AADE-22-FTCE-025

Digital Twins for Automated Treatment of Drilling Fluids

Mehrdad G Shirangi, Reza Ettehadi, Charles Thompson Jr, Edward Furlong, Baker Hughes

Abstract

In drilling operations, it is essential to maintain drilling fluids properties (e.g., density and rheological properties) within a reliable range. This ensures that various functionalities of drilling fluid such as cuttings transport and controlling formation pressure are effective. Treatment of drilling fluids typically involves manual calculations or reliance on experienced engineers. In this work, we present a procedure for automated real-time drilling fluid treatment. We consider the oil-based fluids. We first present and apply a procedure for building accurate digital twins for fluid rheological properties based on its composition. This involves the use of various machine learning algorithms. We then incorporate the predictive models in a prescriptive model that provides a new mix-sheet (allocation of chemicals) when new fluid properties are desired.

Introduction

During drilling operations, the drilling fluid (mud) is treated by adding proper allocation of the base oil, water, and different chemicals. The goal is to maintain or alter properties such as density (mud weight), rheological properties, salinity, HTHP properties, etc. The mud treatment process is typically performed through calculations by the mud engineer at rigsite. In this work, we consider a significant amount of data collected throughout years, particularly from laboratory experiments of drilling fluids when creating a mud plan for field operations. We develop accurate predictive models for drilling fluids rheological properties by utilizing latest algorithms and predictive analytics technologies. The models are then incorporated in a newly developed prescriptive analytics model that recommends proper allocation of different chemicals for automated mud treatment.

Drilling fluids have various critical roles for the success of drilling operations (Mitchell and Miska, 2011). These include, for instance, transporting the cuttings to the surface (Nazari et al., 2010), fluid loss control, and suspending the cuttings when circulation stops. In a previous work, digital twins were developed for monitoring drilling fluid's HTHP rheological properties (Samnejad et al., 2020) and cuttings transport functionality (Shirangi et al., 2020a).

Drilling fluid rheological properties are measured at surface conditions using rotational viscometer (an inline rheology measurement unit) and by sensors (Ofei et al., 2021). Rheological properties are characterized by the measurement of shear stress at some standard viscometer RPM. The standard viscometer RPMs are 3, 6, 100, 200, 300, and 600, according to API Recommended Practice 13D (API, 2006). In this work, we use the Herschel-Bulkley model for describing drilling fluid rheological behavior (shear-stress shear-rate relationship). The Herschel-Bulkley model is shown to be the model better describing the rheological behavior of most drilling fluid (Kelessidis et al., 2011; Saasen & Ytrehus, 2018).

The effect of allocation of different chemicals on rheological properties of drilling fluids has been considered in previous work. (Sami, 2016) presented experimental results to study the effect of magnesium salt contamination. Various authors presented machine learning approaches for predicting rheological properties using other measurement data as inputs (Oguntade et al., 2020; Liu et al., 2021; Shirangi et al., 2021; Makinde et al. 2011). To the best of our knowledge, the use of chemical allocations as inputs for predicting rheological properties of oil-based drilling fluids has not been considered in previous work.

The rest of this paper is organized as follows. Methodology is described in the next section. The section after that contains results and discussion. In the final section, we present conclusions.

Methodology

Our methodology for developing a prescriptive design model involves: 1) setting up data pipeline to ingest raw data into databases, perform data preprocessing, data cleansing, and other data engineering tasks, 2) predictive analytics and generating accurate digital twins; and 3) incorporating predictive models into a prescriptive data analytics model that provides exact change in mud composition to achieve a desired set of mud properties. Our goal at step 2 is to predict drilling fluid properties by utilizing compositional analysis data and pressure and temperature data. We apply an automated predictive analytics tool to perform massive computations for feature selection, feature engineering, and model selection. In Step 3, we develop a prescriptive data analytics model that given the set of "desired mud properties" and the current ones, the model recommends the appropriate change in allocation of different chemicals of drilling fluid. This is achieved through formulating and solving an optimization problem.

Development of a Digital Twin

The schematic of our process for building a digital twin is demonstrated in **Figure *1***. The first step of building a digital twin is collecting relevant data. The data is ingested into a database and a data pipeline is set up to cleans, parse, transform, and enrich raw data to obtain preprocessed data in corresponding tables. After merging data and obtaining a master table, machine learning algorithms are applied to obtain an accurate predictive model for each desired target variable.

