Universitetet
i Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Study programme/specialisation: Engineering Structures and Materials/ Mechanical Systems	Spring / ~~Autumn~~ semester, 20.19. ~~Open~~/Confidential
Author: Per Håkon Jøranli Sandmark	...........Per Håkon Sandmark........... (signature of author)
Programme coordinator: R. M. Chandima Ratnayake Supervisor(s): Aleksandar Primozic/ Arnaud Barré	
Title of master's thesis: Use of Machine Learning Techniques for Risk Based Inspection and Integrity Management of Pipeline Systems	
Credits: 30	
Keywords: - Machine Learning - Integrity Management - Oil and Gas - ILI - Corrosion	Number of pages: ....73.............. + supplemental material/other: .13........ Stavanger,.15.06.2019........... date/year

University of Stavanger
Stavanger, Spring 2019

Use of Machine Learning Techniques for Risk Based Inspection and Integrity Management of Pipeline Systems

Using Extensive Inspedtion Data

Per Håkon J. Sandmark
Internal Supervisor: R. M. Chandima Ratnayake,
External Supervisors: Arnaud Barré, Aleksander Primozic

Master thesis, Engineering Structures and Materials,
Major: Mechanical Systems

University of Stavanger

This thesis was written as a part of the Master of Science in Engineering Structures and Materials, Mechanical Systems at UiS. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Acknowledgements

I would like to thank Professor R.M. Chandima Ratnayake for the support and opportunity to work on such a challenging project, my external supervisors Aleksandar Primozic and Arnaud Barré for their support and guidance over the last 6 months. I would also like to thank Wood PLC Stavanger Norway for sharing the in-line inspection data and giving me full access to their knowledge base. Finally, I would like to thank the team at Wood PLC Stavanger for their invaluable input and guidance in addition to generously sharing their industry expertise.

University of Stavanger

Stavanger, 2019

Per Håkon Jøranli Sandmark

Abstract

This thesis demonstrates the use of machine learning for integrity management and assessment of pipeline integrity using various types of classification algorithms on comprehensive in-line inspection data. Machine learning is a subfield of artificial intelligence that includes abstruse statistical techniques enabling machines to improve problem solving experience and excels at identifying underlying statistical patterns enhancing in example predictive models, anomaly detection, and operational monitoring. Machine learning has been used with tremendous success on a range of problems across industries, and while it is important to be grounded and have realistic expectations for the value gain, there is no reason to believe that the oil and gas industry are any different as the literature study show.

Pipeline corrosion is one of the primary causes for pipeline failure. It is necessary to continuously monitor and analyze the pipeline to predict possible failures and ensure safe operation. In the oil and gas industry, In-line inspection (ILI) is an essential part of the integrity management (IM) of pipeline system. Corrosion is one of the primary concerns for the IM of a pipelines due to the potential for leakages and catastrophic failures. ILI allows for routine inspections of pipelines with high accuracy and is a great tool for identifying corrosion damage, and if necessary, is used to decide whether further detailed investigation is necessary. As a proactive IM strategy, it is highly dependent on the ability to accurately predict rate of corrosion growth.

The thesis gives the necessary theoretical background to integrity management of pipelines with the inherent risks of operation, failure modes, and defect assessment. The literature study focuses on the historical development and current developments for machine learning in the oil and gas industry, followed by a client case where various supervised machine learning models is developed and used to determine the structural integrity of sections of a pipeline based on data from in-line inspections with the objective of determining the suitability of machine learning for defect assessment and potential use in assisting predictive corrosion models.

The various machine learning models are developed through extensive experimentation with inspiration from other research and problem domains. As the ILI inspection supply comprehensive data, the importance of imposing constraints is important in order to make findings relevant in practice. The machine learning models[1] performed well on the available pipeline data, and models such as the XGBoost Classifier predicted class labels with accuracy of 100.0%. There are limitations to the models as they were not developed to take into account the assessment of complex shaped defects but the results are promising, and we can argue the case demonstrates the wider potential for machine learning in future work on defect assessment and integrity management.

Keywords – Machine Learning, Integrity Management, Oil and Gas, ILI, Corrosion.

