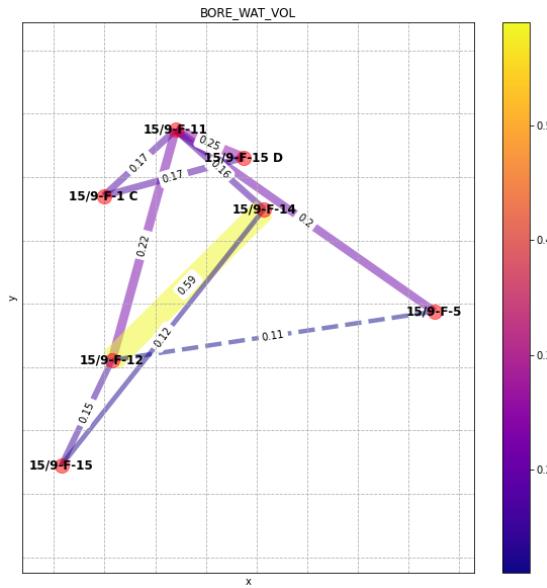


Team "Opps, I did it again"



Team "Opps I did it again"



Anton Voskresenskii
Data Scientist
Russia
ITMO University

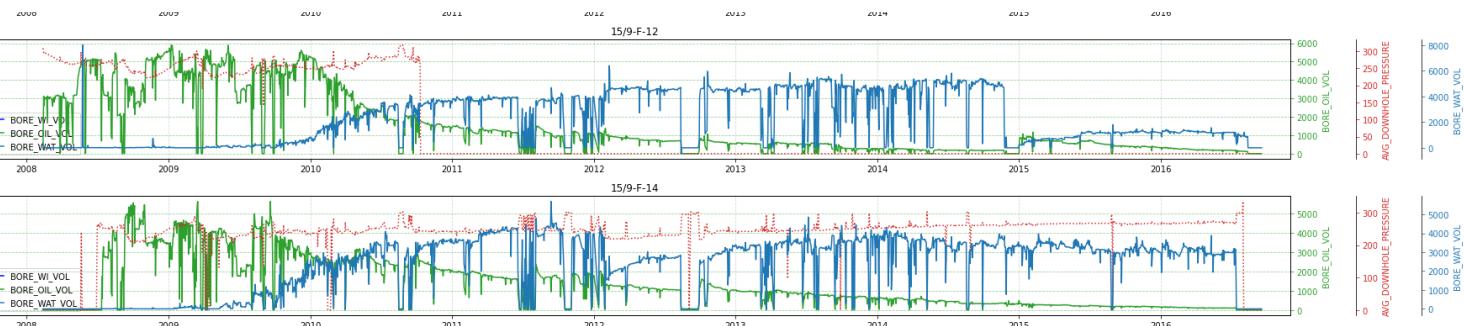
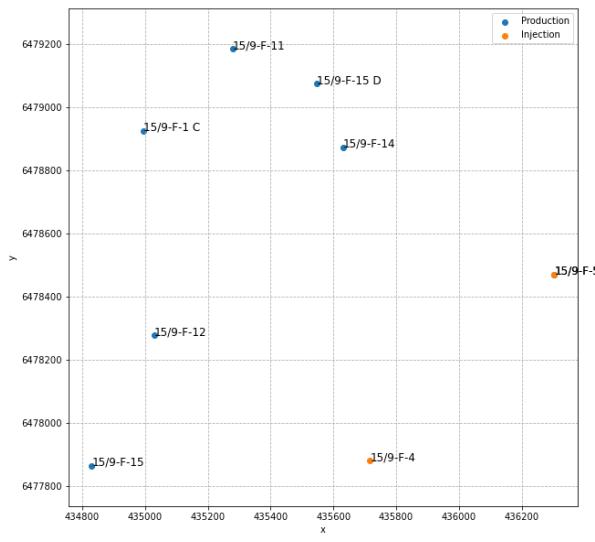
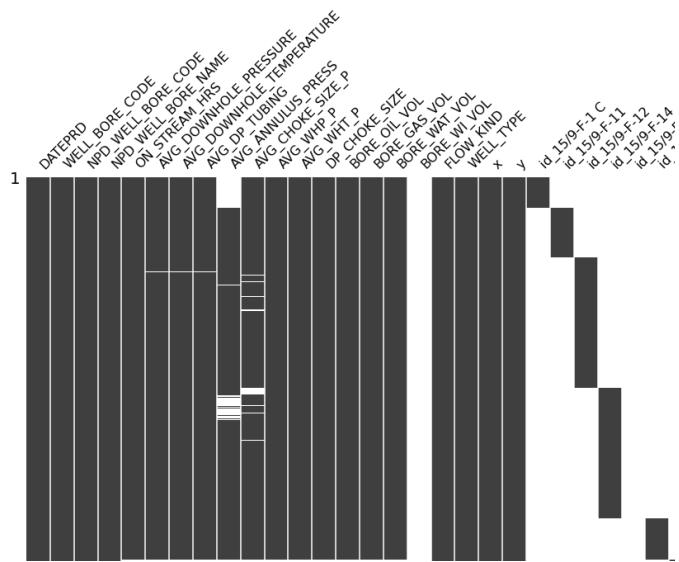


