KPIT

22nd October 2019

Navigating the data maze - tools & techniques for successful prognostics – Part II

Agenda

1. What's 'Not' Predictive Maintenance
2. Advanced Analytics – Predictive Diagnostics Quadrants
3. Evolution of Predictive Diagnostics
4. Case Study – Predicting RUL
5. Redefining the status quo – Predictive Diagnostics
6. Takeaways

What's 'Not' Predictive Maintenance

Dashboarding

No PM on new Automobiles

Expensive

It's Complex

Advanced Analytics – Predictive Diagnostics Quadrants

"Predict the future failure point of a machine"

Predict Remaining Useful Life (RUL)

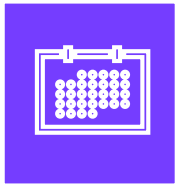
- **How much time left before the failure happens?**
- Solved using: Regression based Deep learning / Machine Learning Algorithms.

Predict failure within a time-window

- **Whether the component will fail in next 1 hr.?**
- Solved using: Classification based Deep learning and Machine Learning Algorithms.

Flagging Anomalous behavior

- **Is this behavior normal?**
- Solved using: Classification based Deep learning and Machine Learning Algorithms.



Failure Probability Over Time

- **Can my past learnings on failure be saved over time?**
- Solved using: Bayesian Networks and Markov Chain Models.

Evolution of Predictive Diagnostics

Threshold Based

- Single Value
- Correlation

Limitations:
- Emphasis on few parameters
- Limited to domain.
- Sensitive to noise

Learning Based Algorithms

- Pattern Recognition
- Machine Learning

Limitations:
- Limited to domain.
- Explicit Feature Engineering
- Limited to Linear relationships

Advanced Self Learning Algorithms

- **Automated Domain Knowhow**
- **Real-time Prognosis**
- Deep Learning

Case Study:
Predict Remaining Useful Life

Case Study I - Predicting Remaining Useful Life (RUL) of a spark plug

Customer:

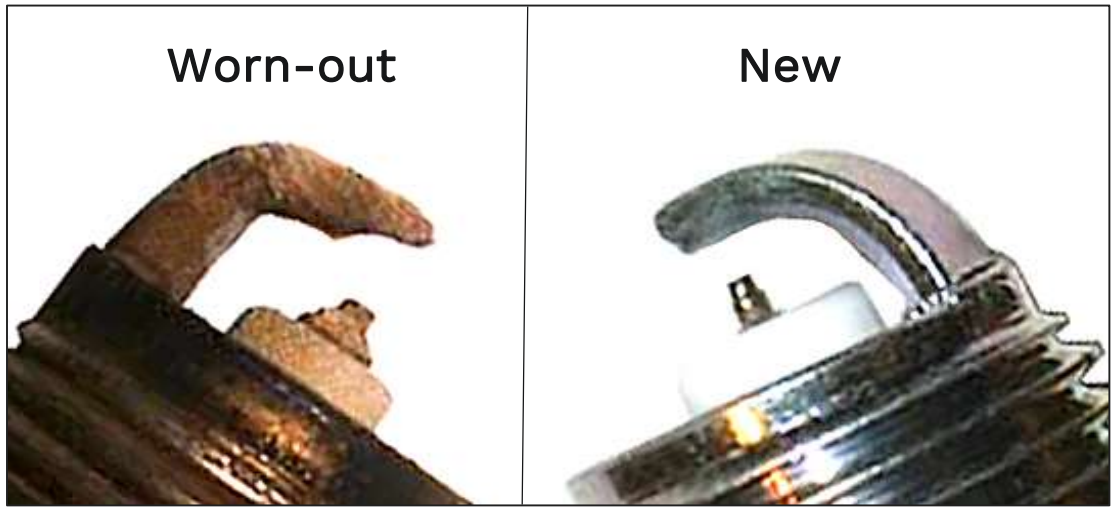
"Spark plugs are failing invariably; can you estimate when would they fail?"