[1]The Python-code for data preprocessing, training and evaluating the classification models is available on request to: perhaakon@live.no

Contents

List of Figures

List of Tables

List of Abbrevations

AI Artificial Intelligence

ML Machine Learning

RF Random Forest

RMSE Root mean squared error

RNN Recurrent neural network

CNN Convolutional neural network

MAE Mean absolute error

MLP Mult-layer perceptron

ILI In-line inspection

IM Integrity management

MAOP Maximum allowable operating pressure

ASD Allowable stress design

LRFD Load and resistance factor design

MFL Magnetic flux leakage

PCS Pressure control system

PoD Probability of detection

PoF Probability of failure

RP Recommended practice

OS Offshore standard

SC Safety class

SMTS Specified minimum tensile strength [N/mm^2]

SMYS Specified minimum yield stress [N/mm^2]

SO Stand-off data (distance from probe to pipe wall)

SORM Second order reliability method

ULS Ultimate limit state

UT Ultrasonic Technology

WT Wall thickness

WTSO Sum of SO and WT

Glossary

Anomaly	An indication, generated by non-destructive examination of an irregularity or deviation from sound weld or base pipe material, which may or may not be an actual flaw.
Arc strike	Localized points of surface melting caused by an electrical arc (also referred to as hot spot).
Construction feature	Feature that arises during pipe manufacture, transport or constructing of the pipeline, including a girth weld anomaly, arc strike and grinding.
Corrosion	An electrochemical reaction of the pipe wall with its environment causing a loss of metal.
Crack	A planar, two-dimensional feature with displacement of the fracture surfaces.
Dent	Distortion of the pipe wall resulting in a change of the internal diameter but not necessarily resulting in localized reduction of wall thickness.
Detection threshold	The minimum detectable metal loss.
Feature	An indication, generated by non-destructive examination, of an anomaly, change in nominal wall thickness, casing, reference magnet, pipeline fixture or fitting including tees, offtakes, valves, bends, anodes, buckle arrestors, external supports, ground anchors, repair shells and CP connections.
Grinding	Reduction in wall thickness by removal of material by hand filing or power disk grinding.
Gouge	Mechanically induced metal loss, which causes localized elongated grooves or cavities.
Heat affected zone	The area around a weld where the metallurgy of the metal is altered by the rise in temperature caused by the welding process. For the purpose of this specification it is considered to be within 3A of the center line of the weld, where "A" is the geometrical parameter related to the wall thickness.
Intelligent pig	A pig that can perform a non-destructive examination.
Metal loss feature	An area of pipe wall with a measurable reduction in thickness.
Mid-wall feature	Any feature which does not run out to either the internal or external surface.
Measurement threshold	The depth of metal loss or remaining wall thickness from which the width "W" and length "L" of an anomaly are measured.
Nominal wall thickness	The wall thickness required by the specification for the manufacture of the pipe.
Pig	A device which is driven through a pipeline by the flow of fluid, for performing various internal activities (depending on the pig type) such as separating fluids, cleaning or inspecting the pipeline.
Pig trap	An ancillary item of pipeline equipment, with associated pipework and valves, for introducing a pig into a pipeline or removing a pig from a pipeline.

Anomaly	An indication, generated by non-destructive examination of an irregularity or deviation from sound weld or base pipe material, which may or may not be an actual flaw.
Pipeline	A system of pipes and other components used for the transportation of fluids between (but excluding) plants. A pipeline extends from pig trap to pig trap (including the pig traps), or, if no pig trap is fitted, to the first isolation valve within the plant boundaries or a more inward valve if so nominated.
Pipe mill feature	A feature that arises during manufacture of the pipe, as for instance a lap, sliver, lamination, non-metallic inclusion, roll mark and seam weld anomaly.
Probability of Detection	The probability of a feature being detected by the intelligent pig.
Probability of Identification	The probability that a feature will be detected and correctly classified by the intelligent pig.
Reference wall thickness	The actual undiminished wall thickness surrounding a feature.
Reporting threshold	A parameter, which defines whether or not a metal loss feature will be reported. The parameter may be a limiting value on the depth of metal loss or it may be a function of depth, width and length of a metal loss feature.
Sizing accuracy	Sizing accuracy is given by the interval within which a fixed percentage of all metal loss feature will be sized. This fixed percentage is stated as the confidence level.
Spalling	Abrasion of the pipe surface resulting in shallow surface laps and possibly hardening of the material below.
Weld feature	Feature in the body or the heat affected zone of a weld.

1 Introduction

1.1 Problem Definition

How can Machine Learning techniques be used to improve risk based inspection and integrity management of pipeline systems?

