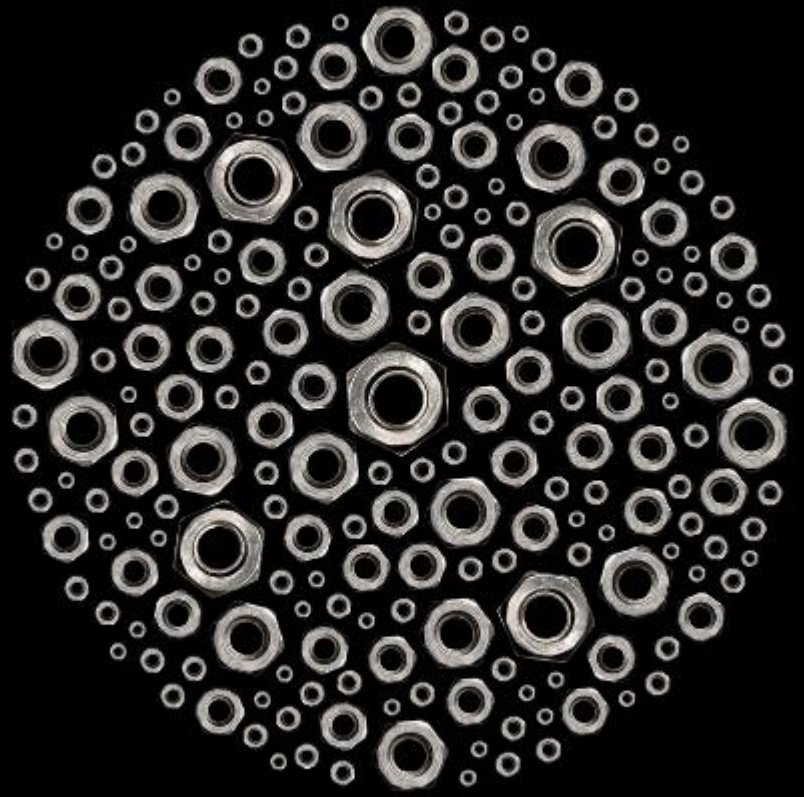
Deloitte.

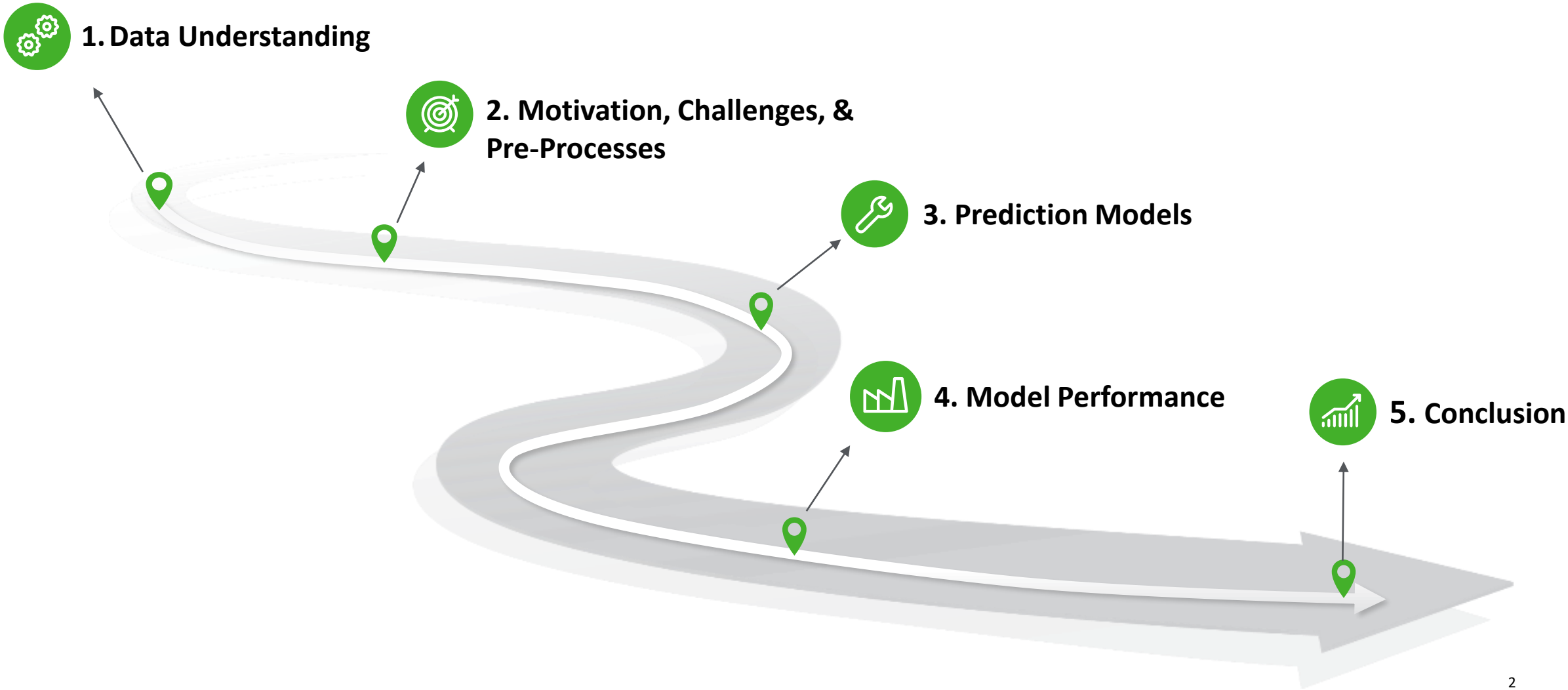


Optimal Machine Failure Prediction Model For Toyota

Tolulope Akinlabi, Brandon Kokin, Alexandria Lee-Robinson,
Diana Seo, Hanwen Zuo

MAKING AN IMPACT THAT MATTERS since 1845

Agenda


1. Data Understanding
2. Motivation, Challenges, & Pre-Processes
3. Prediction Models
4. Model Performance
5. Conclusion

Data Understanding

Significance of machine failure in vehicles

Situation

- Machine parts failure resulting in catastrophic losses (Toyota)
- 5 failure types and machine aspects

Complication

- Hard to identify tell-tale signs
  - If able to, often too late
  - Financial pressures

Solution

- Predictive modelling with high accuracy
  - Random Forest

Motivations, Challenges, and Pre-processes

Importance of pre-processing techniques

Challenge

Creating multiple reliable models consuming an **imbalanced dataset** with both **discrete** and **continuous variables**.

Motivation

Working with the same preprocessed dataset would enable **comparability** among the models created

Pre-Processing Techniques

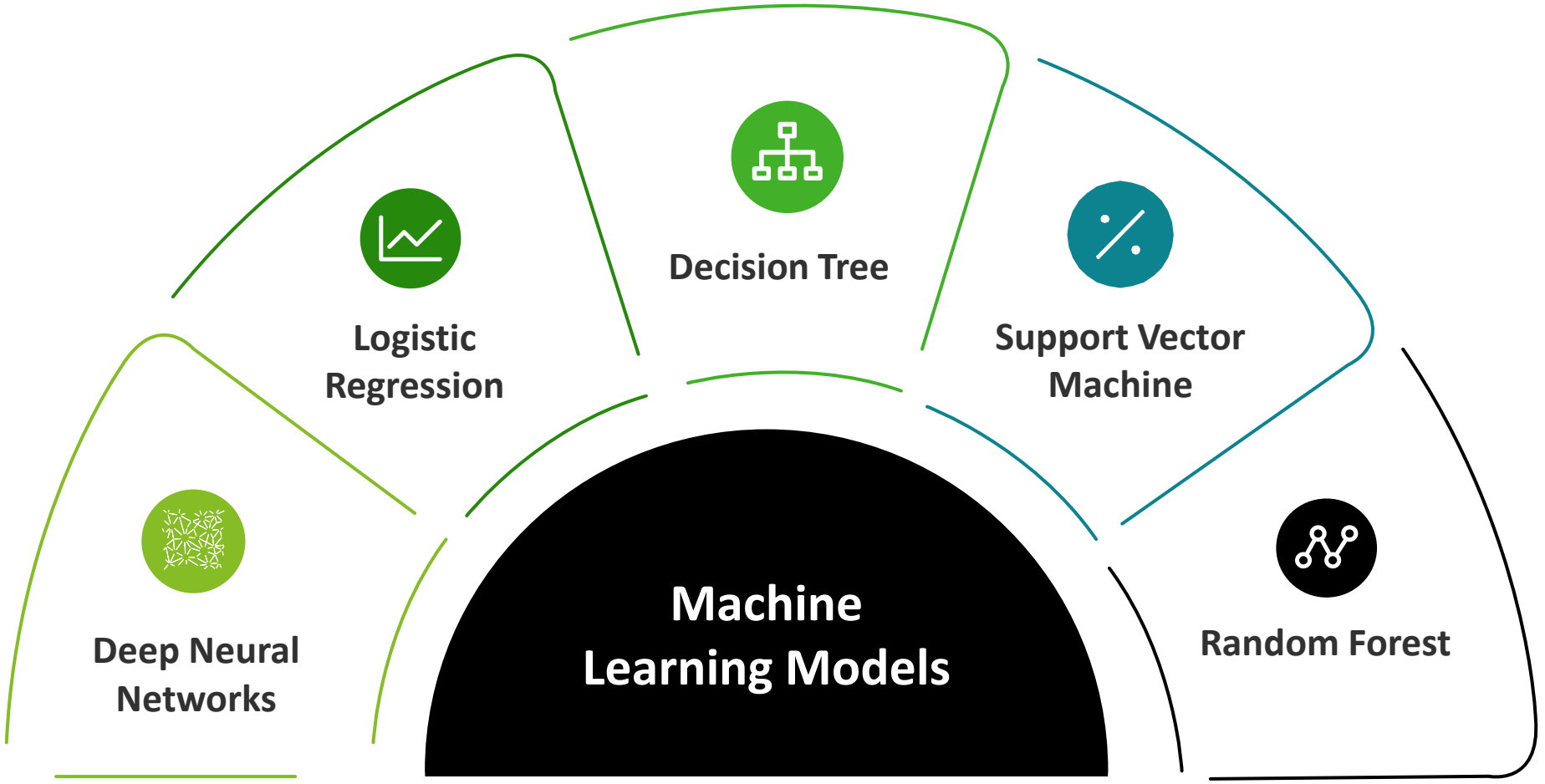
- **SMOTE** resampling
- Pandas **Get Dummies**
- **Consolidation** of failures

Results

- Dataset with **balanced failures** and non-failures
- Columns with multiple categories transformed into **binary data**
- Multiple columns to indicate failure **transformed into 1 column**

Prediction Models

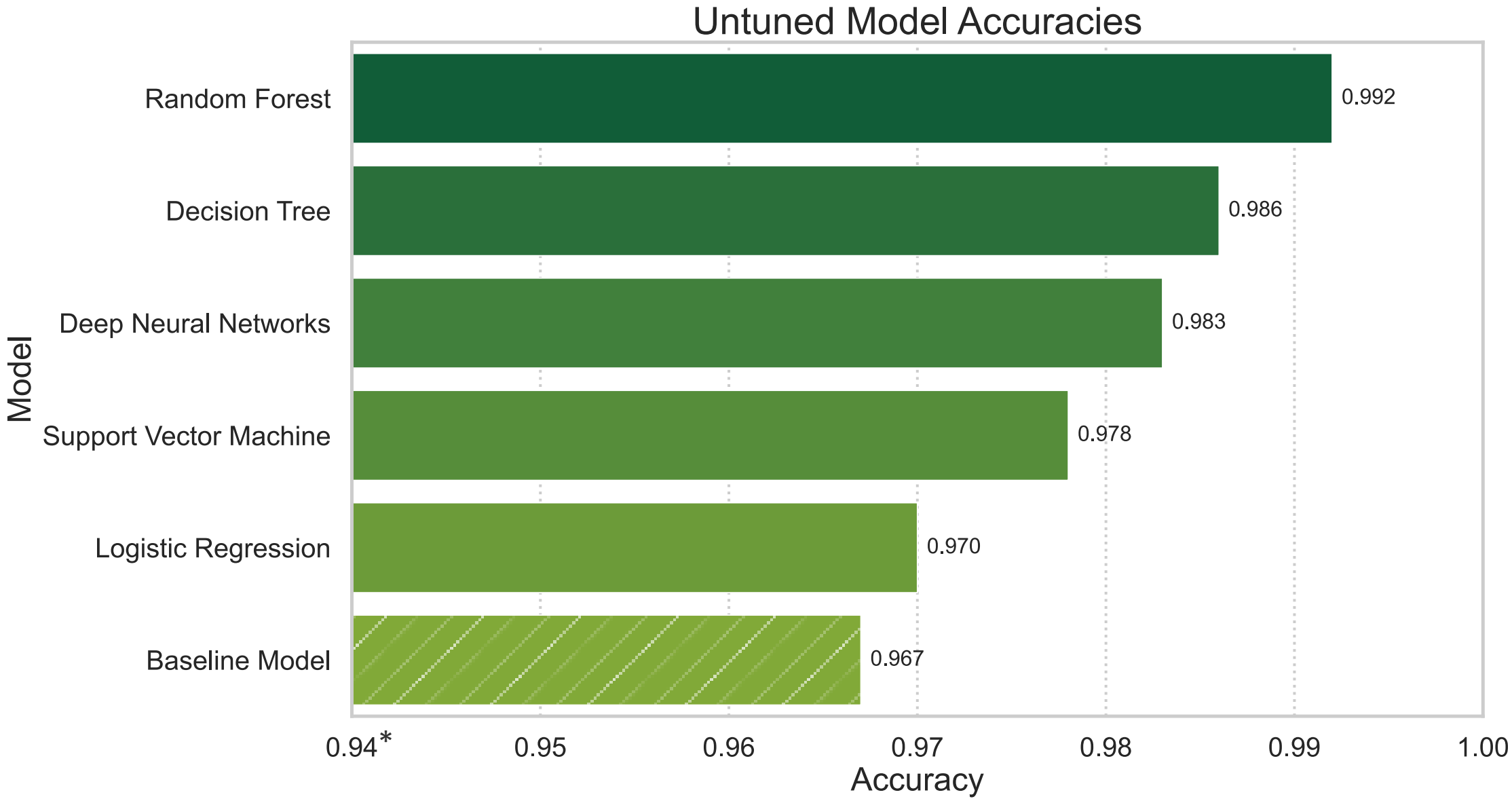
Selecting ML models for predicting machine failure



