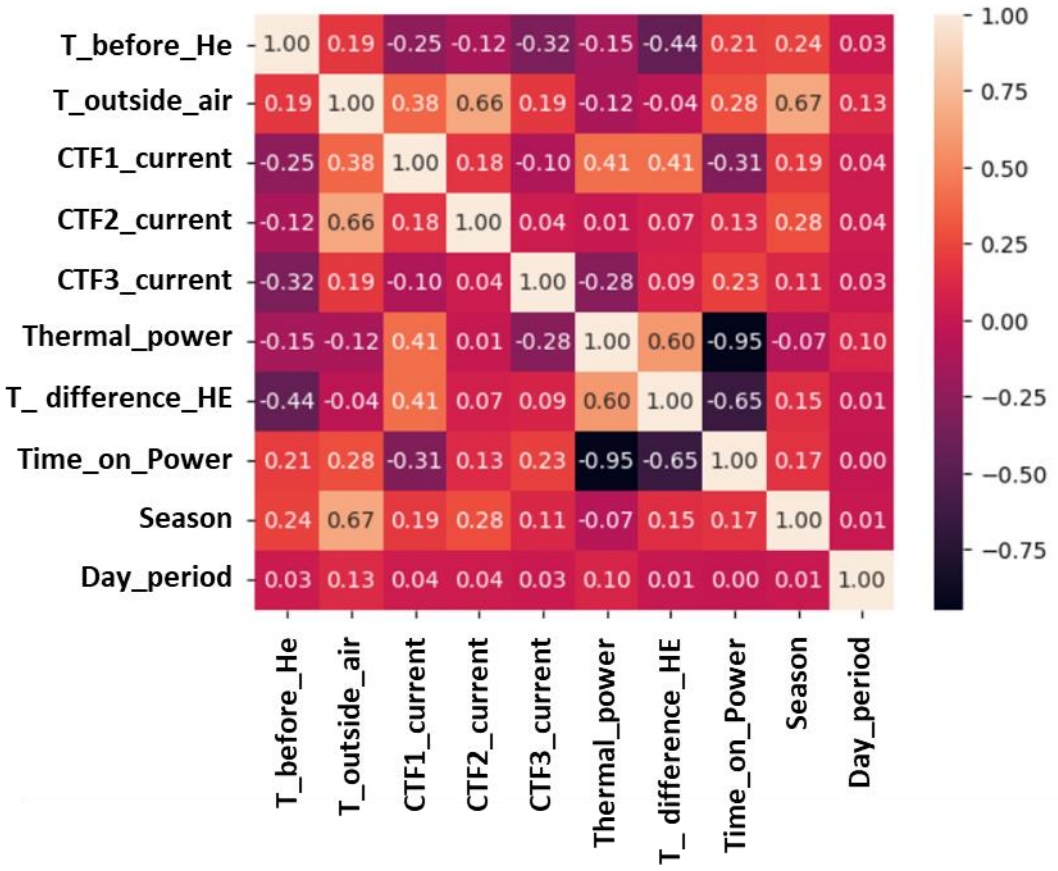
DATA ENGINEERING





Initially that was decided to determine **total heat transfer coefficient**, but due to complexity of calculations work was switched to studying on **thermal power and temperature difference** of the secondary circuit because these parameters are proportional to each other.

