Predictive Maintenance for Facilities Failure Prediction using Machine Learning

International Research Experience for Students (IRES) 2022

Jessica Hankele Torres, Lorenzo Dino Mendoza, Sidney Brower

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

North Carolina
Agricultural and Technical
State University

UNIVERSITI
TEKNOLOGI
PETRONAS

ENGINEERING
TEXAS A&M UNIVERSITY

NSF

Our Extended Team

Faculty Advisors:

- Dr M Hilmi Hasan (Primary Advisor, UTP)
- Prof. Dr Bimal P Nepal (Co- Advisor, TAMU)
- AP Dr Jafreezal Jaafar (Co-Advisor, UTP)
- Hajar M Razip (Co-Advisor, UTP)
- Madhusudhan Gopalaraju (Graduate Assistant, TAMU

UTP Students:

- Farzana Athirah Abd Razak
- Nursyafinas Norazli
- Nur Anissa Zolkifli

Problem Background

- Probabilistic forecast from time to event are crucial in many industries including healthcare, manufacturing, meteorology, and the energy sector.
- If calibrated properly a forecast's usefulness increases with its confidence.
- Regression models tend to have much higher error than that of classification models
- Multiple linear regression models are growing in popularity and accuracy

Motivation for Research

- Predictive maintenance informs technicians about the type of failure that is likely to occur and when it will occur.
- It also enables pre-emptive investigations, maintenance scheduling, and repairs to be performed before the asset fails.
- Decreases downtime, which in turn creates more opportunities for profit.

Background Research

Predictive maintenance has been widely covered in extant literature

- Li et al. (2016) developed a Bayesian network model that investigated the probability of failure of a submarine pipeline.
- Ziolkowski et al. (2019) used artificial neural network to determine what affects the failure occurrence rate of concrete machine foundations.
- IBM et al. (2020) educational run down on what machine learning is, and how it is applied at IBM.
- TWI LTD. (2022) explains what predictive maintenance is, while describing the advantages and disadvantages.

Problem Statement

- The aim is to develop a machine learning-based prediction model which can foresee any facilities failure in industry such as in refinery or plant.

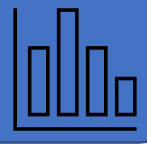
Research Methodology

Data Collection

Dataset was provided to UTP and CeRDaS by industry and represents the buildup to a failure event.

Data Analysis

- Data preprocessing
- Mutation
- Standardization
- Correlation Analysis

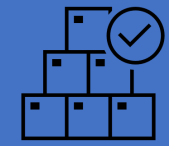


Models Development

- Date-Time indexing
- Creating countdown
- Arguments and parameters

Model Check

Is model reliable and able to be used in industry?



Data Analysis Platform: Python

Pandas

- Data manipulation and analysis
- Data structures and operations
- Manipulating numerical tables and time series

scikit-learn

- Machine learning library
- Statistical modeling
- Classification, regression, and clustering algorithms