Elvira Selivanova
Economics & CS
New York
NY University



Daniil Zakharov
Data Scientist
Russia
MSU

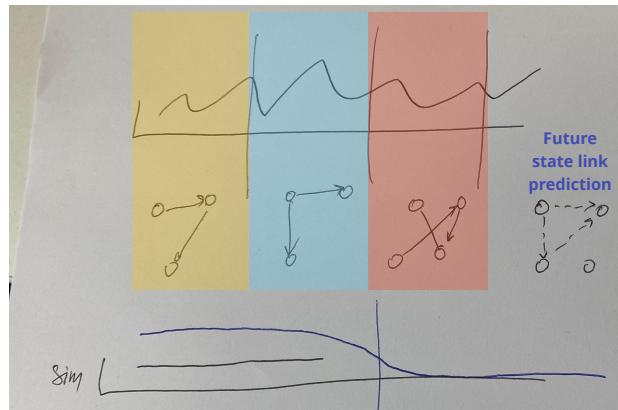
Volve dataset



Team "Opps, I did it again"

Task

- Develop a product that determine similarity between wells using machine learning approach.
It could be used for increasing quality of history matching
- Evaluate connectivity and interaction between production and injection wells and make predictions of future state of the reservoir using only wells that matter (have connectivity between target well)
- Approach could be applied to similar tasks in O&G such as feature selection of well logs or seismic attributed in order to predict some properties taking into account spatial relationship
- It can be applied in another industries such as retail



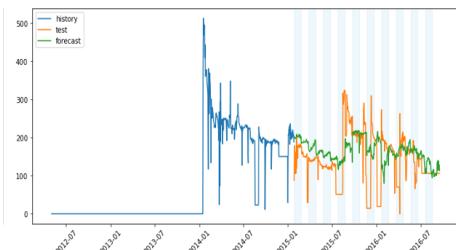
Team "Opps, I did it again"

Pipeline

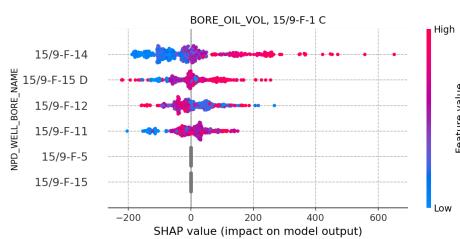
Build production profile forecasting model

Apply SHAP

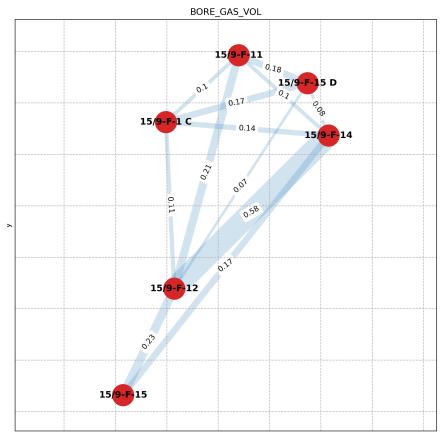
Build a number of graphs that represent spatio-temporal relationship between wells