REGRESSION MODEL BASIS



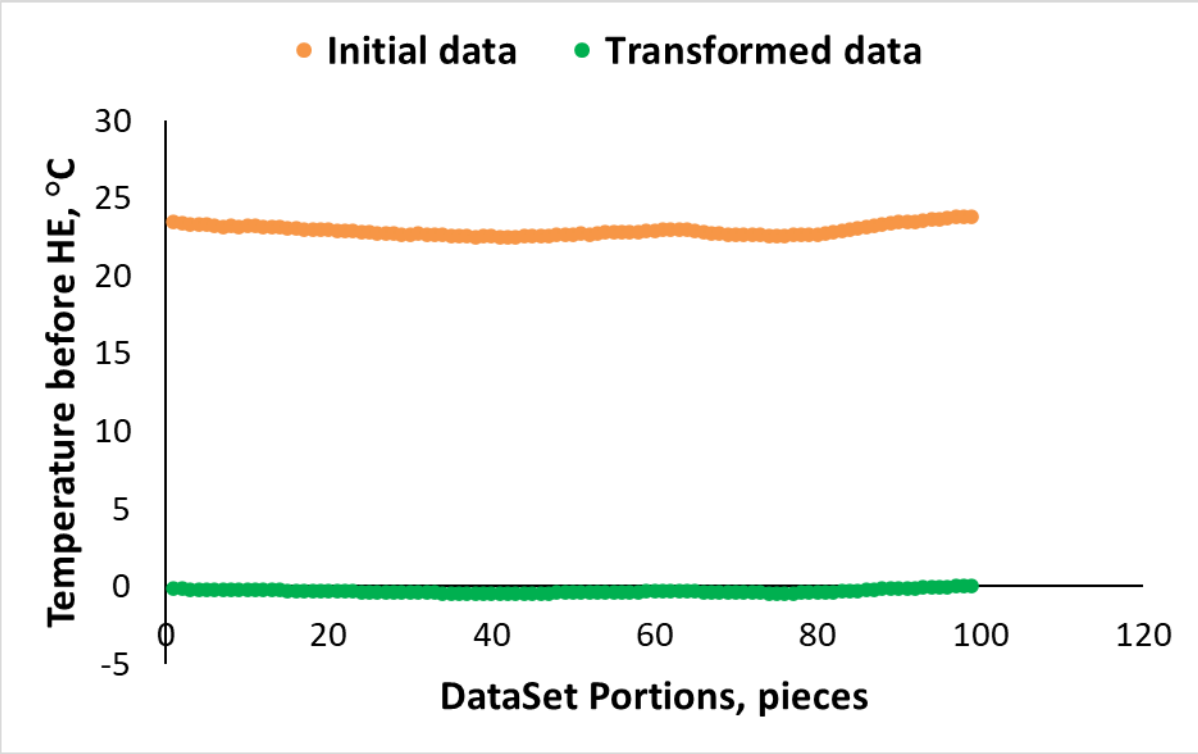
That was decided that **the most important parameters** for analysis are:

- Temperature before heat exchangers;
- Temperature of the outside air;
- Cooling tower fans currents;
- Temperature difference on heat exchangers.

For proper study of the model it is necessary to build **the heat map of parameters** to see the correlation between parameters.

For example, dT **increases** when CTF1 in operation, and **decreases** with increasing of temperature before heat exchangers.