Data Acquisition and Processing

Data from various aspects of drilling fluids has been considered. After careful investigation, data from drilling fluids lab experiments were chosen due to their accuracy in controlled conditions. The majority of these data were obtained when designing a drilling fluid composition at well planning stage. In this process, the lab technician receives an order with a set of desired specifications and a selected set of chemicals. The specifications include the specified density (mud weight) for each section of the well, range of desired electrical stability (ES), range of desired rheological properties (readings at different RPMs, 10 minute and 10 second gel strength value, yield point, etc.) and in some cases, a range of desired fluid loss properties, and specifications for static-aged properties. The technician comes up with an initial fluid composition (mix-sheet) and iteratively changes the allocation of different chemicals to ultimately obtain a composition whose properties meets the desired specifications. This process is time-consuming as each new composition has to go through a hot-rolling process, and in cases, static-aging process first before measuring its properties. Hot-rolling process may take 24 hours and the static-aging process typically takes longer times.

A data ingestion pipeline is developed to automate the process (**Figure 2**). Each report is loaded and parsed into proper table format to be saved in a database. It is not uncommon to use different names to refer to the same chemical. Therefore, a reference dictionary is created as “similar items” where each key is the designated name of the chemical, and the values correspond to the different names used in the reports. The parsed reports together with the “similar items” dictionary are fed into a postprocessing pipeline to obtain postprocessed reports in new tables. Each postprocessed report is a table that is enriched with additional features for each record. The enriched features are either binary (such as hot-rolled/static-aged, presence of contaminant) or continuous (such as hot-rolling temperature).

In the postprocessing step, each feature is tagged with a keyword. For chemicals, the tags include base-oil, rheology modifier, weight material, salt or brine, emulsifier. These tags are provided in the item-types dictionary for each chemical. The postprocessed tables are then merged to create a master table.

Automated machine learning for predictive analytics

Our goal is to obtain accurate predictive models for drilling fluid properties where the input parameters include the fluid composition (allocation of different chemicals), temperature, and additional features such as hot-rolling temperature. In this work, we focus on prediction of rheological properties; however, the same process applies for prediction of other properties. This prediction problem can be formulated as a regression problem where algorithms such as artificial neural network (ANN), support vector regression, and gradient boosting regression are considered.

We utilize a previously developed automated machine learning (autoML) tool (Shirangi et al., 2020a,b) here. Creating an accurate machine learning model involves the process of model selection where for each algorithm considered, the hyperparameters are optimized to obtain the best model (the model with the highest accuracy) from that algorithm. The final selected model corresponds to the algorithm that provides the highest accuracy. The Hyperparameters are the parameters that must be specified a priori and are not learned during training. In case of ANN, the neural network architecture (the number of hidden layers and the number of hidden neurons at each layer) and the choice of activation function (sigmoid versus ReLU) are considered the hyperparameters.

We use the coefficient of determination (R2) evaluated through K-fold cross-validation process as the metric to determine the best model.

Prescriptive Design for Automated Treatment

We consider adjustments for density and rheological properties. When adjusting density, we use a rule-based approach. Increasing density typically involves increasing allocation of weight material(s) whereas reducing it, involve dilution where the new fluid addition corresponds to smaller allocation of weight material and higher allocation of brine and base oil.

When treating for rheological properties, allocations of rheology modifiers are iteratively changed until the obtained solution corresponds to properties within a tolerance of the desired ones. This is achieved through the use of an optimization algorithm and the machine learning predictive model. The flowchart of the algorithm is presented in **Figure *3***.

Results and Discussion

We consider data for oil-based drilling fluids in global offshore operations. Data includes measurements for initial mix, hot-rolled samples, and static-aged samples. We only considered the hot-rolled samples to closely approximate conditions seen in the field. Samples that included contaminants were removed.

After performing preprocessing, postprocessing, and merging data, the master table included 5,000 data records for hot-rolled samples. Total number of features (chemicals and enriched features such as temperature and oil-to-water ratio) is 93. Each data record is sparse as only 15% of entries are non-zeros. Data is split into train (80%) and test (20%) datasets. The algorithms are trained on the train data set and test data set is only used to assess the accuracy of the best model on unseen data once model selection is completed.