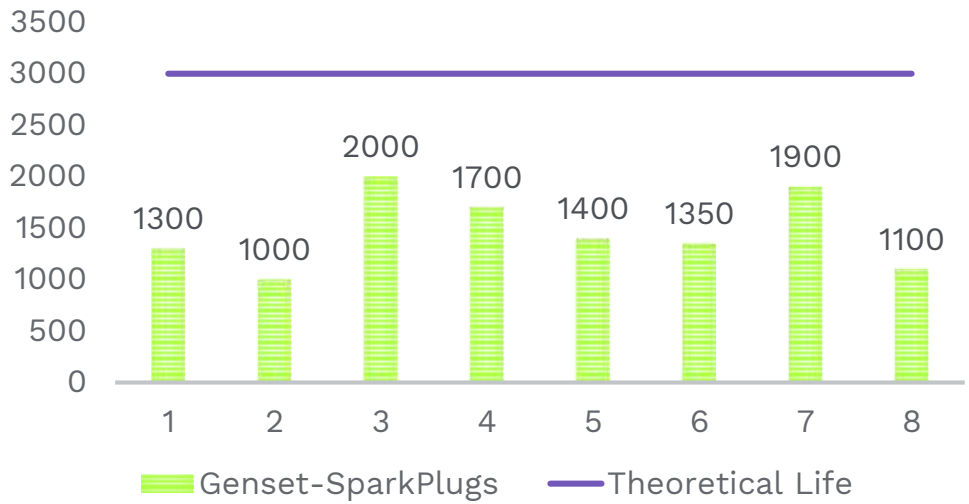
Data:

- 2 Years time-series data for 8 Gensets
- Data sampled at 2Hz; 120GB worth of Data
- Logs when the spark plugs were replaced.

Challenges:

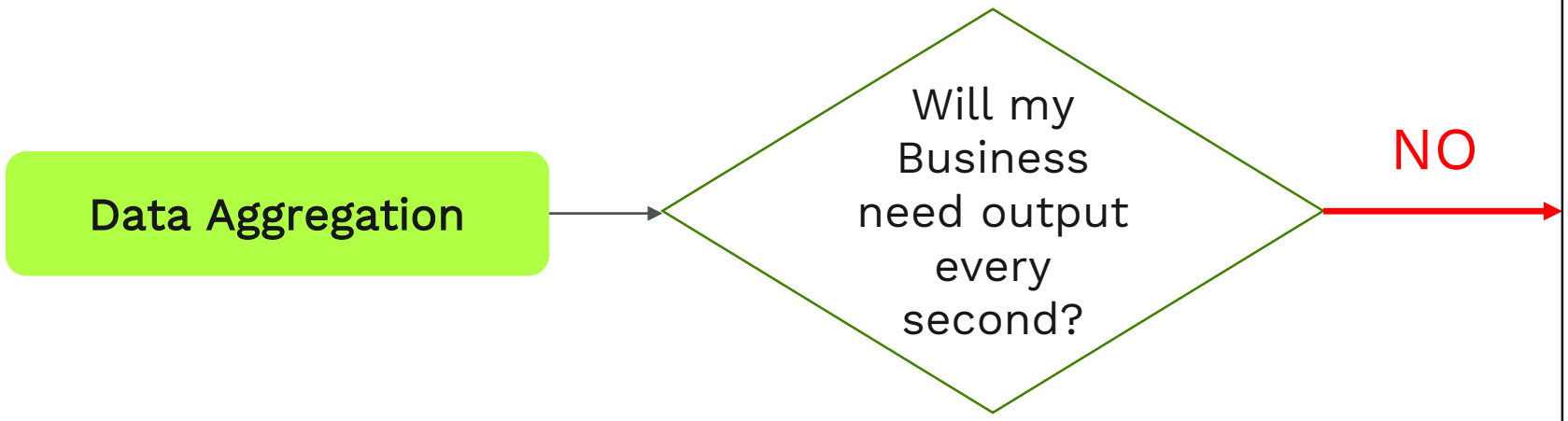
- Gensets with variable specifications.
- Different operating Conditions.
- Diverse make of spark-plugs across time.