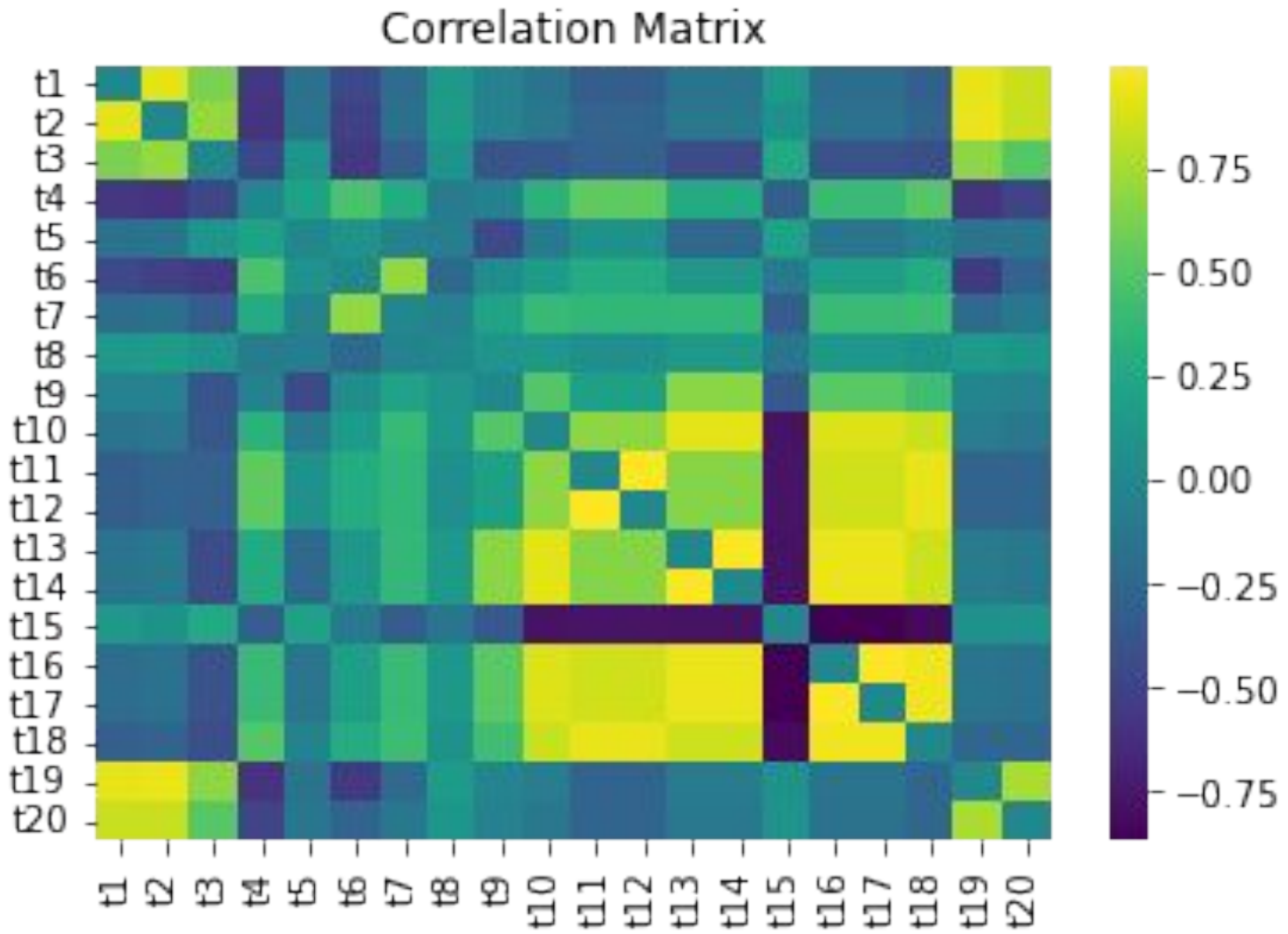
Data Analysis and Models Development

- Data Pre-processing
- Correlation Analysis

Multiple Machine Learning Models were Developed and Analyzed