TRANSFORMERS FOR DATASET

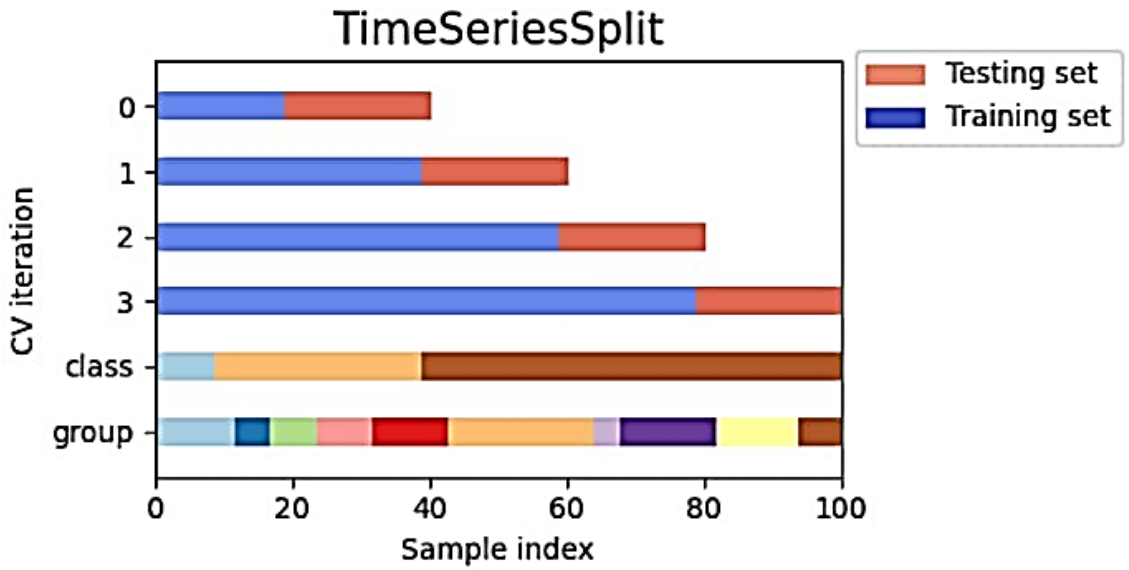


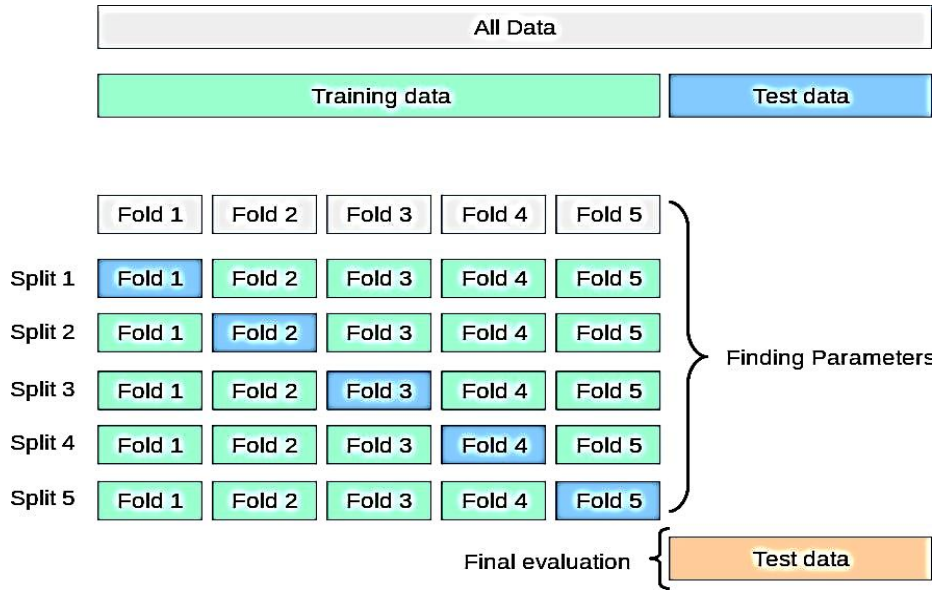
The idea was to develop **Column Transformer** function. The preprocessed data has to be changed in the view which is easier to machine to learn.

The thing is about taking the values of the dataset, finding the average value, **which is represented as 0 and matching deviations from 0 in positive or negative side.**

For example was taken little part of data set for temperature before HE. Average temperature was nearly **22.8 °C** so that was prescribed as **0** and deviations are formed. This method allows to improve accuracy of calculations for all parameters.

REGRESSION MODEL BASIS





Splitting time series on the parts is a simple idea. The supervisor needs to choose a part as a training set (usually bigger) and as a test set. Then the process of **cross validation** comes when different splits are chosen for training and test. This way the model gets studied.

The **more trainings** – the **better result** in the end.

Using **80/20 for training/testing** the data leads to **smoothing** the data and seems to be optimal for machine learning.