Model Performance

Holdout set accuracy before tuning

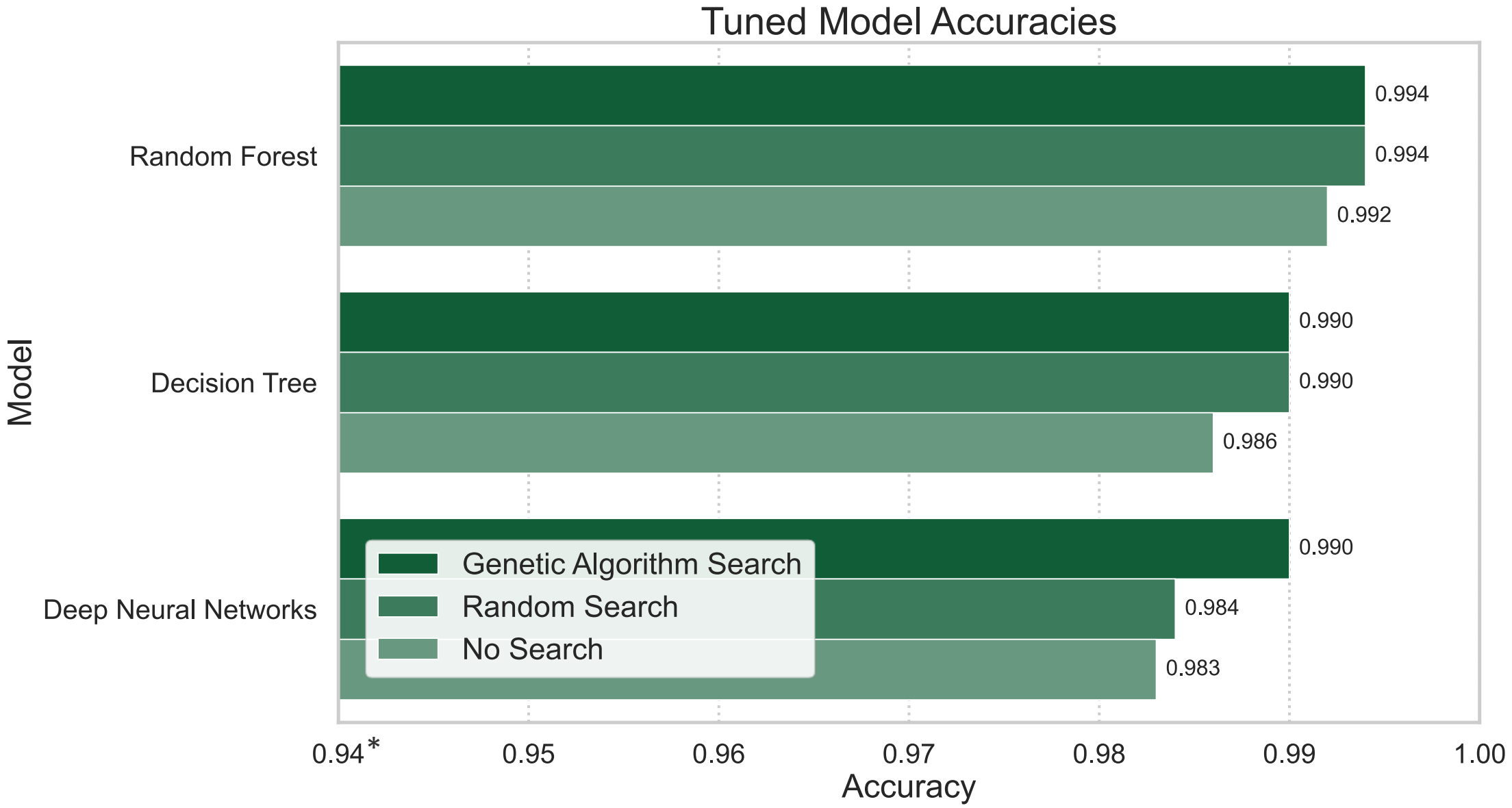
* Starts at 0.94; ticks represent 0.01 change