- Regression Models
- Decision Tree
- Random Forest

Data Preprocessing: Correlation Analysis

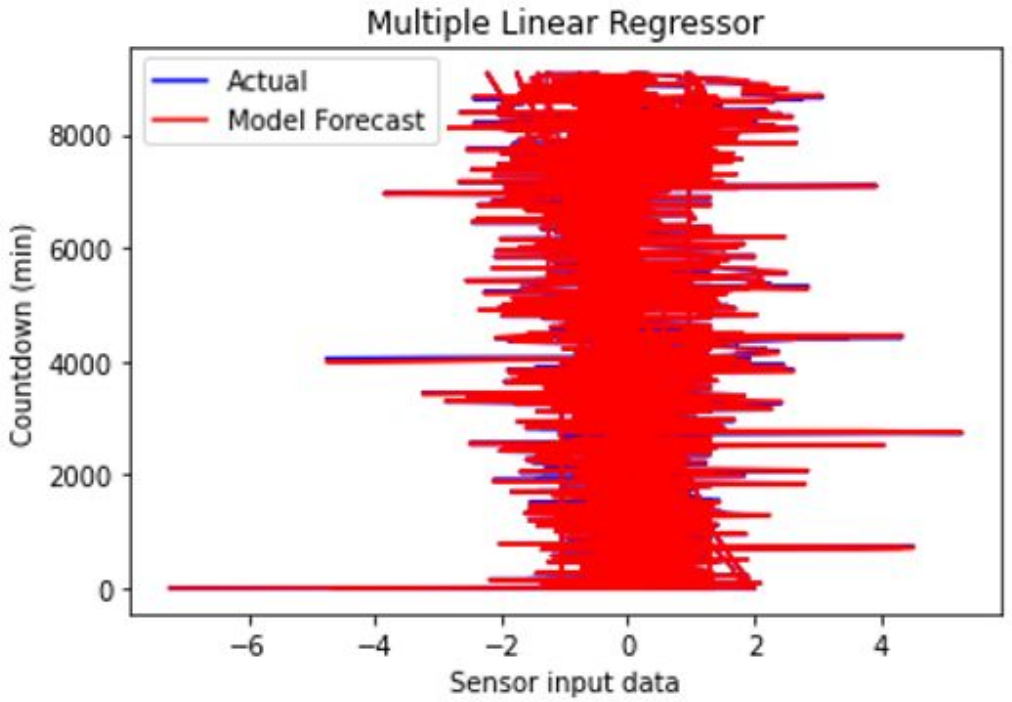


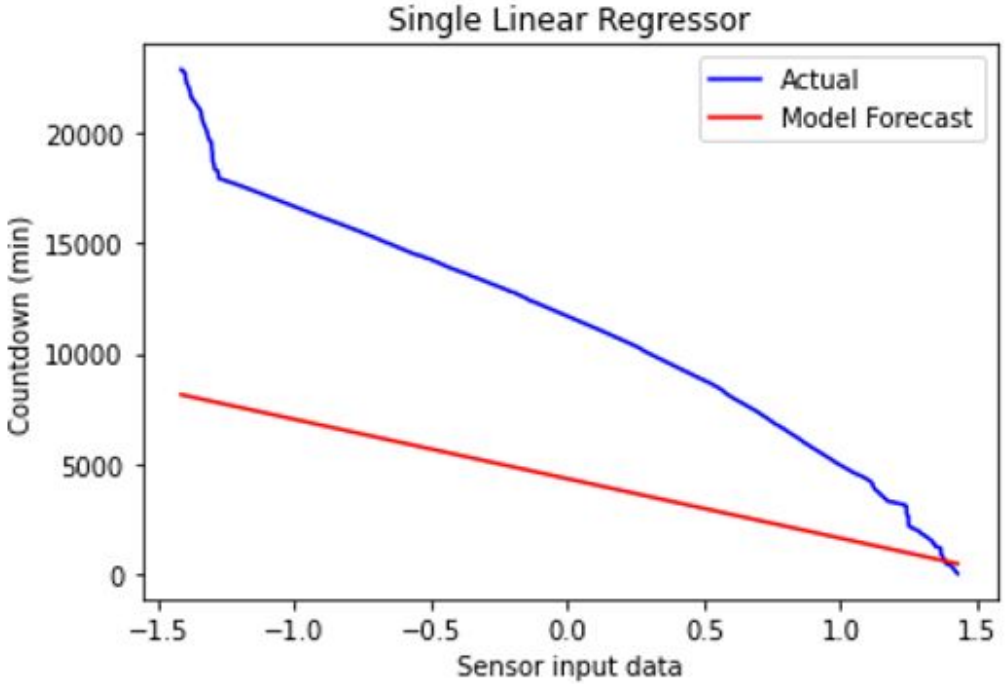
Models Development