Aggregating Data – Predict RUL

1



- Aggregate over sliding time window – min, max, mean, std-dev, median.
- Right Aggregation strategy – No loss in data granularity.
- **Reduces** the processing data by 85% i.e. to ~10GB
- Big-Data Processing tools – Apache PySpark

Data Exploration & Validation

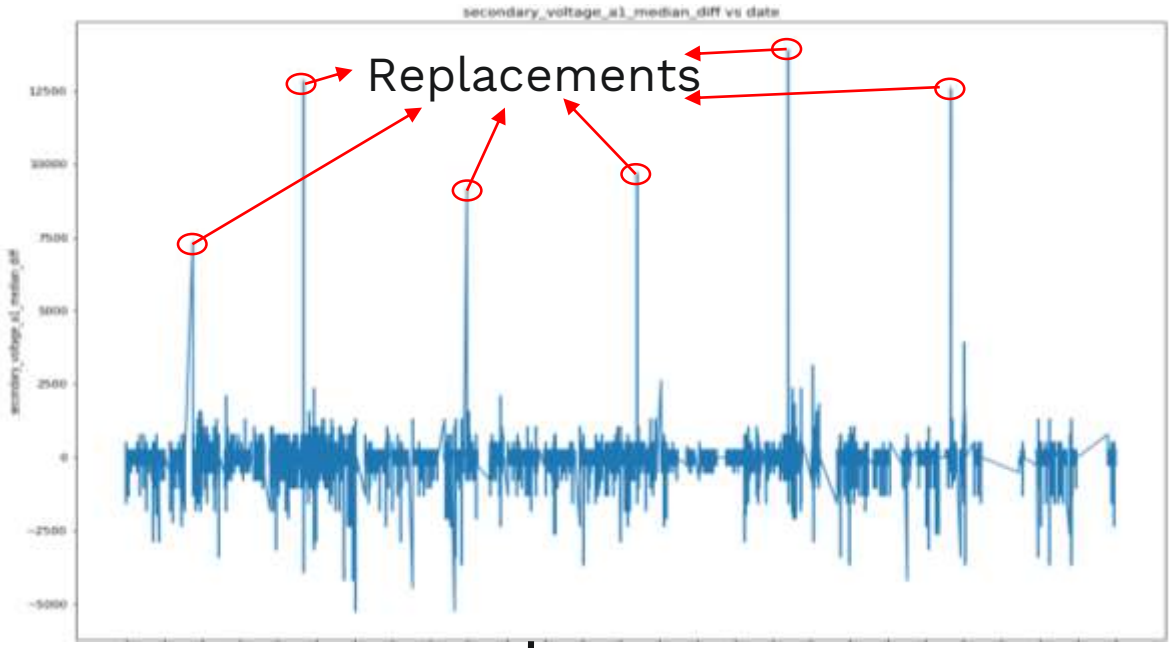
2

Data Exploration and Validation

Can the provided data solve the use case?

YES

NO



Domain Validation ?

NO

YES

Not the right problem to solve!

