Predictive Machine Maintenance Project (systemRx)
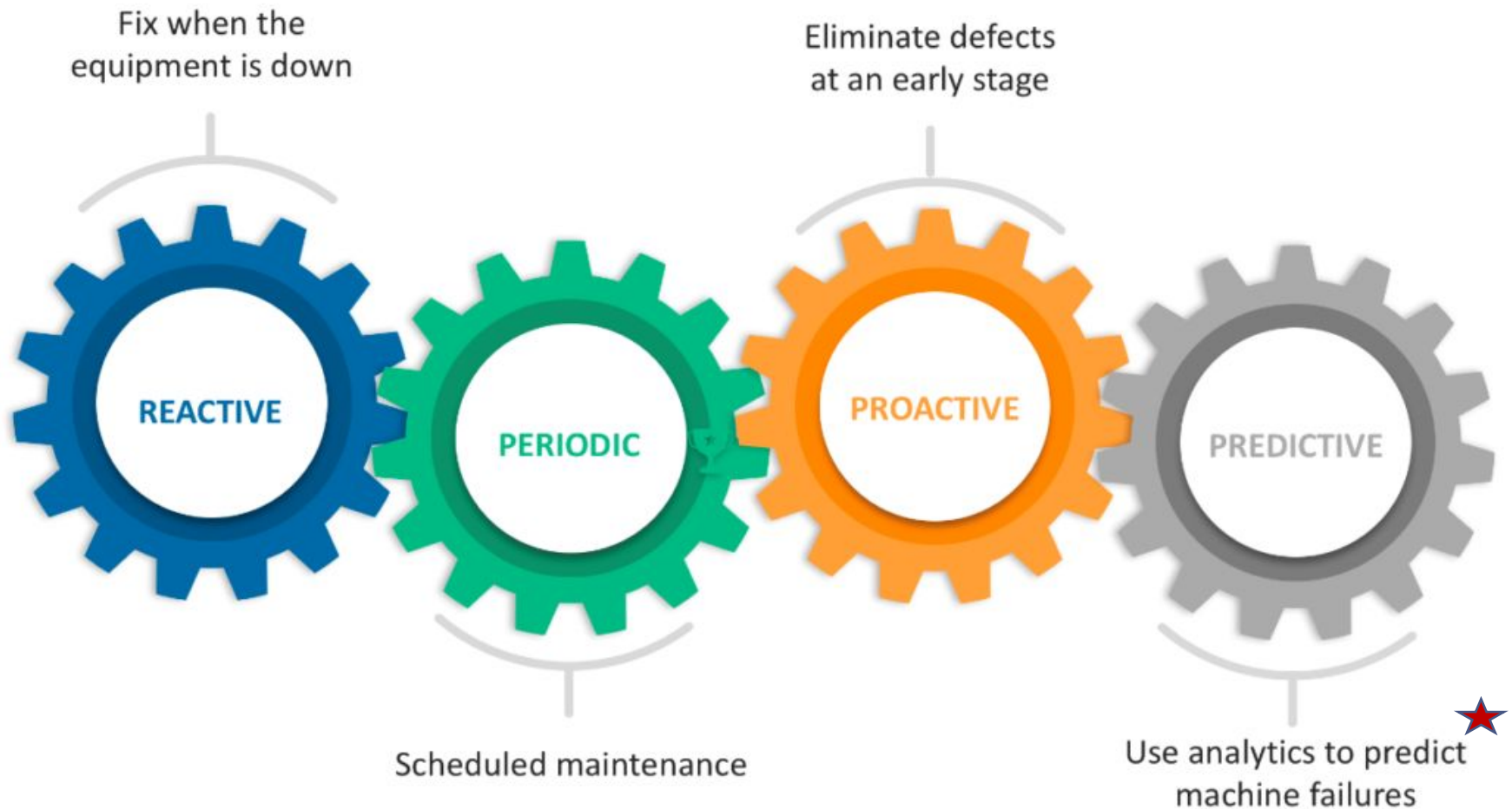
Fix when the equipment is down