An example of performance of different algorithms is shown in **Figure 4** where the target parameter is 600 rpm reading. The highest mean R2 from K-fold cross-validation (with K=5) is obtained for the XGBoost algorithm. This model is therefore selected for deployment. Scatter plots for test data and train data versus prediction for the selected model (XGBoost) and the second-best model (neural network) are shown, respectively, in **Figure *5*** and **Figure *6***. It is evident that the XGBoost algorithm matches the training data very closely but its prediction for test data shows a similar spread to that from neural network. The R2 score for prediction of test data for both models are similar and equal to 0.96.

We also applied the autoML to generate model for other RPM readings. The accuracy of the best models is shown in Figure 7. All models correspond to a mean R2 above 0.8. Note that the yield point and plastic viscosity can be obtained from the predictions at 600 RPM and 300 RPM.

An example of prescriptive design is presented here. The chemical composition of the initial mix is shown in **Table *1***. The chemical composition is shown for chemical types as we were asked not to disclose the actual names of the chemicals. The sample is hot-rolled, and its properties are measured, displayed in **Table 2**. The hot-rolled properties from the initial mix violate the specifications for 3 rpm reading and for the yield point. The prescriptive design algorithm was able to obtain a solution whose properties satisfy those specifications.

Conclusions and Future Work

In this work, we developed digital twins for predicting drilling fluid rheological properties given its composition and other parameters. The results demonstrate that fairly accurate models could be obtained for RPM readings.

A prescriptive model is then developed utilizing the machine learning models and an optimization algorithm. Given an initial mix-sheet (allocation of chemicals) and the range of specifications, the algorithm provides an updated mix-sheet that satisfies the specifications while the density and oil-water-ratio are unchanged.

The prescriptive model will be incorporated in an automated drilling fluid (mud) treatment system at rigsite. We envision a system based on artificial intelligence (AI) trained on past historical field/lab data that makes recommendations about the mud treatment. This AI system makes the recommendations based on real-time rig data and data from sensors on various drilling fluid properties.

We considered oil-based drilling fluids in this work. Extension to water-based fluids should be investigated in future. In addition, we only considered the lab experiment data here. Collecting and analyzing drilling fluids rig data should be performed. Laboratory and field fluids can act differently sometimes due to formation effects and presence of low gravity solids, etc.

Acknowledgments

We thank the Baker Hughes Company for permission to present this work.

Nomenclature

ANN = *Artificial neural network*
AutoML = *Automated machine learning*
RPM = *Round per minute*
R2 = *coefficient of determination*

References

1. API. (2006). Recommended Practice on the Rheology and Hydraulics of Oilwell Drilling Fluids.
2. Kelessidis, V. C., Dalamarinis, P., & Maglione, R. (2011). Experimental study and predictions of pressure losses of fluids modeled as Herschel–Bulkley in concentric and eccentric annuli in laminar, transitional and turbulent flows. *Journal of Petroleum Science and Engineering, 77*, 305–312.
3. Liu, N. a. (2021). Real-Time Measurement of Drilling Fluid Rheological Properties: A Review. *Sensors*.
4. Makinde, F. a. (2011). Modelling the effects of temperature and aging time on the rheological properties of drilling fluids. *Petroleum & Coal*, 167--182.
5. Mitchell, R., Miska, S. (2011). *Fundamentals of drilling engineering.* Society of Petroleum Engineers.
6. Nazari, T., Hareland, G., & Azar, J. J. (2010). Review of cuttings transport in directional well drilling: systematic approach. *SPE Western Regional Meeting.*
7. Ofei, T. N., Lund, B., Saasen, A., & Sangesland, S. (2021). The Effect of Oil-Water Ratio on Rheological Properties and Sag Stability of Oil-Based Drilling Fluids. *Journal of Energy Resources Technology*, 1–26.
8. Oguntade, T. a. (2020). Application of ANN in predicting water based mud rheology and filtration properties. *SPE Nigeria Annual International Conference and Exhibition.*
9. Saasen, A., & Ytrehus, J. D. (2018). Rheological properties of drilling fluids: use of dimensionless shear rates in herschel-bulkley and power-law models. *Applied Rheology, 28*.
10. Sami, N. A. (2016). Effect of magnesium salt contamination on the behavior of drilling fluids. *Egyptian Journal of Petroleum*, 453--458.
11. Samnejad, M., Shirangi, M. G., & Ettehadi, R. (2020). A digital twin of drilling fluids rheology for real-time rig operations. *Offshore Technology Conference.*
12. Shirangi, M. G., M., Aragall, R., Ettehadi, R., et al. (2021). Development of Digital Twins for Drilling Fluids: Local Velocities for Hole Cleaning and Rheology Monitoring. *International Conference on Offshore Mechanics and Arctic Engineering.* American Society of Mechanical Engineers.
13. Shirangi, M. G., Ettehadi, R., Aragall, R., et al. (2020a). Digital twins for drilling fluids: advances and opportunities. *IADC/SPE International Drilling Conference and Exhibition.* Galveston, Texas, USA: SPE-199681-MS. https://doi.org/10.2118/199681-MS