REGRESSION MODELS CHOICE

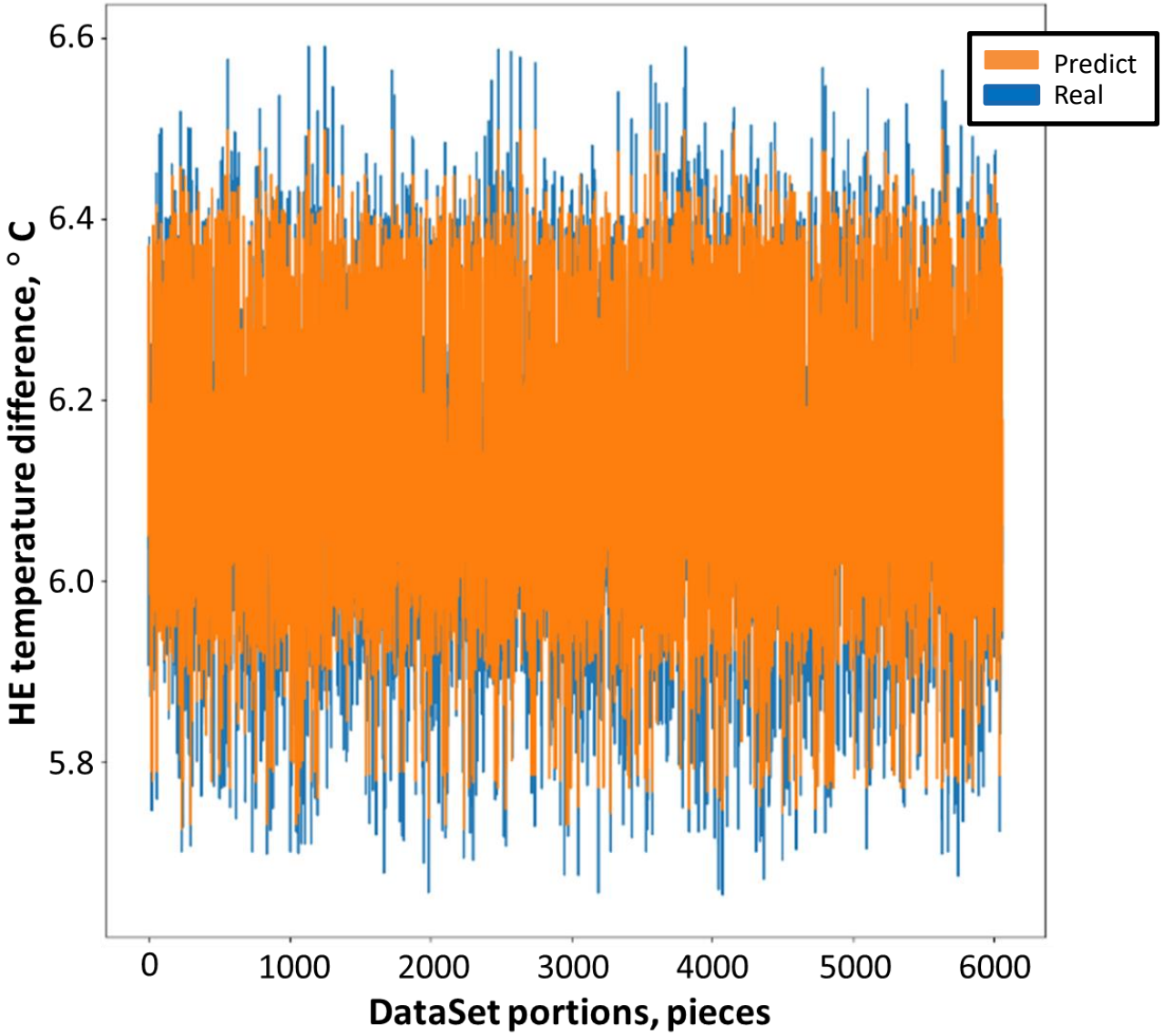
JOINT INSTITUTE FOR NUCLEAR RESEARCH

TOMSK POLYTECH

	With transformers		Without transformers	
Model	Mean Error, °C	Root Error, °C	Mean Error, °C	Root Error, °C
KNeigbours	0.14 +/- 0.02	0.16 +/- 0.02	0.18+/- 0.08	0.26 +/- 0.08
Multylayer perception	0.55 +/- 0.08	0.63 +/- 0.08	1.46 +/- 0.08	1.23 +/- 0.05
Gaussian processed regression	0.13 +/- 0.03	0.15 +/- 0.03	0.52 +/- 0.02	0.75 +/- 0.04
Gradient boost regression	**0.12 +/- 0.01**	**0.14 +/- 0.02**	**0.13 +/- 0.02**	**0.15 +/- 0.02**
Stochastic Gradient Descent with Warm Restarts	0.76 +/- 0.07	0.88 +/- 0.08	Crashes during calculations	Crashes during calculations

That is clearly visible that **GBR** shows **the least error**. **0.1 of °C** is a perfect result for machine learning. Transformers is really important step because it decreases deviation.