Eliminate defects at an early stage

REACTIVE

PERIODIC

PROACTIVE

PREDICTIVE

Scheduled maintenance

Use analytics to predict machine failures
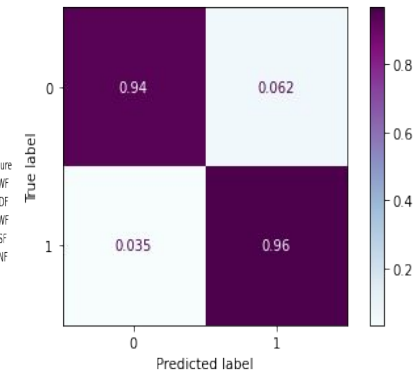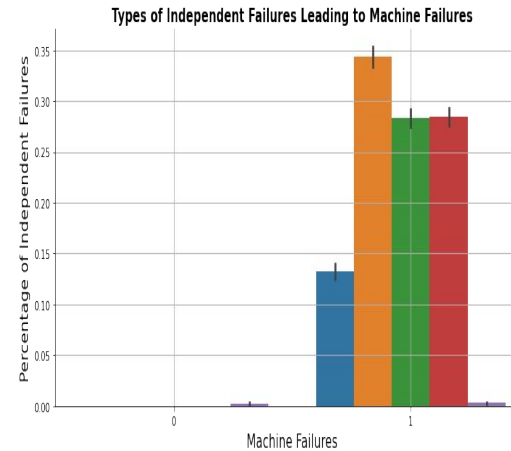
Image Source: digitaldoughnut.com

Presenter: Sheneka Allen
June 2022

Predictive Machine Maintenance Project (systemRx)

Objective

- Develop a modeling tool that accurately predicts the percent likelihood of machine failure. If possible, identify common contributors to machine failure to minimize equipment downtime and maintenance costs.
- Prediction performance goals: >90% correct positive predictions, < 5% for false negative predictions
- Technology & tools: predictive maintenance (pDm) dataset (synthetic) of 10K rows, data analytics

Visualizations

Types of Independent Failures Leading to Machine Failures

Type failure: TWF, HDF, PWF, OSF, RNF

Confusion matrix:
	0	1
0	0.94	0.062
1	0.035	0.96

True label / Predicted label

Model 'Problem-Solving' Value

Strengths:

- Helps to easily identify machine failure trends and patterns
- Potential decrease in maintenance & labor costs
- Potential increase in average time between machine failures

Limitations:

- Limited public availability for predictive maintenance data to re-train model
- No date-time values in dataset to calculate machine average runtime or availability
- Must collect a good mix of machine data to ensure quality model performance (e.g., balanced dataset improved performance by **31%!!**)

Recommendations

1) Focus initial maintenance tasks on <u>highest percentage causes</u> for machine failures
2) Migrate from scheduled maintenance culture to formal condition monitoring program

- Identify & collect machine data on <u>critical assets</u> (e.g., safety and revenue loss focus)
- Determine <u>what typically goes wrong</u> on machines that cause failures (e.g., 34% heat dissipation (HDF), 28% power failures (PWF) for this dataset)
- Consider altering testing frequency between degraded machine operation and breakdown

Presenter: Sheneka Allen
June 2022

Backup - More systemRx Visualizations

Classification Metrics for Confusion Matrix (Slide 2)

```
Classification Reporting for Test Set (XGB):
              precision    recall   f1-score   support

          0        0.96      0.94       0.95      2405
          1        0.94      0.96       0.95      2426

   accuracy                            0.95      4831
  macro avg        0.95      0.95       0.95      4831
weighted avg       0.95      0.95       0.95      4831
```