1.2 Objectives and Approach

The objectives of this thesis are to perform exploratory work on the application of Machine Learning techniques towards risk based inspection and integrity management of pipeline systems, identify sections of corrosion features, determine severity of identified corrosion features, train, validate, and test selected Machine Learning models on selected datasets extracted from inspection reports, and compare test results with results from traditional defect assessment methods in accordance with DNV-RP-F101 Corroded Pipelines to determine the suitability of machine learning for the assessment of anomalies.

The expectations for the client are to get an enhanced understanding of how machine learning can aid existing procedures related to operational monitoring and integrity management of pipeline systems.

1.3 Limitations

The thesis focuses ILI inspection data from an inspection run performed in 2010. As the data available was limited in severity, some assumptions was made in order to enable the application of machine learning techniques towards the classification problem. The minimum depth threshold for anomalies was set to 10% of reference wall thickness instead of 2.0 mm which is the traditional threshold set when inspecting pipelines due to the low amount of high corrosion features recorded during the inspection run. The assessment performed in the Python kernel is based on the allowable stress approach in DNV-RP-F101 and features were labeled according to the depth percentage feature. Hence, the hypothesis is that the feature with highest importance and correlation with the class label would be depth percentage. It should also be mentioned that the assessment uses the assessment method for single defects that do not interact with other defects. As many of the recorded features are of feature type "main spot", indicating the deepest corrosion feature in a cluster making up a complex shaped defect, this results in some limitations for the validity of the models when performing classification of unseen data. The thesis does also not cover into detail the inner workings behind the different machine learning algorithms as this is beyond the scope of the thesis.

1.4 Thesis Structure

The thesis is structured in 3 major sections. The first part introduced the client case and gives the necessary theoretical background to integrity management of pipeline systems by performing a literature review. The second part focuses on the current application and development of machine learning technology in the oil and gas industry with the objective of performing an exploratory study on the possibilities for improving existing workflows for the client. The third part is about showcasing the comprehensive workflow involved when building predictive models by using machine learning tools. The objective of this client case is to enhance the understanding of how machine learning models can be deployed and how they can improve assessment workflows. The thesis finished with a discussion and conclusion on the potential of machine learning, practical barriers of machine learning, and theoretical limitations and avenues for further research.

2 Client Case

2.1 Enhancing Integrity Management using Machine Learning techniques on In-line Inspection data

Integrity Management of pipelines is essential for the ensuring safe operations and prevent potentially catastrophic failures. The primary concern for pipeline deterioration is corrosion. Hence, in-line inspections (ILI) is a vital part of the monitoring of pipelines subjected to corrosion, supplying detailed information about pipeline features such as corrosion features with detailed geometric measurements of size and position, in addition to other features such as anodes, girth welds, etc. The objective of an ILI is to inspect the level of corrosion and possible damage on the pipeline, locate defects, assess the severity of defects, and use the ILI history from previous inspections to estimate corrosion growth rates for each section of the pipeline. The ILI results are then used as a guideline for determining the Integrity Management Strategies (IMS) for the given system.

The objective of the client case is to apply ML techniques on ILI data from inspections performed with an Ultrasonic Tool (UT) on a pipeline on the Norwegian continental shelf in order to train, test and validate different ML models with the end goal of determining the level of integrity of the pipeline. The predicted results from the ML models are then cross validated against results from the traditional corrosion flaw assessment in accordance with DNV-RP-F101.

The case will be a binary classification problem where Class 0 indicates good integrity, i.e. low corrosion levels, or Class 1 indicating severe corrosion levels and subsequent reduced level of integrity. As a part of the case, an exploratory data analysis (EDA) is performed on the different features with the objective of identifying patterns in the data which can aid the predictive models.

The end goal of the client case is to enhance the client's knowledge about ML techniques, how to apply the tools, and how ML can be used to improve IM of pipeline systems.

3 Theoretical Background

3.1 Integrity Management of Pipeline Systems

Corrosion as considered as one of the primary concerns for pipeline integrity (Ahammed and Melchers, 1996; da Cunha, 2016; Choi et al., 2003). Hence, it is of utmost importance to implement necessary integrity management strategies to ensure safe operation by reducing risk and preventing potential failures (Kishawy and Gabbar, 2010). To enable the implementation of sufficient strategies it is important to understand the fundamental concepts of corrosion including the causes, growth behaviour, consequences, and mitigation techniques. Pipeline integrity management is a program which manages methods, tools and activities necessary for assessing the health condition of pipelines in addition to scheduling adequate inspection and maintenance activities to reduce the risk and costs (Xie and Tian, 2018). Pipeline integrity management programs consist of three major steps. Defect detection and identification, defect growth prediction, and risk-based management.

