Well logging verification using machine learning algorithms

A. Tsanda, A. Bukharev, S. Budennyy (MIPT Center for Engineering and Technology)
A. Andrianova (Gazpromneft Science and Technology Center)

Introduction

Well logging analysis plays a crucial role in the design of oil field development. The analysis determines the location of the oil saturated reservoir and its thickness, which influences the estimation of oil reserves.

Results of well logging analysis are highly dependent on the expert knowledge level, in addition a great part of manual work is demanded. The geological data itself contains a great number of uncertainties, which complicate the interpretation problem making it stochastic. As a result, the interpretation problem is time-consuming and contains a high subjectivity.

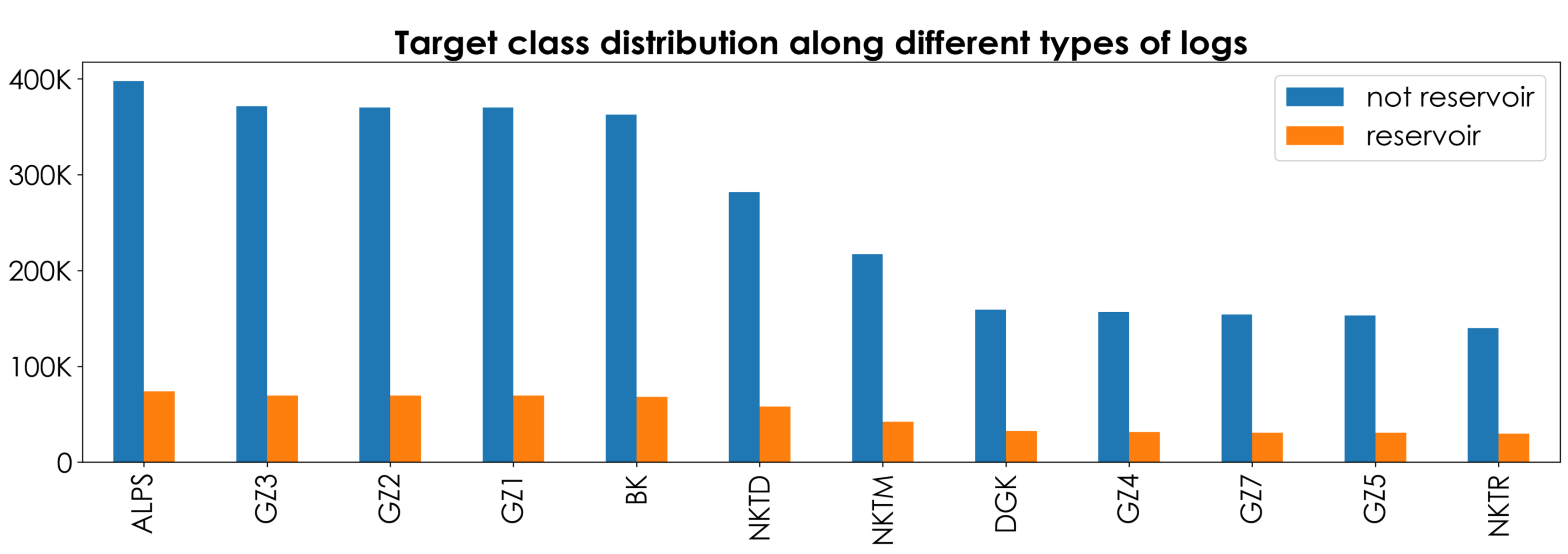
To reduce uncertainties in well logging data interpretation and make this process unbiased, an additional independent expertise is required. The paper proposes an independent expert system based on machine learning algorithms.

Workflow

We have considered the problem of reservoir identification along wellbore exploiting well logging data, which is reduced to the problem of binary classification. The first proposed methodology predicts reservoir presence by logging data for given depth while another one uses an interval of depths and localizes reservoir for the whole interval. Both approaches are based on machine learning algorithms (gradient tree boosting and convolutional neural network (CNN), correspondently, showed the highest results).