Heat Dissipation Failure - Feature Analysis

Rotational Speed and Process Temp Effect on Heat Dissipation Failure (HDF)
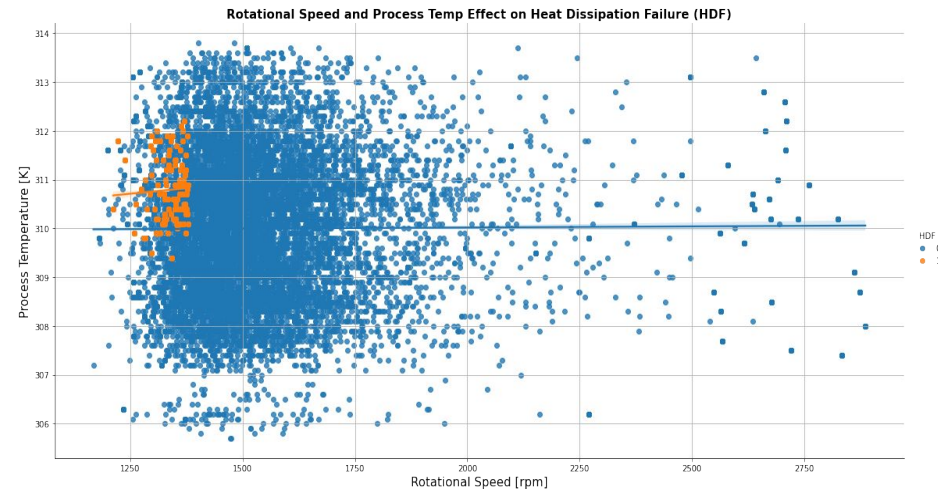
Imbalanced Data & Model Performance

Imbalanced datasets cause model predictions to be _skewed_ in favor of the majority class versus the minority class (target). The prediction accuracy is 99% but of ZERO analysis value! You MUST correct the data imbalance and re-train the model to improve performance quality and ensure its value to stakeholders.
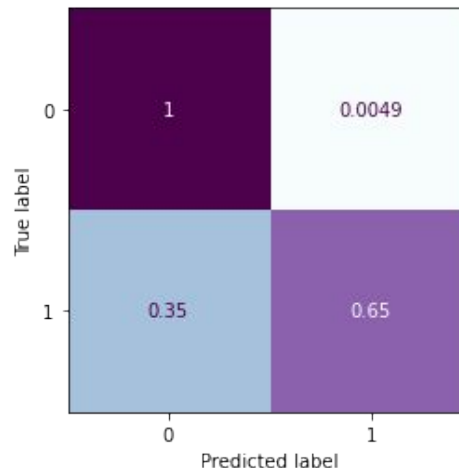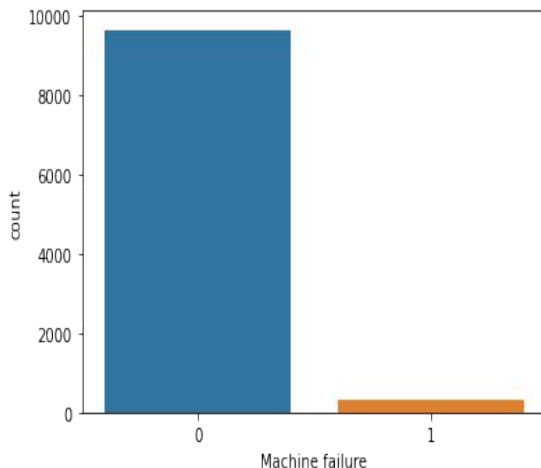

```
Classification Reporting for Test Set (XGB):
              precision    recall   f1-score   support

          0        0.99      1.00       0.99      2428
          1        0.80      0.65       0.72        72

   accuracy                            0.99      2500
  macro avg        0.89      0.82       0.85      2500
weighted avg       0.98      0.99       0.98      2500
```

Presenter: Sheneka Allen
June 2022