- Time series 20 fold cross validation
- 20 days forecasting horizon



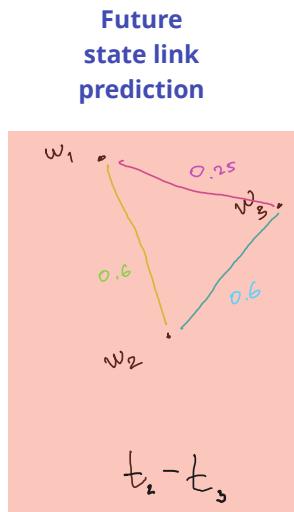
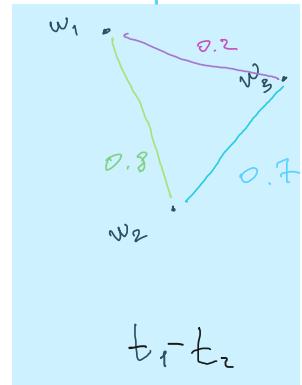
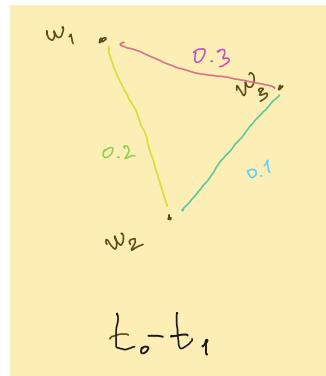
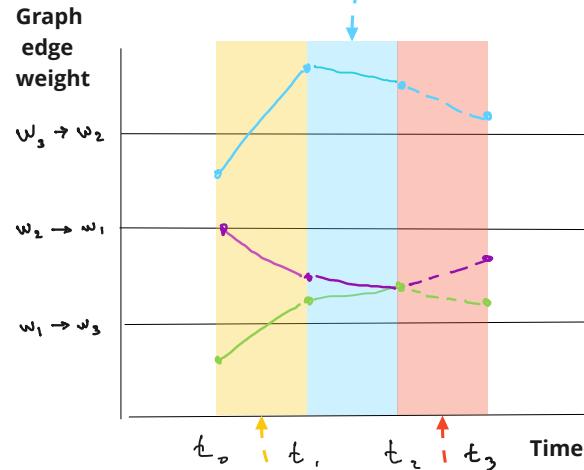
- Calculate Shapley values of features



- Absolute average value of each feature as graph edge weight
- Graph node signifies well location

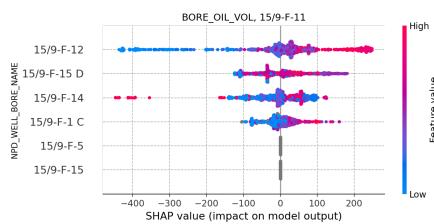
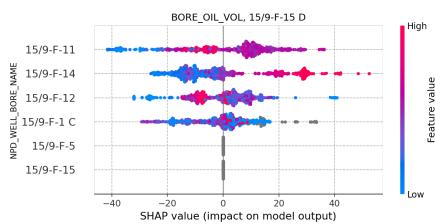
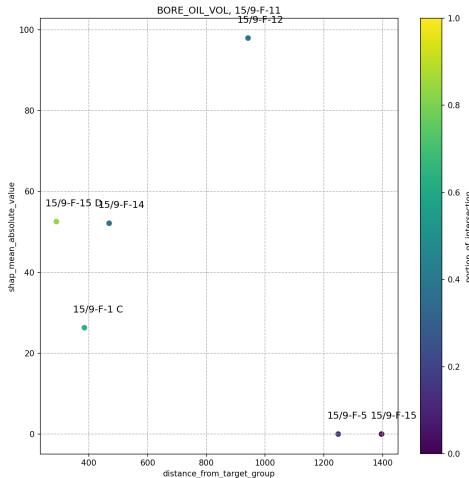
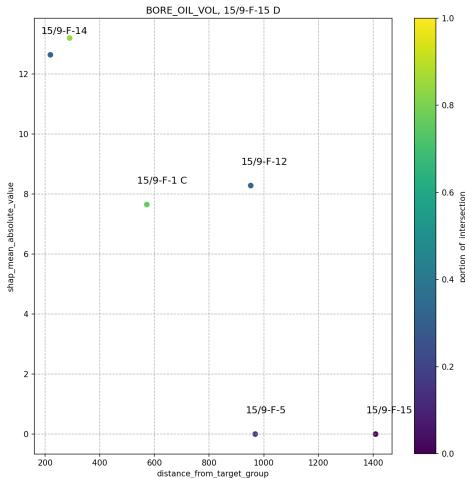
Team "Opps, I did it again"

Spatio-temporal relationship



Team "Opps, I did it again"