ILIs are performed periodically using smart pigging tools in order to detect defects and anomalies such as corrosion and crack features. Significant advances are needed as there are great challenges associated to accurately evaluate anomalies based on ILI data, predict growth rate of corrosion features, and optimize integrity activities in order to prevent failure of pipelines. The following chapter gives a comprehensive review on pipeline integrity management using ILI data.

3.1.1 Introduction

Pipelines can suffer from a variety of different defects such as corrosion, fatigue cracks, stress corrosion cracking (SCC), dents, etc. which can affect the level of integrity of the pipeline. If the pipeline is not properly managed, the consequences of a failure can be of catastrophic scale for the local environment. Failures can be in the form of either leaks or ruptures, and in addition to the potential environmental consequences, the potential expenses related to cleaning and downtime of production will be a heavy cost for the company operating the pipeline. Integrity is the top priority for pipeline operators ensuring reliable and safe operations. Reliable and safe operations lead to increased productivity, stable production, reduced costs, and prevents possible damage to the environment. It is of utmost importance to ensure safety, security of supply, and compliance with legislation and relevant codes for the offshore industry. Pipeline integrity tools are developed to manage risk, ensure compliance, and improve business performance. By implementing integrity management practices, we can reduce probability and consequences of failure, and improve the pipeline operating company's business performance by correctly assess and manage detected anomalies and defects. The role of a pipeline integrity management program is to monitor and predict the effect of defects in order to ensure when, where and what type of inspection, maintenance and repair is performed.

A great pipeline integrity management program should prevent failures from occurring, manage risks adequately, reduce costs for operators, and control damage effectively. The IM program can be divided into three major steps as listed earlier:

1. Anomaly and defect detection and identification
2. Defect growth prediction
3. Risk-based management

The detection and identification of anomalies and defects are performed by using tools which collects the necessary data such as visual inspection, nondestructive evaluation (NDE), hydrostatic testing, or in-line inspection. ILI is the most used inspection technology and the focus of this thesis. Hence, the thesis focuses on data resulting from ILI tools. Defect growth prediction is about predicting, as accurate as possible, at what rate the feature grows and remaining time before pipeline failure will likely occur. Features that are a threat to pipeline integrity are metal loss, cracking, third party damage, dents, etc. The third step, risk-based management, determines the suitable inspection intervals, maintenance and repair actions to ensure safe operation. Management models affect primarily the first and second step by potential changes to inspection actions and defect status classification. The primary objective of an integrity program is to achieve accurate defect prediction in order to find the most effective balance/compromise between reliability and costs, and it is important to note that some studies consider the design phase as an integral part of the management model. This is reasoned for by arguing that from a lifecycle perspective, design is an integral part as better design practices would lead to better confidence in the integrity of the pipeline. Palmer et al. (2004) provides a detailed introduction to the design stage of subsea pipelines and Antaki (2003) introduces an approach for taking into account the design stage of integrity management in the lifecycle cost modeling.

The following sections covers ILI tools and the major technologies to detect and identify defects, the performance and applications

3.1.2 Detection of Anomalies by In-line Inspection

Failure of pipelines can have dramatic consequences. Due to possible pipeline failures including leakage, environmental damage and high costs related to repair and potentially expensive replacements, accurate pipeline monitoring and inspection is essential. Varela et al. (2015) discusses and summarizes major methodologies, not limited to ILI, which is widely utilized for inspecting and monitoring of corrosion features located externally in pipelines in addition to discussing pros and cons for the different inspection technologies. Rankin (2004) further elaborates pipeline inspection techniques such as hydrostatic testing, ILI tools, tools designed for inspection of non-piggable pipelines, etc. The procedure for performing an inspection with a pigging tool and the capabilities of the tool are as follows. First, the tool is inserted into the pipeline and pushed through the pipe by the fluid flow. As it travels through the pipeline, the tool gathers all specific information related to the health and condition of the pipeline. The tool is capable of classifying the types of anomalies and their features including orientation, size and specific location in the pipeline (Shaik, 2015). R. Walker et al. (2010) elaborates on how to achieve reports of higher quality from ILI data, giving greater insight into what practices and technologies ILI service providers should invest in and embrace in order to ensure reliable service delivery.