Modelling

3

- **RUL** linearly decreases with time – Assumption
- Predictor is **RUL**

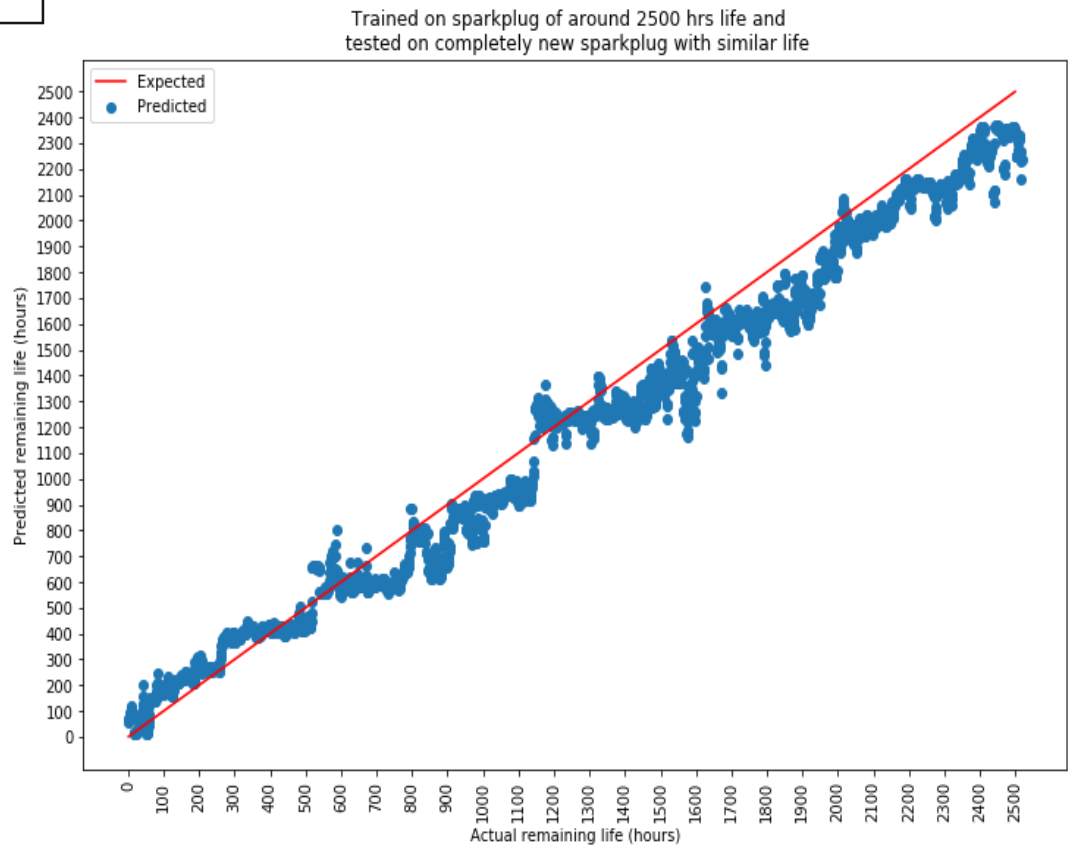
ML Ensemble Regression Models

Feature Snippet:

- Secondary voltage standard deviation
- Mixture temperature mean
- Mixture pressure A mean
- Mixture pressure B mean
- Electrical Power mean