Multiple Machine Learning Models were Developed and Analyzed

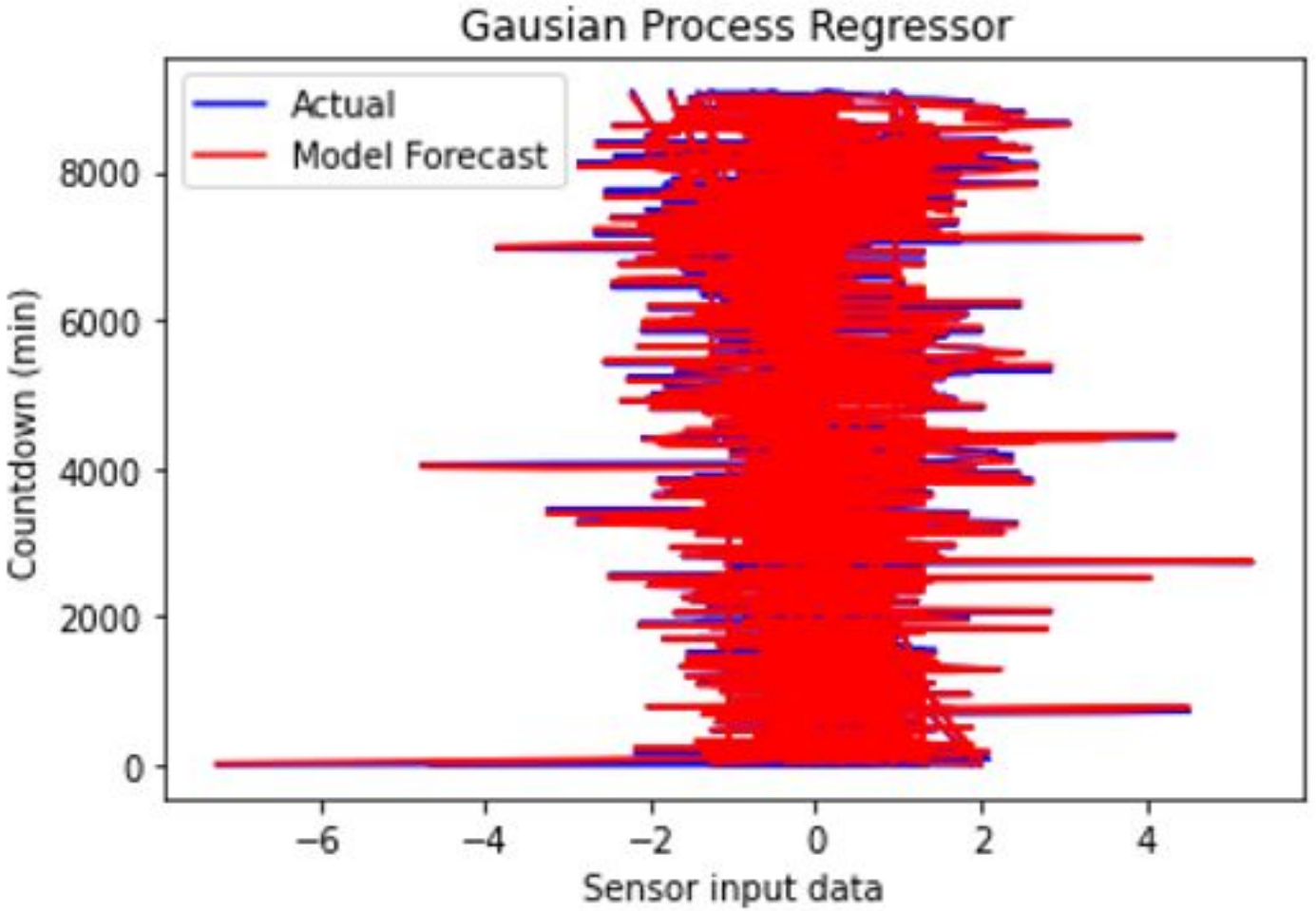
- Simple and Multiple Linear Regression Models
- Gaussian Regression
- Decision Tree
- Random Forest