Metal loss defects can be categorized into two main types. The first type is pressure-based defects and the second type is depth-based defects (Shaik, 2015). When considering depth-based defects, Shaik (2015) argues that the pipe is assumed failed when the defect depth reaches 80% of the initial wall thickness of the pipeline. For pressure-based defects as in example corrosion, the failure state is determined by the failure pressure, uncertainty model and selected safety factors.

3.1.3 In-line Inspection Technologies

To ensure the integrity of the pipeline during the operational life it is necessary to perform routine monitoring and evaluate the impact of the subsea environment on the pipeline. These inspections can be performed in two ways. The first approach is by external inspection carried out by either a remote operated vehicle (ROV) or by an autonomous underwater vehicle (AUV). The choice of inspection method is dependent on the type of corrosion features expected in combination with the location of the features (Nash, 2011). The most widely used approach for internal inspection is the use of in-line inspection tools such as smart pigging which is a device inserted and run through the pipeline recording comprehensive data of the pipeline integrity through corrosion features and all essential geometric data. The pigging tools uses ultrasonic, magnetic flux or visual inspection techniques to examine the condition of the pipeline.

The inspection devices used, referred to as "pigs", are devices which are inserted into either onshore or offshore pipelines to perform the desired task. Tasks can range from cleaning the inside of the pipeline by removing residues and objects which prevent or reduces the flow, to internal inspections for detecting corrosion and cracks which are a threat to the integrity of the pipeline. Pigs enables the assessment of pipeline integrity providing the basis for decisions related to maintenance and repair, preventing failures such as leakages and ruptures. The two main inspection technologies used for subsea pipelines are *ultrasonic testing* and *magnetic flux leakage* where the selection of technology is based on the objective of the inspection work. As the technology used for the ILI inspection is ultrasonic inspection technology, this will be further elaborated on in the next sections.

3.1.4 Ultrasonic Inspection Technology

Ultrasonic technology enables detailed inspection and measurement of pipelines in order to detect metal loss features at an early stage in order to prevent loss of integrity and leakages of potentially catastrophic scale. The following section covers the basic principles of ultrasonic technology and how it enables the detection and evaluation of e.g. corrosion features.

The concept of ultrasonic technology is based on the perpendicular incidence of ultrasound into the wall of the pipe. The pulse is then reflected from the back of the pipe wall forth and back until the energy dissipates. The received signal is a sequence of rear wall echos (RWE) which enables the measurement of remaining wall thickness. In principle, the process of determining wall thickness is to measure the time t between entry echo and the first rear wall echo, or measure the time t between two rear wall echos. With the basic assumption that the sound velocity v in steel is 5920 m/s, the wall thickness d can be determined:

$$d = \frac{v * t}{2} \tag{3.1}$$

In addition to measuring the depth of the defect, the tool does also have the capability of detecting if the feature is located internally or externally. This is a result of the stand-off distance between the probe and the wall. The pipeline is filled with a medium such as oil with a known velocity of sound. This enables the stand-off distance to be calculated directly from the entry echo and time-of-flight. The stand-off measure enables the distinction between internal and external metal loss features as the stand-off signal will change when an internal metal loss feature is present, illustrated in figure 3.1. In addition, ultrasonic technology allows for the detection of a wide variety of anomalies such as in example inclusions, slag, lamination, etc. This is due to inhomogeneities reflecting the signal from within the pipe wall.