14. Shirangi, M. G., Furlong, E., Sims, K. S. (2020b). "Digital Twins for Well Planning and Bit Dull Grade Prediction." *SPE Norway Subsurface Conference*. https://doi.org/10.2118/200740-MS

Table 1: initial mix and the solution mix from prescriptive design algorithm.

Component	Unit	initial mix	solution
Emulsifier	lb/bbl	14	12
Rheological Modifier	lb/bbl	2.25	2.25
Rheology / Filtration 1	lb/bbl	3	3.5
Rheology / Filtration 2	lb/bbl	4	4
Rheology / Filtration 3	lb/bbl	4	5
Salt	lb/bbl	35.12	34.98
Weight Material	lb/bbl	191.9	191.73
water	bbl/bbl	0.2846	0.2834
Base Oil	bbl/bbl	0.4834	0.4877

Table 2: Properties of the initial mix versus those from the solution. It is evident that the solution satisfies the specifications. The properties were measured at hot-rolled condition.

Properties	Specifications	initial mix	solution
ES @ 120ºF, volts	**>400**	550	540
600 rpm		69	90
300 rpm		43	58
200 rpm		34	46
100 rpm		24	33
6 rpm	**12-20**	**12**	**18**
3 rpm	**12-18**	**11**	**18**
Plastic Viscosity, cP		26	32
Yield Point, lb/100 ft^2	**25-38**	**17**	**26**
10 Second Gel, lb/100 ft^2	**12-22**	**15**	**23**
10 Minute Gel, lb/100 ft^2	**14-34**	**27**	**32**
Oil to Water Ratio	65:35	65:35	65:35
Mud Weight	11.7	11.7	11.7

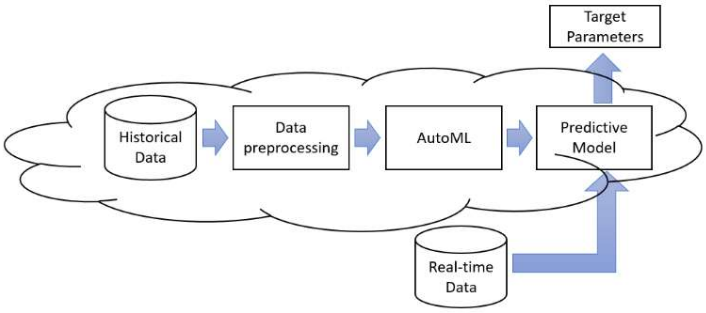

Figure 1: Workflow for building a digital twin. The predictive model takes the real-time parameters as input to produce the estimates for target parameters.