Model Performance

Holdout set accuracy after tuning

* Starts at 0.94; ticks represent 0.01 change



Model Performance

Confusion matrix

Random Forest		Decision Tree		Deep Neural Networks	
Precision:	0.97	Precision:	0.83	Precision:	0.88
Recall:	0.86	Recall:	0.86	Recall:	0.80

Recommendation

Random Forest Model is the most optimal choice for Toyota in predicting machine failures

- Increased downtime
- Higher reactive maintenance costs
- Reduced operational performance
- Decrease in product quality and brand reputation

Current Status

- Best accuracy
- F1 score
- Solid precision and recall

Random Forest Model

- Decrease downtime
- Enhanced operational performance
- Reduce maintenance cost
- Increase product quality
- Protect brand reputation

Result

Deloitte.

[Tolulope Akinlabi]

[Brandon Kokin]

[Alexandria Lee-Robinson]

[Diana Seo]

[Hanwen Zuo]

Thank You For Considering Our Recommendation

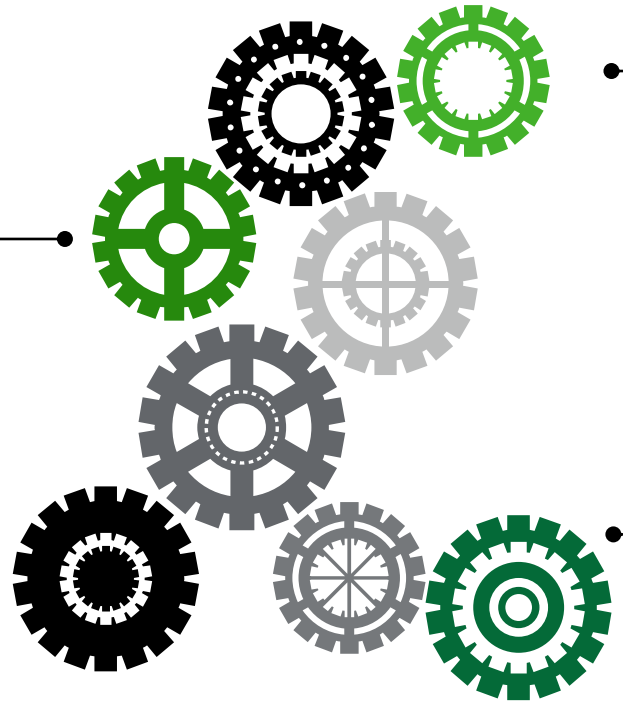
Any Additional Questions?

Model Performance

Functions and corresponding evaluations for accuracy

Deep Neural Networks

- Optimize a loss function across layers of nodes
- Accuracy = 0.983
- F1-score = 0.754



Decision Tree

- Split data to minimize entropy
- Accuracy = 0.986
- F1-score = 0.803

Random Forest

- Combine votes from decision trees
- Accuracy = 0.995
- F1-score = 0.912