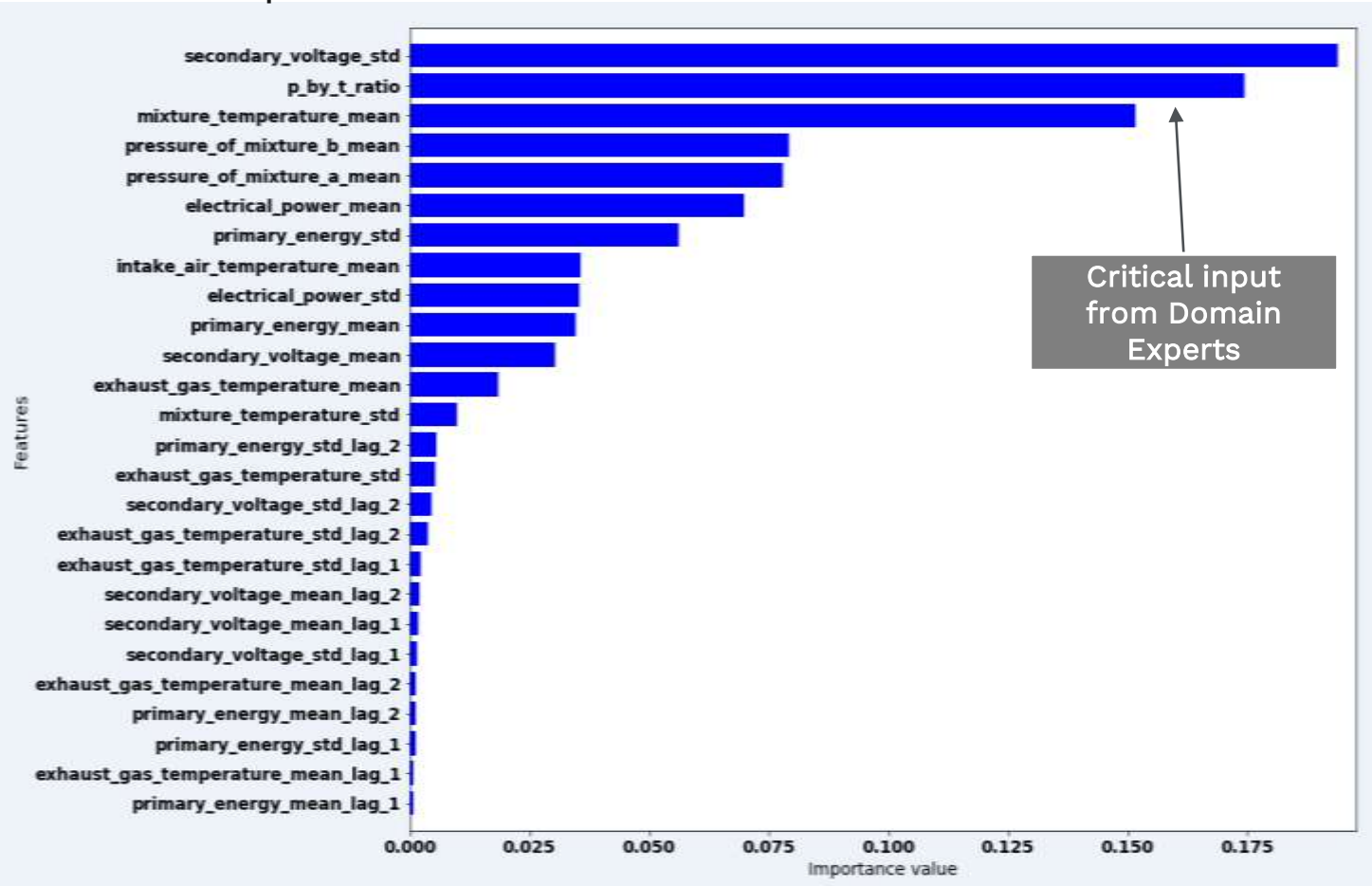
P / T ratio (Derived) → **Input by Domain experts**

and a few many more...