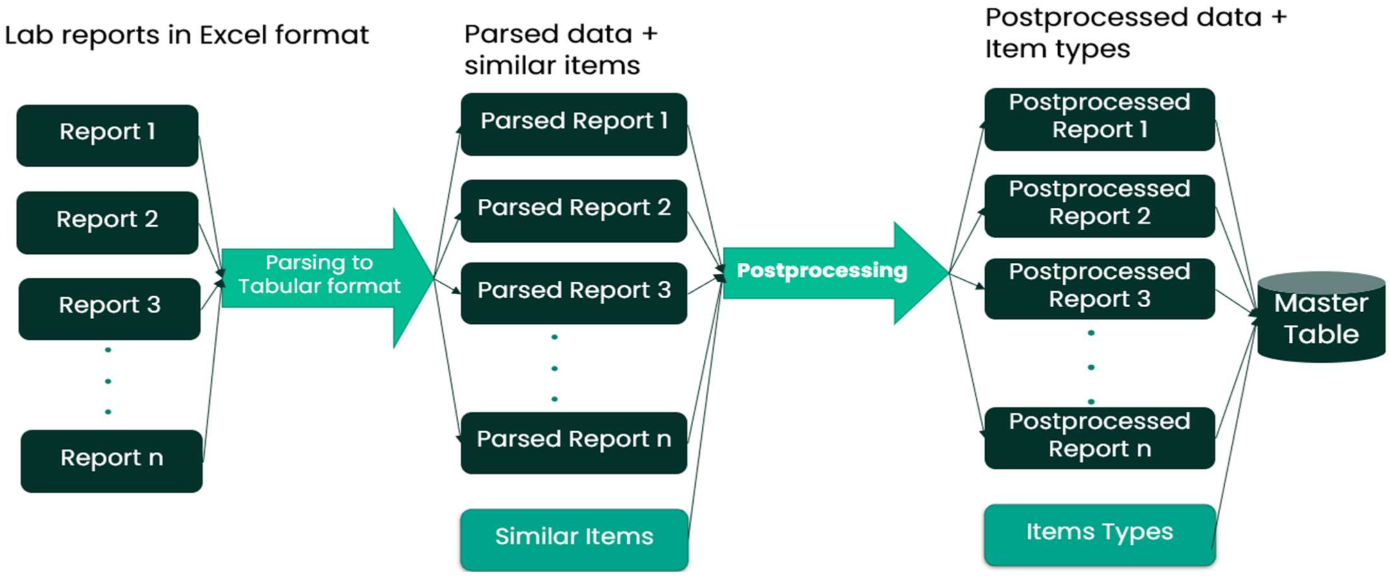

Figure 2:Schematic of data ingestion and processing pipeline

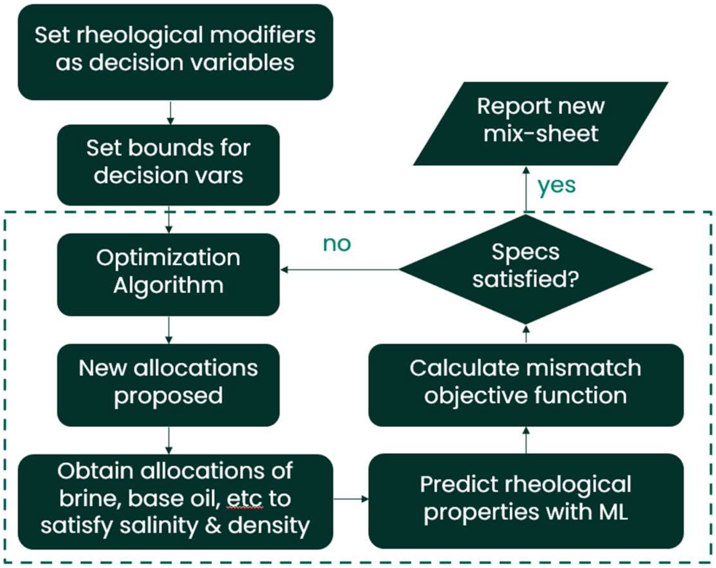

Figure 3: Flow chart of prescriptive design for obtaining a new mix-sheet. The dashed box contains the optimization element.

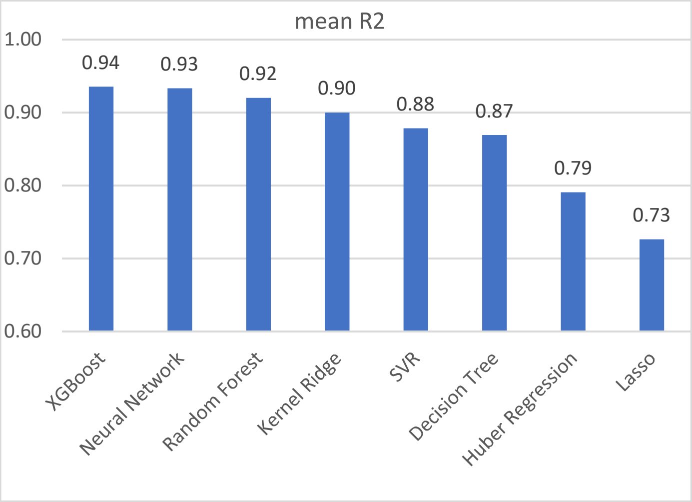

Figure 4: Mean R2 from K-fold cross-validation (K=5) for the best model from different algorithms for prediction of 600 rpm reading (rheological properties).

