

## Jury Member Report – Doctor of Philosophy thesis.

**Name of Candidate:** Viktor Duplyakov

**PhD Program:** Petroleum Engineering

**Title of Thesis:** Machine learning on field data for hydraulic fracturing design optimization

**Supervisor:** Professor Andrei Osipov

**Co-supervisor:** Professor Evgeny Burnaev

**Name of the Reviewer:** Alexander Shapeev

I confirm the absence of any conflict of interest	
<del>(Alternatively, Reviewer can formulate a possible conflict)</del>	<b>Date: 18-09-2023</b>

*The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed report to the thesis defense and to discuss the contents of each report with each other before the thesis defense.*

*If the reviewers have any queries about the thesis which they wish to raise in advance, please contact the Chair of the Jury.*

### Reviewer's Report

Reviewers report should contain the following items:

- Brief evaluation of the thesis quality and overall structure of the dissertation.
- The relevance of the topic of dissertation work to its actual content
- The relevance of the methods used in the dissertation
- The scientific significance of the results obtained and their compliance with the international level and current state of the art
- The relevance of the obtained results to applications (if applicable)
- The quality of publications

The summary of issues to be addressed before/during the thesis defense

The work under consideration is a good example of the application of modern data-driven modeling techniques to outstanding technological problems. In particular, in the work the problem of first predicting the yield of an oil well as a function of design parameters, and then - optimizing the yield is posed and solved.

I particularly praise the choice of the topic and methodology. The problem of predicting and optimizing oil production has been posed for many decades, however, it has not been satisfactorily solved by the conventional computational methods due to the complexity of the underlying processes. The modern data-driven method therefore offers a unique opportunity for us to progress on this problem, which has successfully been done in the thesis. The achieved predictive accuracy is of a clear scientific significance and the recommendation system developed to increase oil production is a notable practical outcome which has a potential to significantly impact decision-making processes in hydraulic fracturing design and optimize production and resource utilization.

Overall, the quality of the thesis is good. It is well-structured and easy to follow. All the aspects including introduction, data collection, forward and backward models, and testing and validation are well described. The language quality is adequate.

The thesis results are published in Q1 Scopus journals (2 papers) and presented in 2 conferences. Also, a patent was submitted and approved.

During the defense I would like to hear responses to the following questions:

- What are limitations and potential sources of bias in the dataset?
- Is the 64% predictive ability a lot or a little? Are there ways to further improve the model's performance?
- What are potential challenges for implementing your recommendation system in a real-world scenario?

#### **Provisional Recommendation**

☒ *I recommend that the candidate should defend the thesis by means of a formal thesis defense*

☐ *I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report*

☐ *The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense*