MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent institution of MAHE, Manipal)

INDUSTRIAL TRAINING

BY-ARUSHI SINHA
MANIPAL INSTITUTE OF TECHNOLOGY

TOPIC

PSEUDO SONIC LOG GENERATION USING MACHINE LEARNING

CONTENTS

ABSTRACT

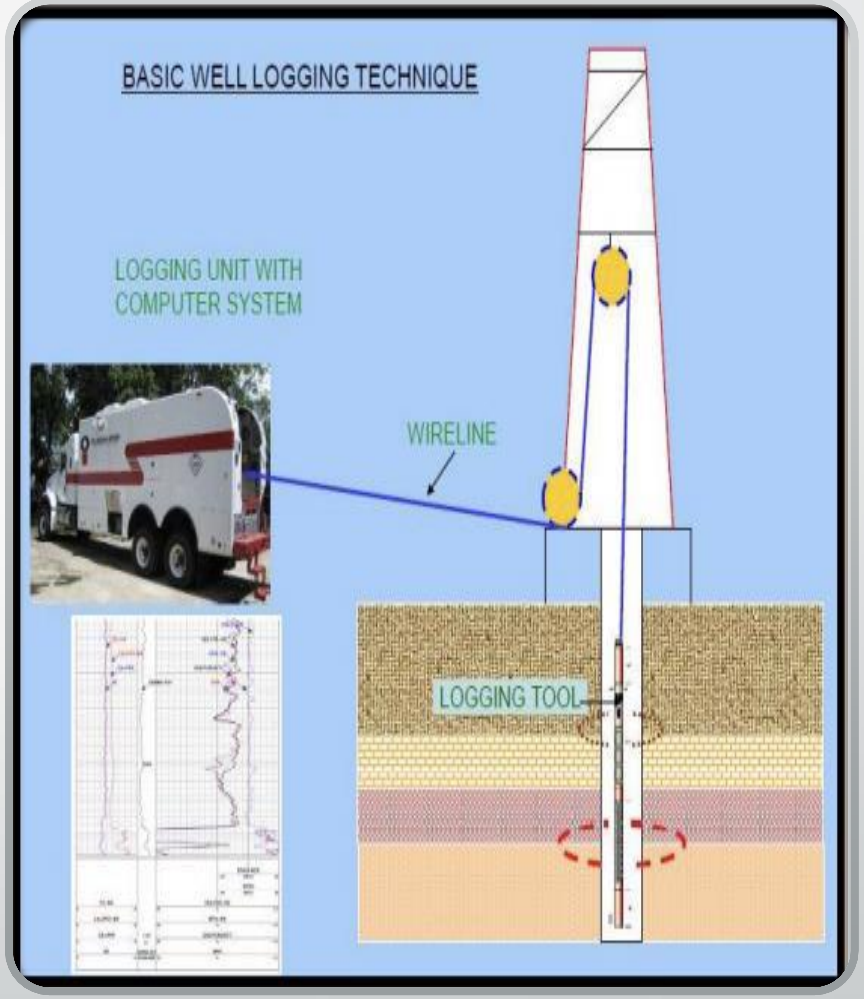
- The following project shines light on application of Artificial Intelligence in Oil Industry. The process of oil extraction is a complicated one and requires drilling of wells kilometres inside the Earth’s Crust.
- Well log data are the eyes of the drillers which guides them while drilling boreholes deep into the Earth. The data is measured though logging tools which are suspended into these boreholes against huge mud and formation pressures.
- Certain well logs, like gamma ray (GR), resistivity, density, and neutron logs, are considered as “easy-to-acquire” conventional well logs that are run in most wells.
- Other well logs, like nuclear magnetic resonance, dielectric dispersion, elemental spectroscopy, and sometimes sonic logs, are only run in a limited number of wells.
- When sonic logs are absent in a well or an interval, a common practice is to synthesize them based on neighbouring wells that have sonic logs. This is referred to as sonic-log synthesis or pseudo-sonic log generation.

INTRODUCTION

- The process of crude oil and natural gas exploration is not an easy one, especially with oil reserves depleting around the world.
- The only way of extracting these reserves is by drilling a borehole into it, which is called a well. However, before drilling a well we need the exact location of these reserves in the earth. We also need to guide our borehole as it drills into the Earth's crust. This is where Well Logging and Data Interpretation comes in.
- Well logging, also known as borehole logging is the practice of making a detailed record (a well log) of the geologic formations (subsurface rocks), penetrated by a borehole.
- The log may be based either on visual inspection of samples brought to the surface (geological logs) or on physical measurements made by instruments (Logging Tools) lowered into the borehole.

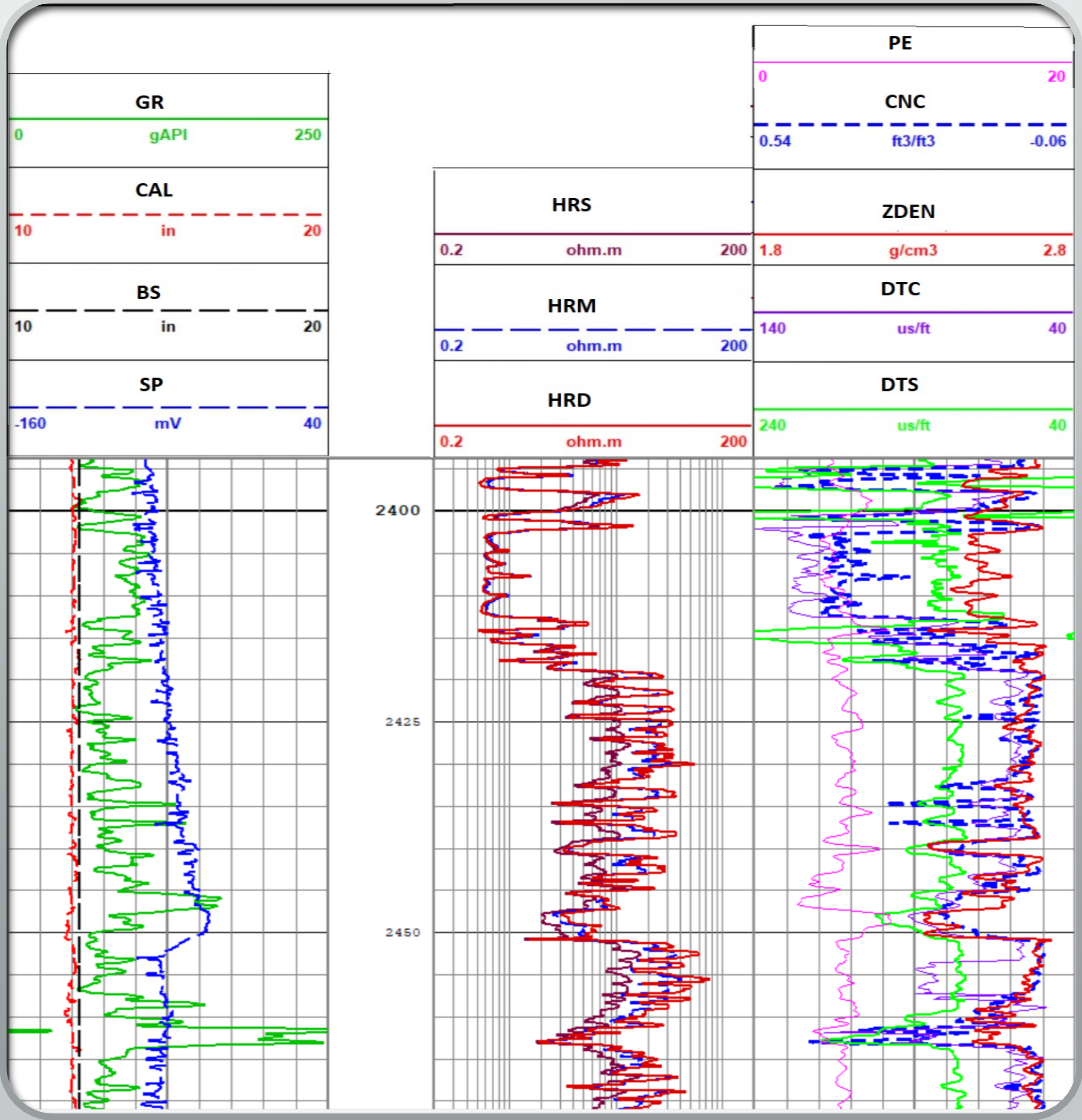
WELL LOGGING BACKGROUND

- **Logging Unit**: A specialized truck installed with a full computer system for log data acquisition & processing.
- The logging tool is lowered into the wellbore by means of the logging cable or wireline. The wireline also connects the logging tool electrically to the surface computer system.
- Data acquired by the tool are transmitted to the surface system over the logging cable using digital telemetry. The surface computer records, processes and plots these data as a function of well depth and produces a “log” or “well log”.



WELL LOG

- Basic Open Hole Logging Measurements:
- Caliper
- Gamma Ray
- Resistivity
- Density
- Photelectric Effect
- Neutron Porosity
- Acoustic Compressional
- Acoustic Shear



FORMATION EVALUATION

- When a well is drilled and well logs have been acquired in a well a quick look interpretation of logs is done to provide answers to the following questions:
  - What zones are potential hydrocarbon producers?
  - Will it be gas or oil?
  - Approximate volumetric analysis?

This evaluation results in following properties:

- Porosity:
- Water Saturation
- Thickness

These parameters are called petrophysical parameters.

SONIC LOG SYNTHESIS

- Sonic travel-time logs contain critical geo-mechanical information for subsurface characterization around the wellbore.
- Often, sonic logs are required to complete the geo-mechanical properties prediction.
- When sonic logs are absent in a well or an interval, a common practice is to synthesize them based on neighbouring wells that have sonic logs. This is referred to as sonic-log synthesis or pseudo sonic log generation.
- Machine learning has gained increasing momentum in petrophysical applications in recent years. It is imperative to prove the capability of machine learning in solving real petrophysics problems.

OBJECTIVE

- During my internship at ONGC, I undertook the study of Machine Learning algorithms on Pseudo Sonic Log generation from available recorded logs.

- I have generated the missing sonic logs using the input data provided to the SPWLA (Society of Petrophysicists and Well Log Analysts) contestants. After successfully implementing this, I have generated the same missing sonic logs using the open data and compared the accuracy of both the results.

PROBLEM STATEMENT

- Compressional-wave travel time (DTC) and shear-wave travel time (DTS) logs are not acquired in all the wells drilled in a field due to financial or operational constraints. Under such circumstances, machine-learning techniques can be used to predict DTC and DTS logs to improve subsurface characterization.
- The goal is to develop data-driven models by processing “easy-to-acquire” conventional logs from Well 1 and using the data-driven models to generate synthetic compressional and shear travel-time logs (DTC and DTS, respectively) in Well 2.
- A robust data-driven model for the desired sonic log synthesis should result in low prediction errors, which could be quantified in terms of Root Mean Square Error by comparing the synthesized and the original DTC and DTS logs.
- The data-driven model used feature sets derived from the following eight logs: Caliper, Neutron, Gamma ray, Deep resistivity, Medium resistivity, Shallow resistivity, Photoelectric factor and Bulk Density. The data-driven model synthesized two target logs: DTC and DTS.

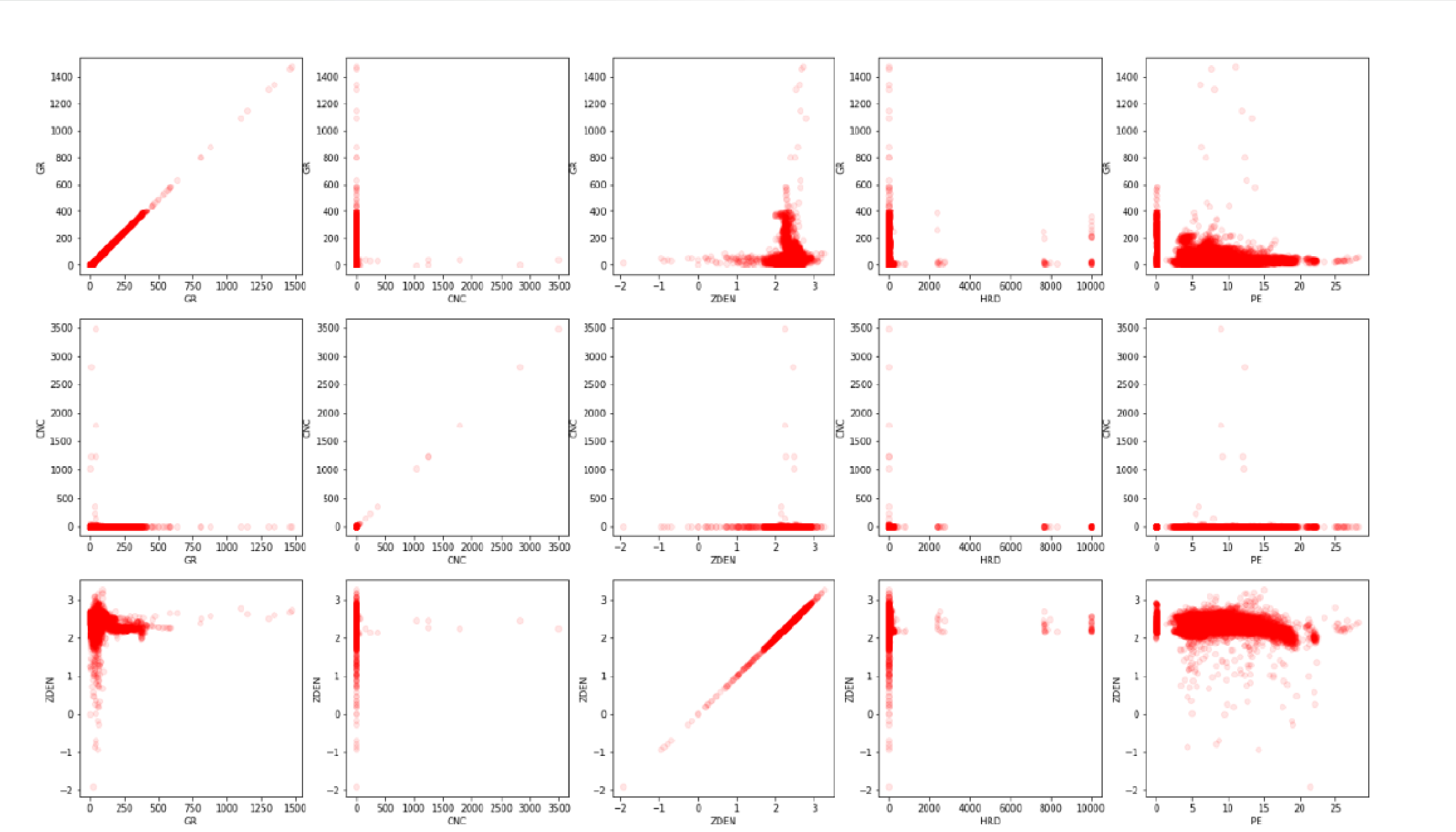
DATA DESCRIPTION

Log Name	Units	Well 1	Well 2
Caliper (CAL)	Inches	Yes	Yes
Neutron (CNC)	Dec	Yes	Yes
Gamma Ray (GR)	API	Yes	Yes
Deep Resistivity (HRD)	Ohm•m	Yes	Yes
Medium Resistivity (HRM)	Ohm•m	Yes	Yes
Photoelectric Factor (PE)	Barn	Yes	Yes
Density (ZDEN)	Gram/m3	Yes	Yes
Compress Travel time (DTC)	µs/foot	Yes	
Shear Travel time (DTS)	µs/foot	Yes	

- Data with varying log properties was provided from the VOLVE open dataset courtesy of Equinor.
- The table shows the log availability for provided wells.
- A yes denotes that the well was provided with that type of log for analysis.

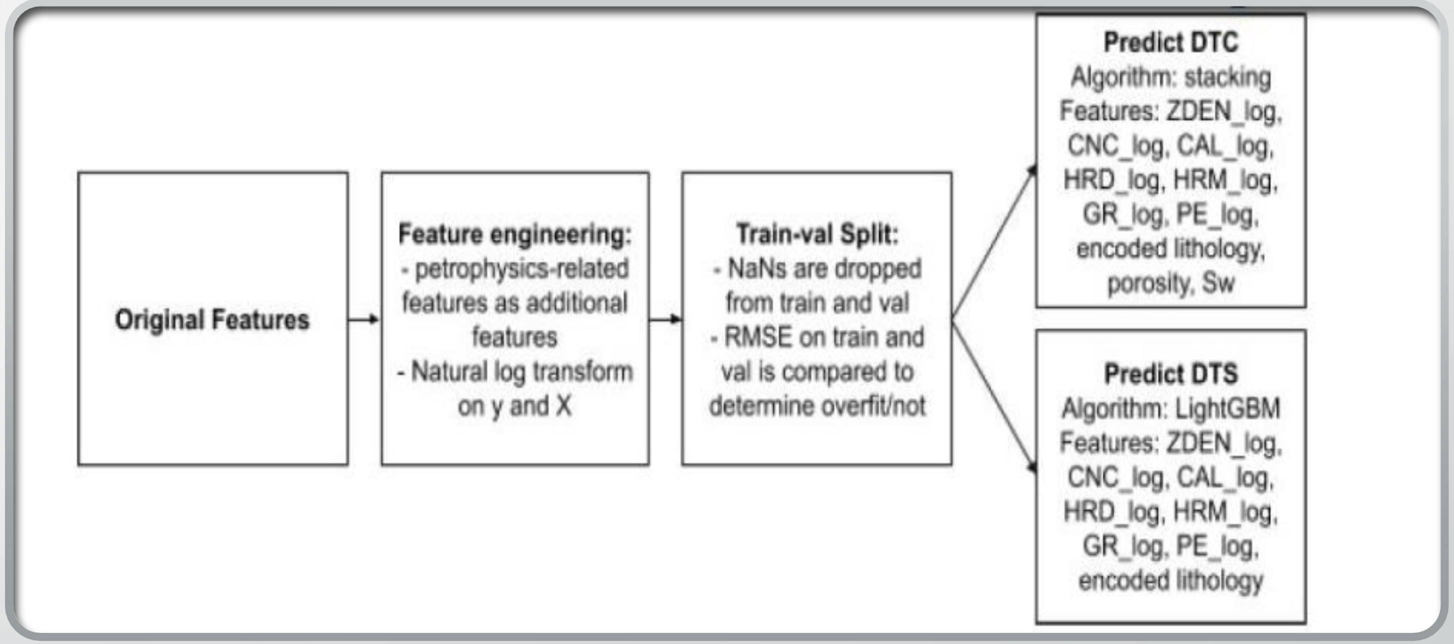
MODEL SUMMARY

- The data set acquired from well 1 was split in the ratio of 80:20. Hence, a training and validation data set was created.
- Using the training data set, we formed the cross plots among the nine well logs. After forming these relations, it was implemented on the validation data set.
- The machine learning technique which gave the lowest RMSE was selected.
- After successfully implementing the algorithm on validation dataset, it was used on the test dataset (without sonic logs) to predict the missing well logs.
- Following this process, I implemented this algorithm to generate the missing sonic logs from the data provided by ONGC.

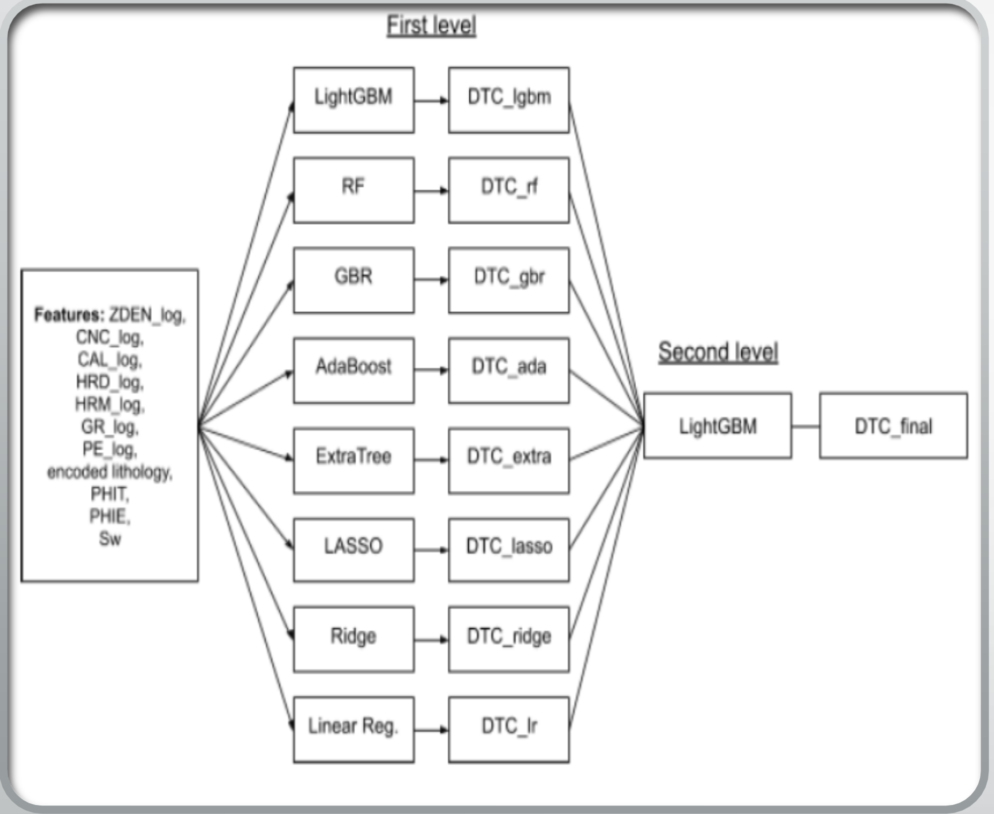

GR
CNC
ZDEN
HRD
PE

MODEL SUMMARY

- The algorithm was a combination of multiple machine-learning algorithms, also known as an ensemble method.
- The off the shelf algorithms had low prediction accuracy.
- Data cleaning, feature selection, and feature engineering were used as a complement to algorithms selection. The figure here illustrates the workflow.

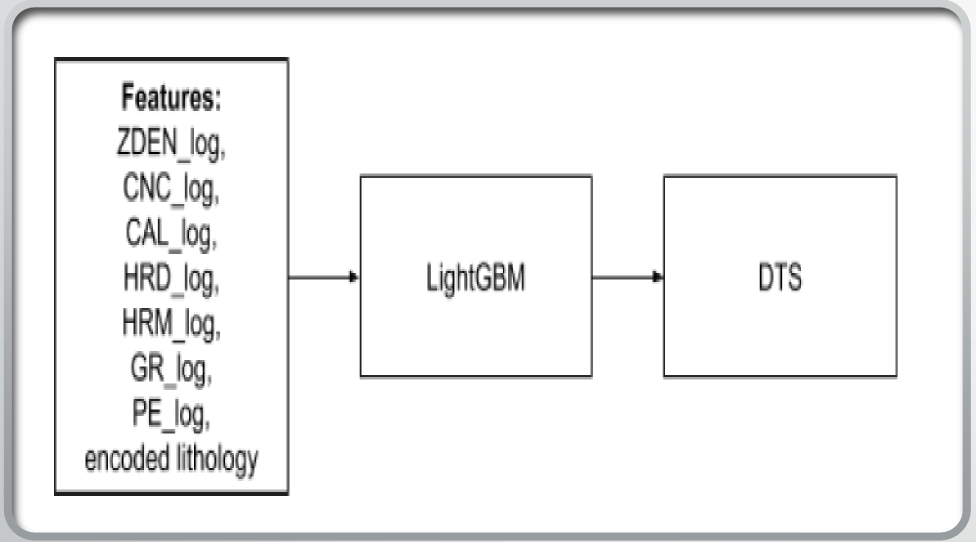


DTC PREDICTION



- To predict DTC, two steps of prediction were used shown in the figure.
- The first step is prediction using an ensemble of popular machine-learning algorithms, namely Random Forest, GBR, AdaBoost, ExtraTree, Lasso Regression, Ridge Regression, and Linear Regression.
- Results from these algorithms were used as features for the second-step prediction. The model that was used to wrap the ensemble was LightGBM.

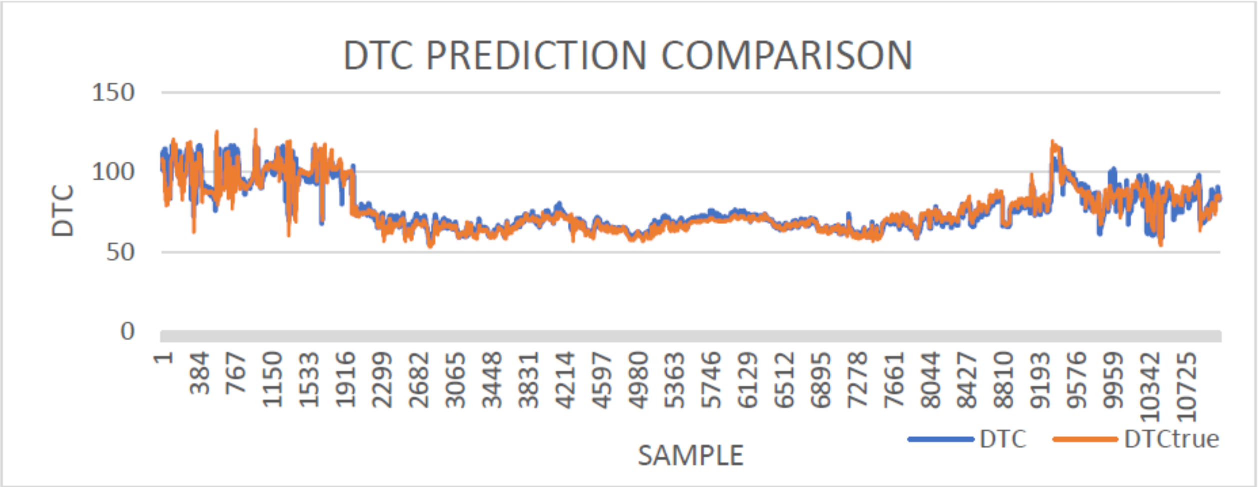
DTS PREDICTION

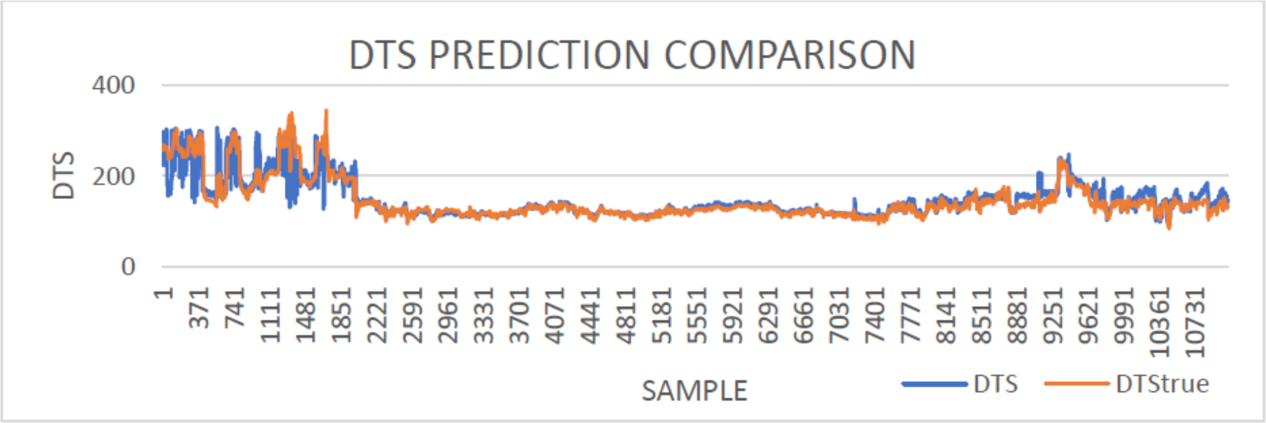


- The prediction for DTS was more straightforward, as shown in the figure.
- Ensemble method was not used as it yields worse results compared to a straightforward prediction by using LightGBM.
- the most reasonable way to make acceptable results in making predictions in this complex situation is by making a consensus from many predictions.

RESULTS

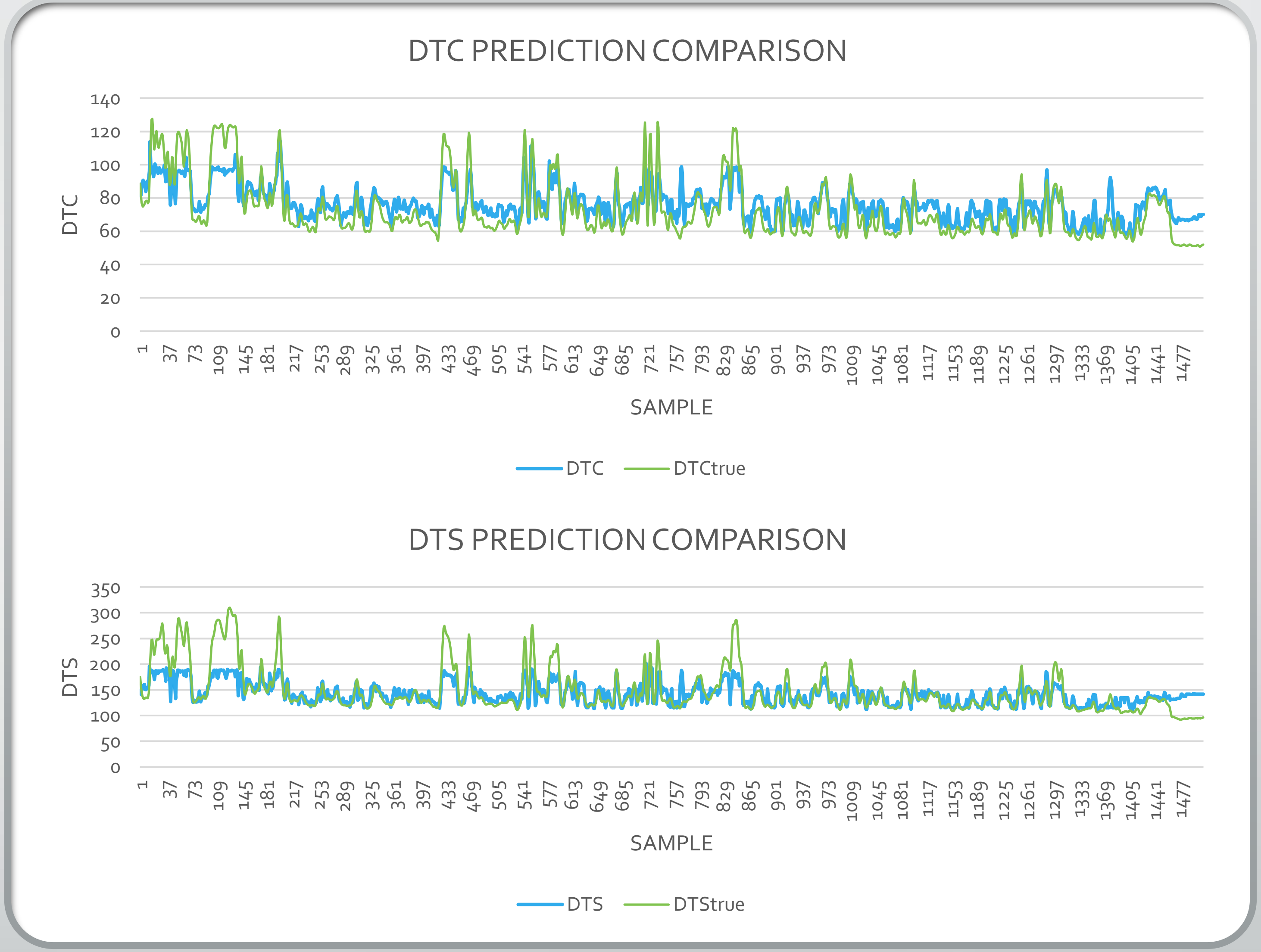
- Results obtained from reference are shown in the plots.
- To predict DTC, the best train and val performance using Extra Trees algorithm is as follows:-
- Using MLA: Extra trees
- RMSE Train: 0.003
- RMSE val: 2.682





RESULTS

- Results obtained from data provided by ONGC are shown in the plots.



CONCLUSION

- The results obtained from the data used in the SPWLA contest was analysed and it was found that the anomaly of the true and predicted DTC and DTS values matched properly with negligible offset.
- The results obtained from the open data was observed and it was found that the anomaly of the true and predicted DTC and DTS values matched to a good extent with some offset in amplitudes.
- The result from machine-learning prediction is quite good, and this competition can open a whole discussion on the practicality of using synthetic DTC and DTS in actual use cases, e.g., advanced petrophysics, geomechanics, etc., in the absence of sonic log data.
- It was also found that simple natural logarithm transformation can improve the prediction dramatically, as it creates a more “normal, algorithm-friendly” data distribution to numerical data.

REFERENCES

- SPWLA_Today_Newsletter_Vol3_No6.pdf
- https://www.spwla.org/SPWLA/Chapters_SIGs/SIGs/PDDA/PDDA.aspx
- Sonic Log Prediction (analytics-blog-d3223.web.app)
- Telecom-churn-case-study-1 (jayantb1019.github.io)
- ML Project Evaluation (notion.site)