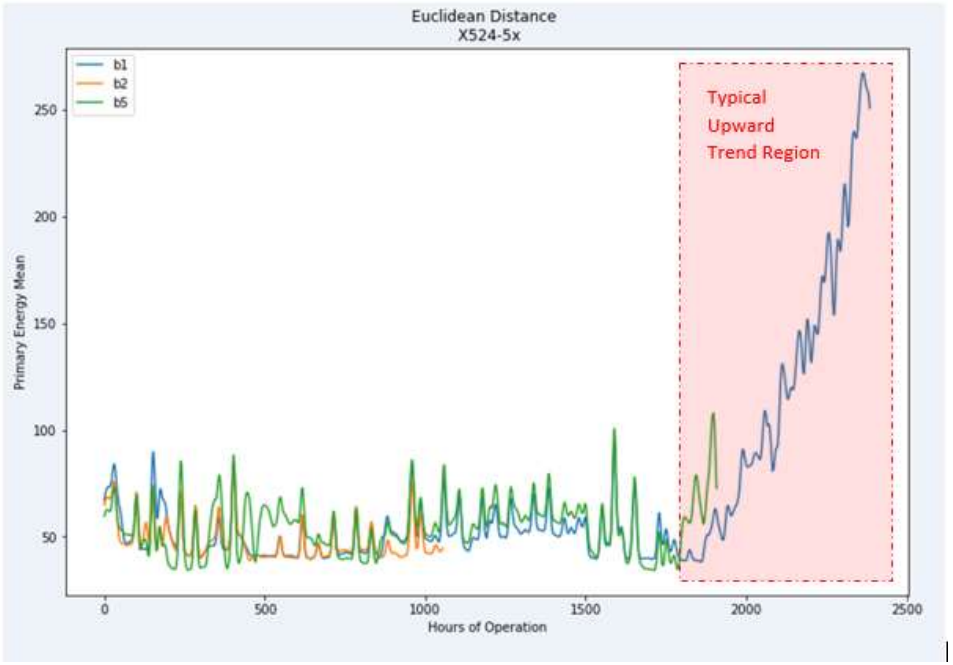


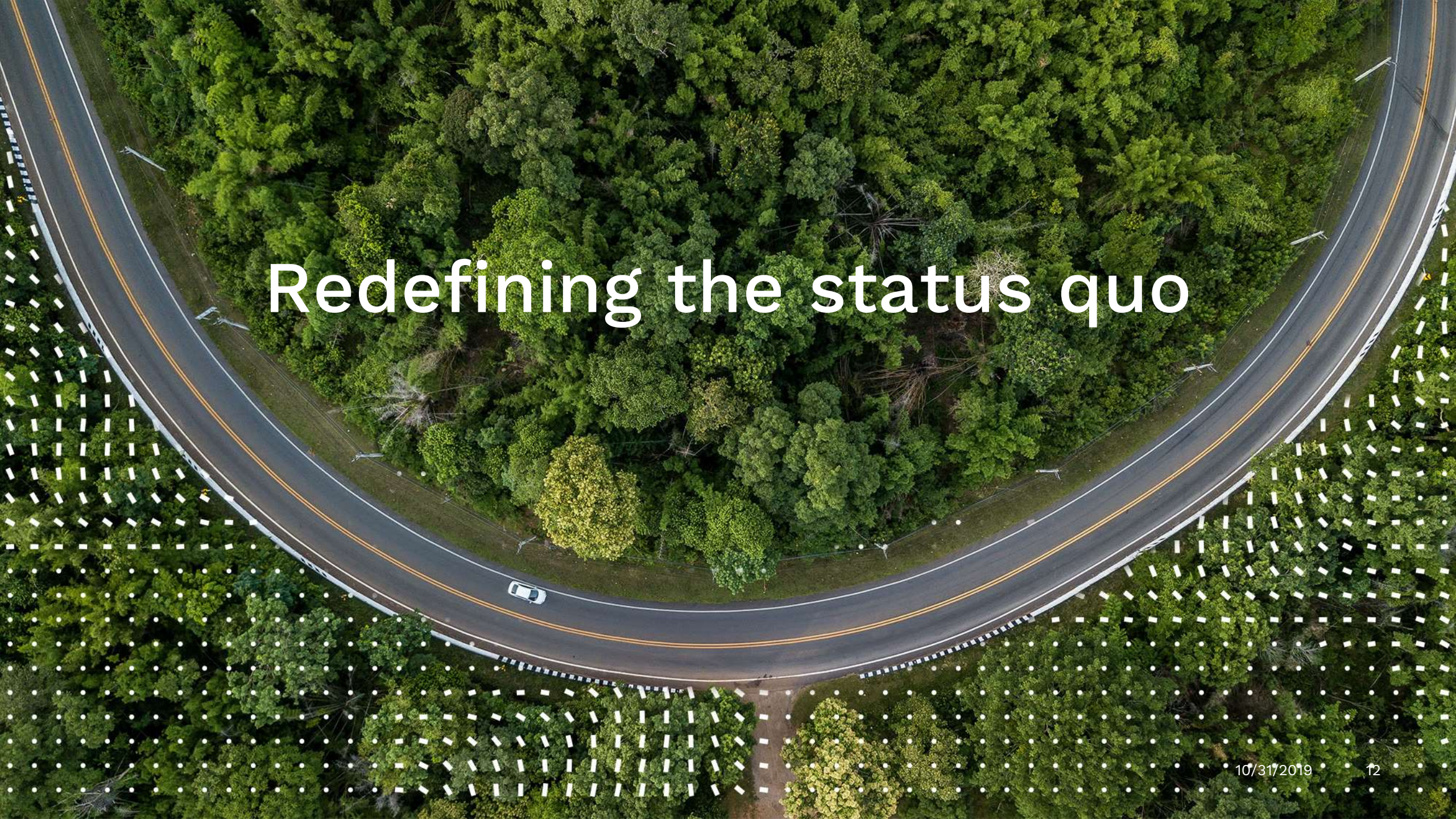
Evaluation

Feature Importance

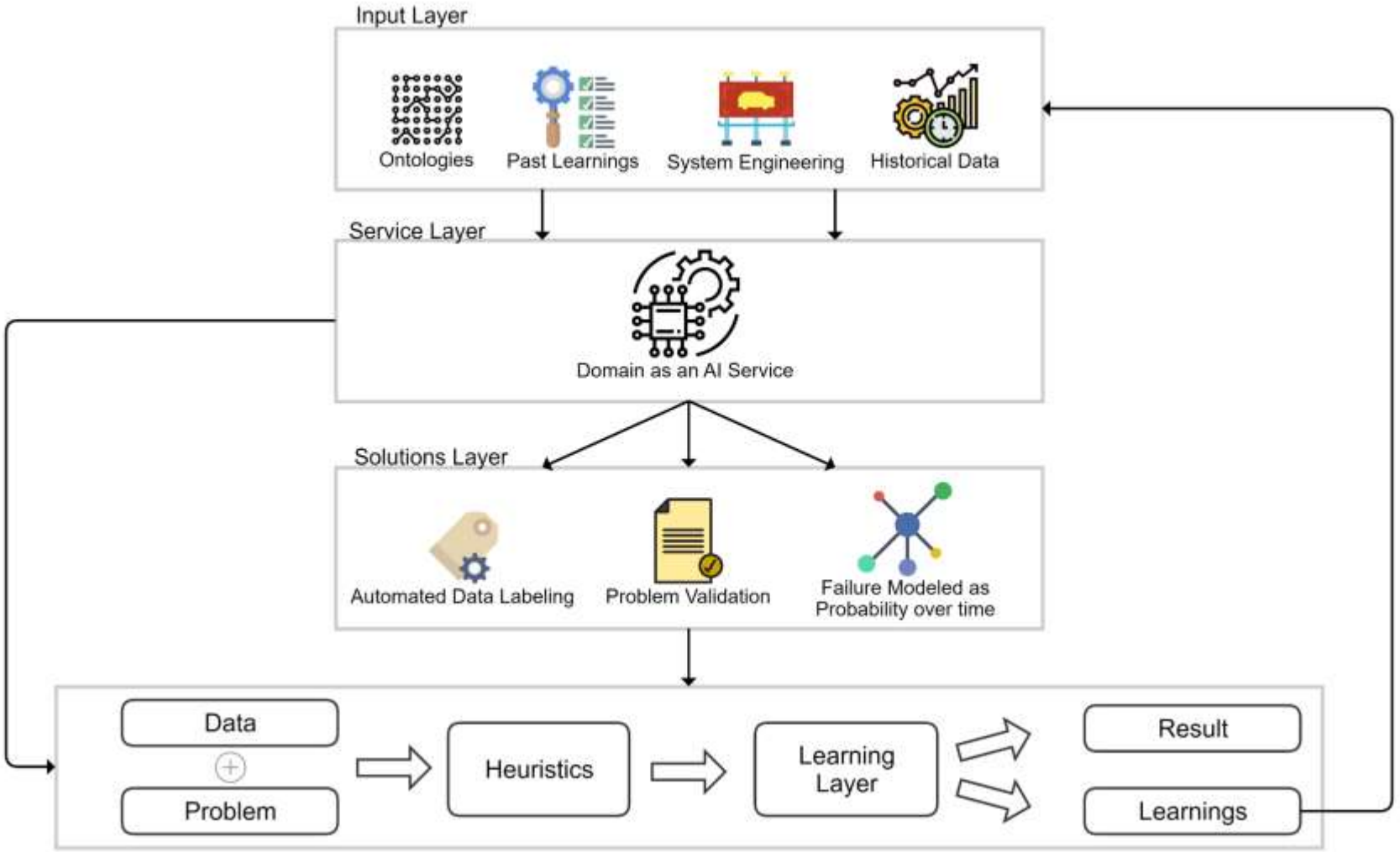


Validation




Redefining the status quo

Encapsulate the Domain



Advantages:

- Neural Networks are **not** a black box anymore!
- **50% Time Reduction** in Problem Solving
- Learnings are **back propagated**!