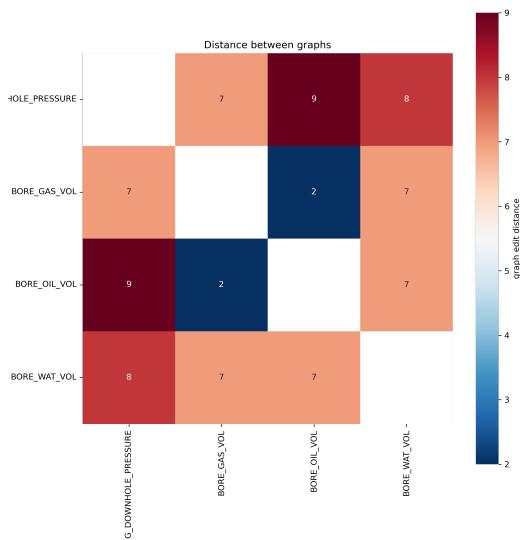
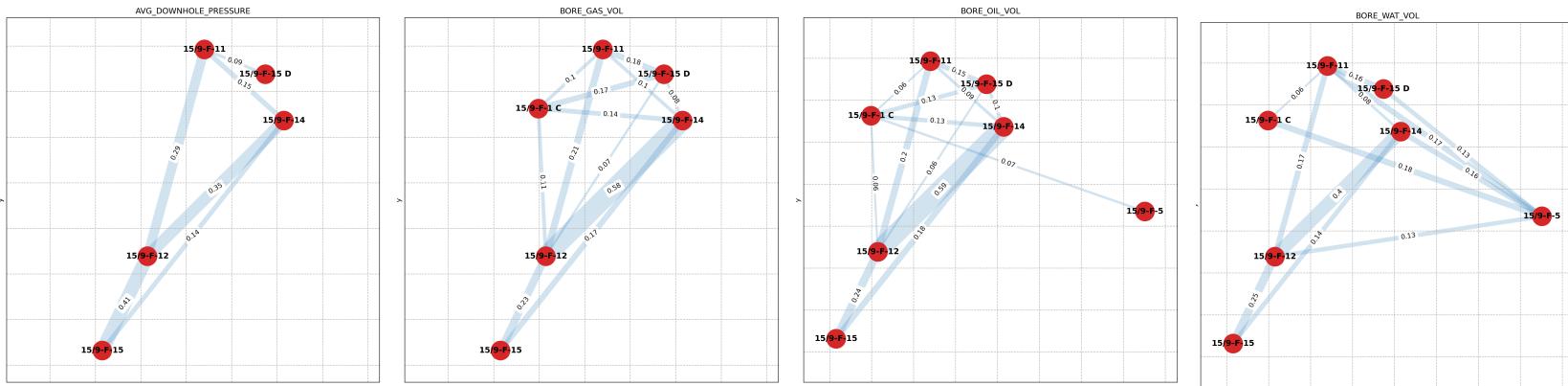
Proof of concept (1/2)



- The scatterplots show distance between target well and other wells vs Shapley value. Color of dots signifies normalised time when wells produced simultaneously
- **Nearest wells have the highest Shapley values**
- **Wells without intersection in time shave Shapley values close to zero**

Team "Opps, I did it again"

Proof of concept (2/2)



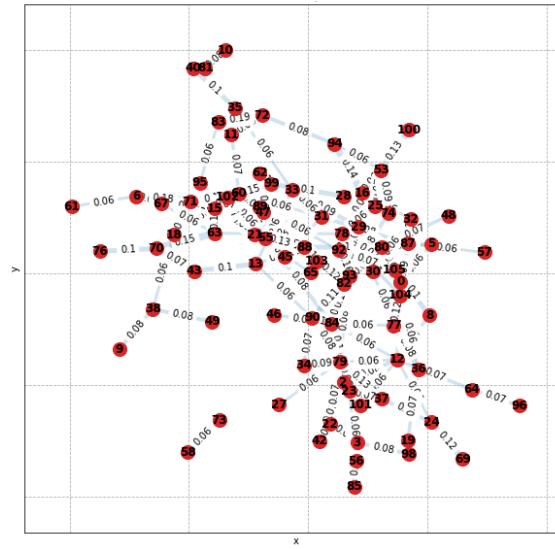
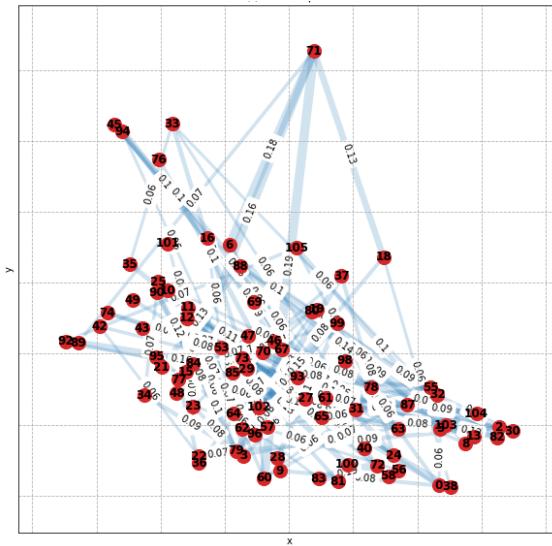
- Similar time series (oil and gas production) have the lowest GED* between each other
- Some edge weights (e.g. 15/9 F-12 → 15/9 F-14) demonstrate robustness despite using different data)
- Charts above show graphs for different targets
- Heatmap signifies GED between graphs

*GED (graph edit distance) - minimum cost of edit path (sequence of node and edge edit operations) transforming graph G1 to graph isomorphic to G2

Team "Opps, I did it again"

Further steps

- Apply approach on similar dataset from recent Russian hackathon (production profile forecasting)
- Build app that evaluates temporal relationship between graph edges
- Implement metric that take into account weight of graph edges
- Find out what to do with such a messy graphs :)



Team "Opps, I did it again"