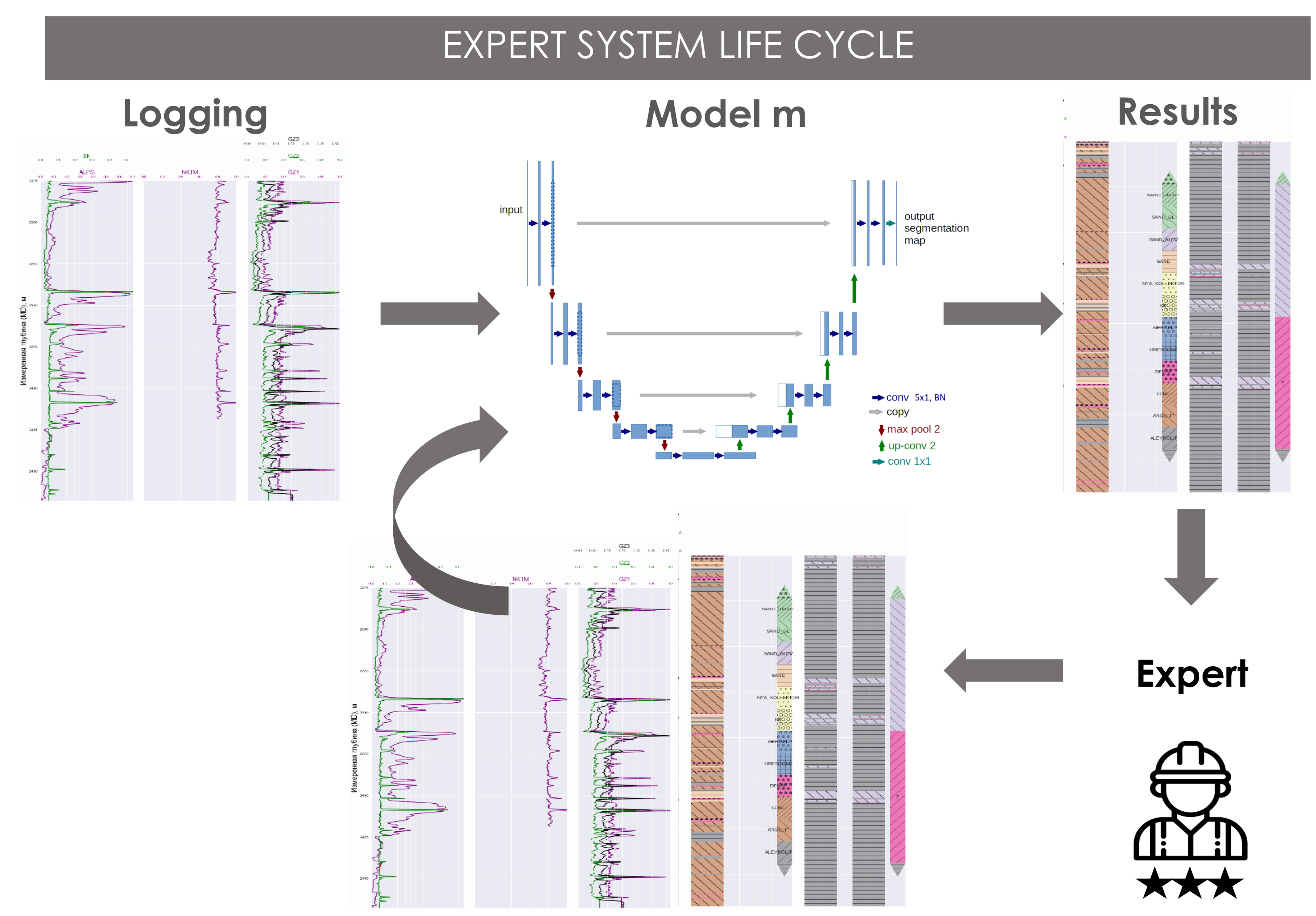
Logging data for training represents the dependence of the magnitude of the physical field on the borehole depth. The data was taken from the real oil field in Western Siberia (logging data was received from 287 wells).

The following types of logging were involved: GZ - lateral logging sounding; DGK – gamma ray logging; NKTD, NKTR – neutron logging; ALPHS – spontaneous potential logging; BK – laterlog. The target feature is binary one: indicates oil saturated reservoir presence for the given depth.