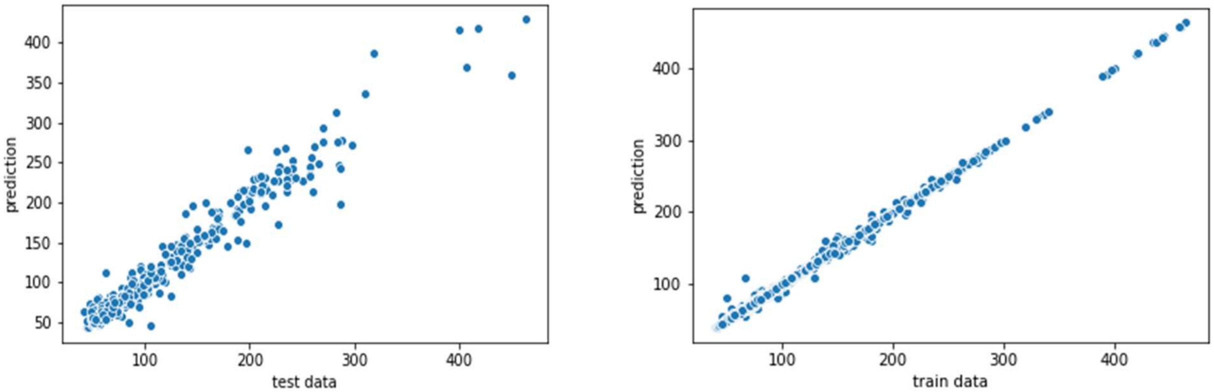

Figure 5: Predictions versus test data (left) and versus train data (right) for 600 rpm reading (XGboost). Predictions are from XGBoost model with a max_depth of 5 and n_estimators set to 1,000.

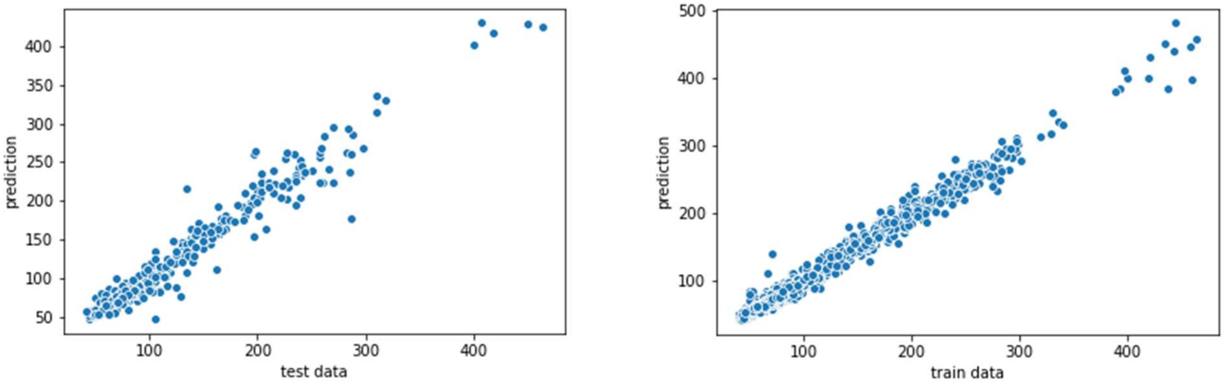

Figure 6: Predictions versus test data (left) and versus train data (right) for 600 rpm reading (Neural Network). Predictions are from a fully connected neural network model with seven hidden layers each having 280 or 281 hidden neurons, and all activation functions set to ReLU.

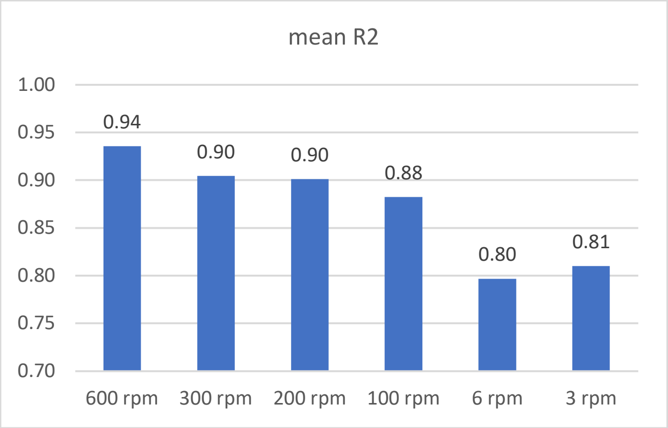

Figure 7: Mean R2 from K-fold cross-validation (K=5) for the best model for prediction of different rpm readings (rheological properties)