DT PREDICTION WITH GBR



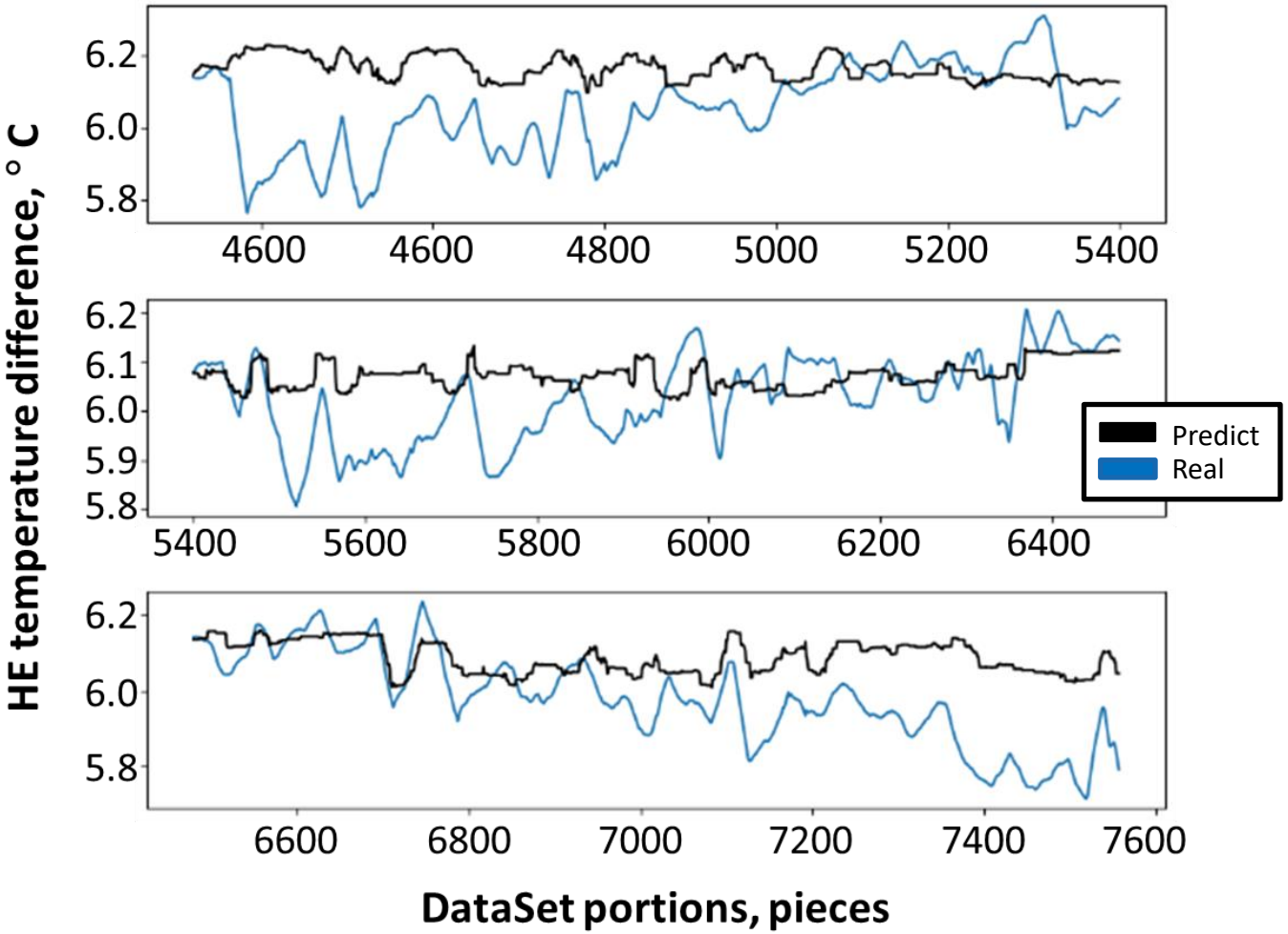
Default split (80/20) for training shows very good result.

Using the **simple square analysis** for the whole DataSet interval such results were obtained.

Prediction figure covers the real one on **0.90.**

So the **accuracy** of prediction for the model is nearly **90 %.**

DT PREDICTION WITH GBR



Final result is represented via graphs from Time Series, where blue line is a real state and black one – testing the model on unknown data.

The most important thing that model sees the peaks of the data, deviation from real state is not more **than 0.1÷0.2 °C.**

Even that model version can be used in manufacturing That deviation **can be smoothed** by adding more data and trainings.

CONCLUSION

- The **unique CSVRefactorer** utility was developed as **preprocessing software** for CSV-files from Simple SCADA system. That allows to reconstruct the files, divide the columns and differ time borders in **1 hour.**

- Transformers **(One Hot Encoder + Standart Scaler)** assembly was used for that study to make data in the more understandable for machine model way. That allows to decrease the error of modelling and improving the accuracy by **20 %.**

CONCLUSION

- The estimation of different machine learning models was made. According to the obtained data of temperature difference the best results is **Gradient Boosting Regressor** model with **accuracy of about 0.1 °C.**

- Using that model for month and week estimation for the personnel **90 % accuracy** of prediction might be obtained.

TOMSK POLYTECH

JOINT INSTITUTE FOR NUCLEAR RESEARCH

EFFICIENCY ASSESSMENT OF IRT-T RESEARCH REACTOR COOLING SYSTEM BY MACHINE LEARNING METHODS

Postgraduation student: Maxim K. Kublinskiy, group A3-43

E-mail: mkk4@tpu.ru

Phone number: +7 (913) 882–29–19

30 October 2023