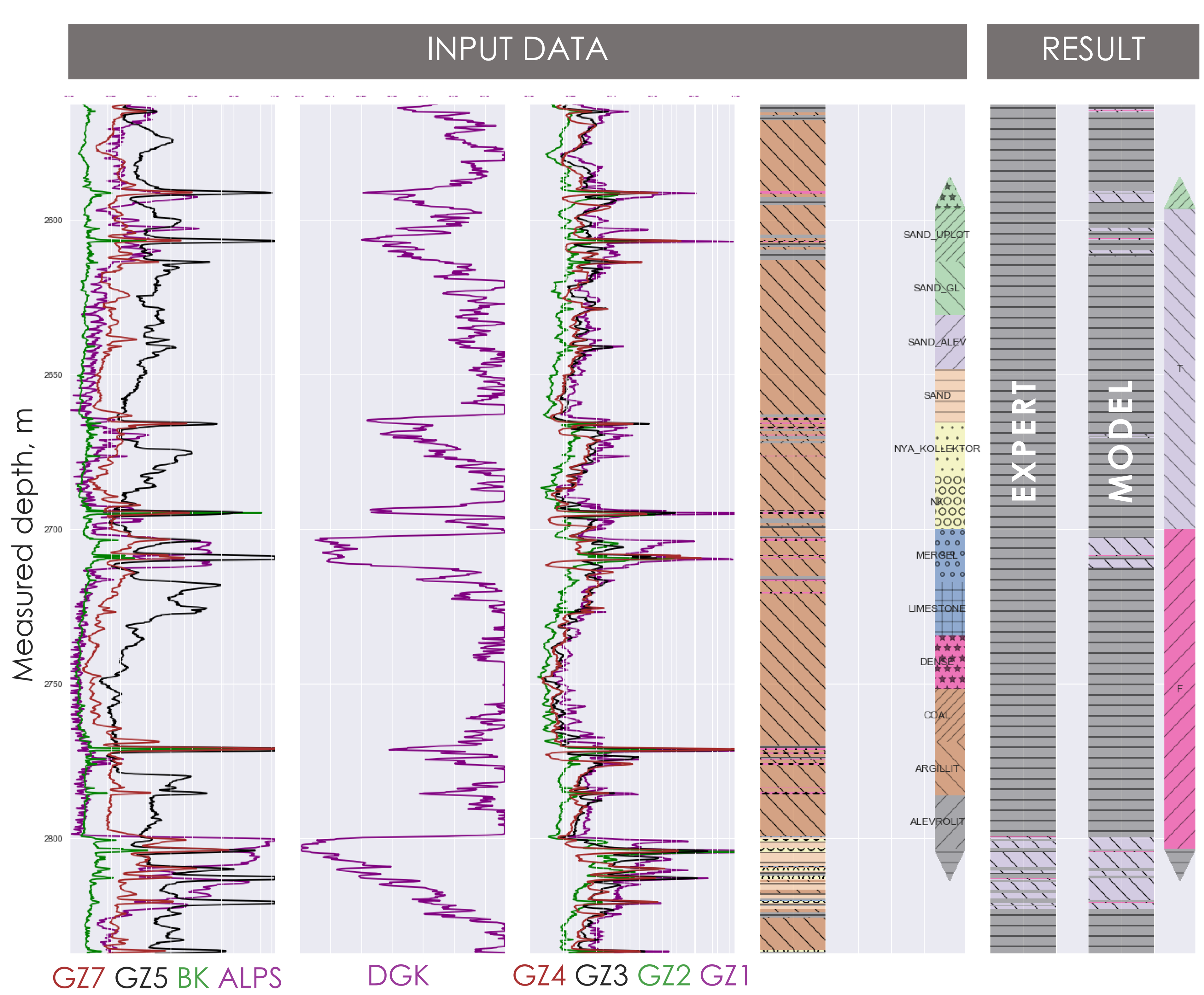


The crucial feature of the data is a misbalance of classes (see fig. above): well logging data is not provided uniformly. To overcome this obstacle class weights were added to the loss function.

The first method is more conventional and requires accurate data normalization. As some geo-physical well logging experiments were conducted for a specific range of depths, the data contains a great number of missing values (~36%). Thus, a robust algorithm is required to predict reservoir presence.



The second approach exploits the interval of depths to make prediction. This way of prediction is more like experts do logging analysis. Thus, the segmentation problem needs to be solved. The architecture of applied CNN was inspired by U-net, which was used for biomedical image segmentation. The CNN consists of two parts: encoder for feature extraction and decoder for making segmentation.



Results

F1-score was used to evaluate a prediction quality. The final score was the mean score over all folds in cross-validation (10 folds, the data was spitted according to wells). Prediction by interval of depths using CNN turned out to be more successful than prediction by sample using gradient tree boosting algorithm. F1-core consisted **0.63** for gradient boosting algorithm and **0.72** for CNN.

Conclusion

The paper introduced a segmentation approach to well logging data analysis. The approach was proven to be more successful (according to F1 score) than traditional approach using absolute values of logging data at specific depth.

The study shows the quality of well logging analysis using machine learning algorithms does not allow to fully automate the process, but it allows to use algorithm as a verification tool. In case of algorithm and human interpretation mismatch additional GWL could be conducted as well as additional expert analysis. Undertaken measures would help to decrease uncertainties and increase quality of well logging analysis.