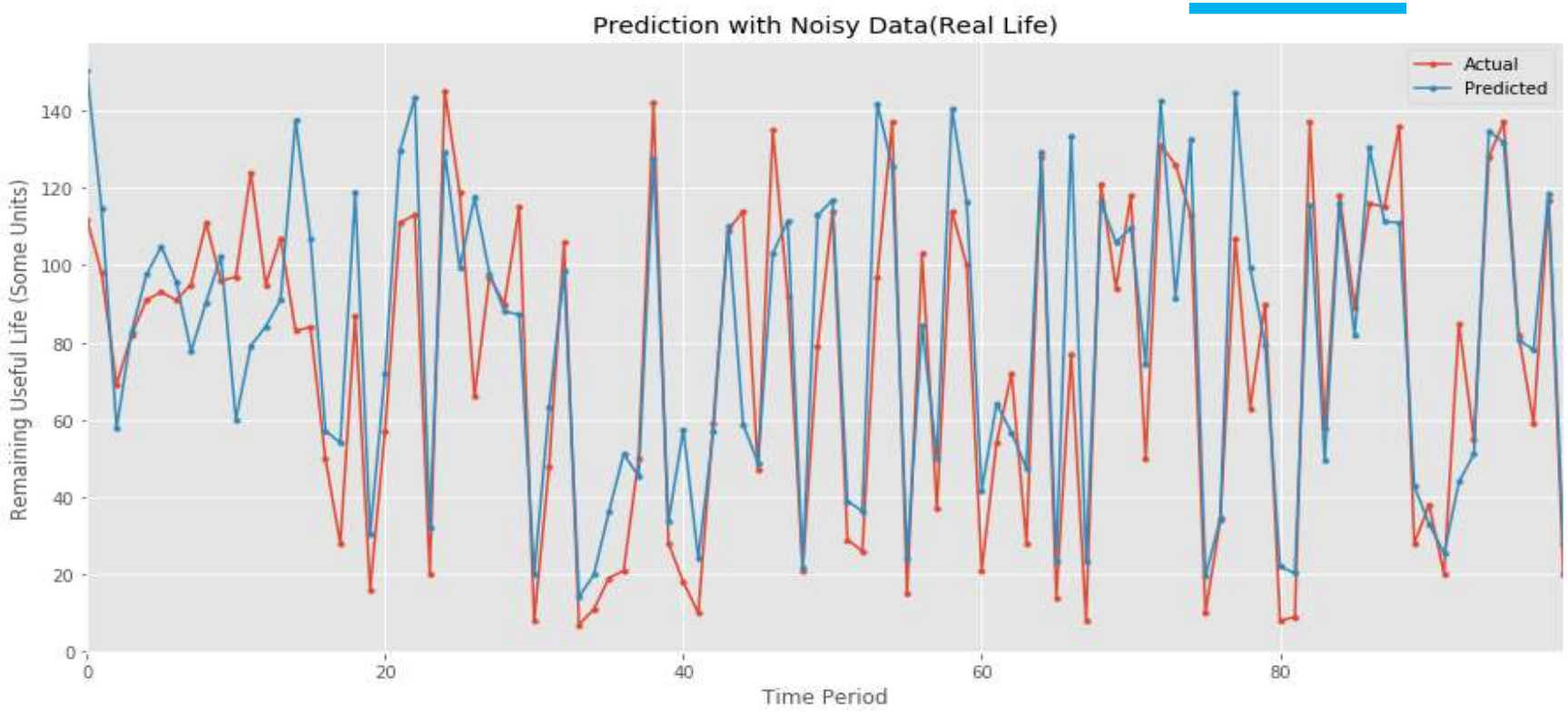
An Ongoing Case Study – Towards Perfection…

- **Recurrent Neural Network** with perfected Architecture
- **No explicit domain feature engineering –** Fully Encapsulated
- Ability to predict **fluctuating values.**

Initial Results

Actual

Predicted



Takeaways

- ✓ Evolution of Diagnostics
- ✓ Case study positively impacting a Business
- ✓ Encapsulating the Domain – The Future