Data Analysis Results: Single and Multiple Linear Regression Models

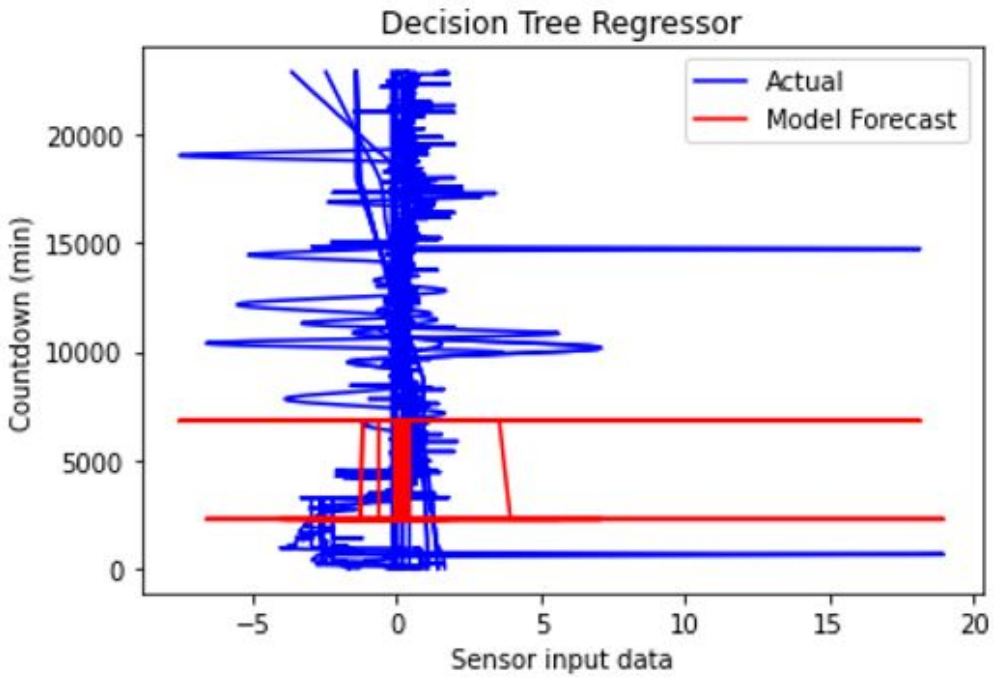


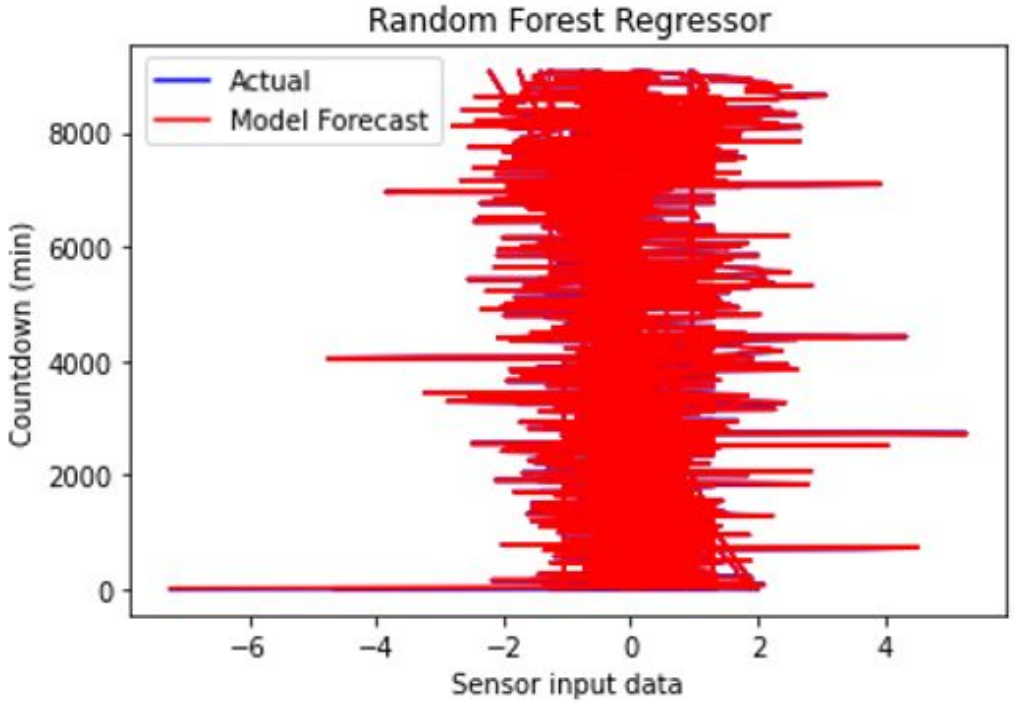


Data Analysis Results: Gaussian Process Regression Model



Data Analysis Results: Decision Tree and Random Forest Models



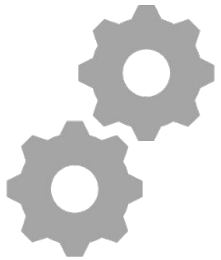


Comparative Analysis of Proposed Regression Models

	Single Linear Regression	Multiple Linear Regression	Decision Tree Regressor	Random Forest Regressor	Gaussian Process Regressor
Mean Squared Error	6.57E+07	2.24 E+02	1.72E+06	4.10 E+01	1.50 E+03
Explained Variance Score	0.636	0.999	0.750	0.999	0.999
Mean Absolute Percentage Error	58.32%	19.40%	39.32%	16.16%	12.24%

Limitations

Confidentiality



Feature Engineering

Dataset Size

Conclusions

- Developed a model to predict time-to-failure of equipment/facility
- Conducted data preprocessing and analysis
- Employing predictive models
- Limitations in our work and process
- The scope of our project investigated the accuracy of regression models in predictive maintenance
- Neural networks, the next step: models for limited a priori knowledge

Acknowledgment

First and foremost, we would like to express our gratitude to our advisors for their invaluable contributions to the success of this research. Second, we would like to thank the graduate students from UTP and TAMU for their assistance. We would also like to thank the National Science Foundation and our universities for making this collaboration possible. Finally, we would like to express our heartfelt appreciation to UTP for hosting us.

Thank you to everyone who has helped!

Acknowledgement of NSF Support:
This work was supported by the National Science Foundation's International Experience for Students (IRES) Site grant (Grant Numbers: OISE# 1952490-TAMU, 2208801-NC A&T State, and 195249-UNLV7). Any opinions, findings, conclusions, or recommendations presented are those of the authors and do not necessarily reflect the views of the National Science Foundation.