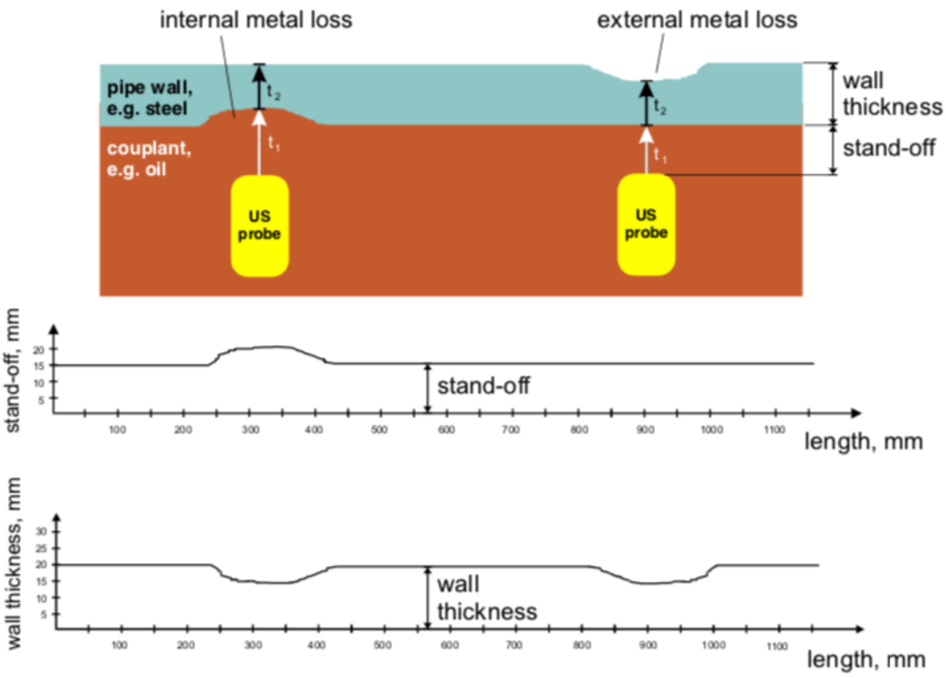


Figure 3.1: Principle of ultrasonic wall thickness measurement

3.1.5 Tool Specifications

This sections covers the ultrasonic inspection tool configuration used for the 2010 inspection run. The tool configuration is listed in table 3.1 and the tool schematics are displayed in figure 3.2.

Table 3.1: Ultrasonic inspection tool configuration.

Maximum wall thickness	60 mm[2]
Velocity at full spec	$\leq$ 1.12 m/s [3]
Temperature range	-50 – $+50$ °C[4]
Maximum pressure	120 bar
Minimum internal diameter in straight pipe	432 mm
Minimum internal diameter in minimum bend	457 mm
Minimum bend radius	3D/90° [5]
Tool length	Approximately 3850 mm
Tool weight	Approximately 620 kg
Number of bodies included sensor carrier	4
Distance range [6]	Approximately 295 km at 1.1 m/s [7]
Battery life time	Approximately 75 hours
Number of sensors	312
Axial sampling distance	Approximately 1.5 mm
Circumferential spacing sensor spacing	Approximately 1.5 mm

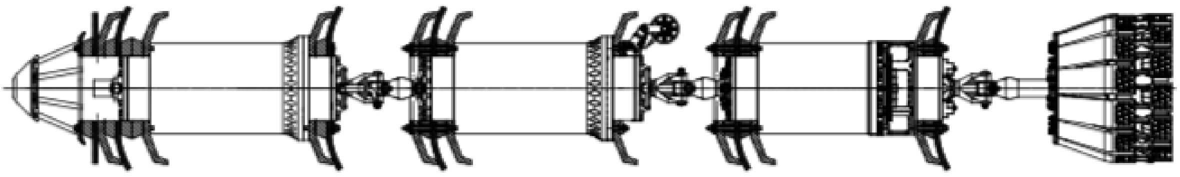


Figure 3.2: Schematics of 20" Ultrasonic inspection tool

3.1.6 Feature Specifications

The following tables list the inspection tool's capabilities for recording features during the in-line inspection. Important to note that the values represent the capabilities and not reporting thresholds.

Table 3.2: Metal loss - Inspection tool capabilities

Detection of metal loss with POD[8] > 95%[9] (due to low echo loss rate and good anomaly detection in the weld area)	**Without depth sizing:**	
	Minimum diameter	7.5 mm
	Minimum depth	1.5 mm
	With depth sizing:	
	Minimum diameter	15.0 mm
	Minimum depth	0.4 mm[10]
Discrimination int./ext.		Yes
	Resolution of wall thickness	0.06 mm
Depth sizing accuracy	Resolution of stand-off measurement	0.014 mm
	Accuracy of depth sizing at 95% confidence level	±0.4 mm
Length sizing accuracy	At 95% confidence interval	±3 mm
Width sizing accuracy	At 95% confidence interval	±6 mm
Accuracy of wall thickness measurement	Purpose of verifying nominal wall thickness	±0.2 mm

Table 3.3: Other features - POI > 95%

Mid-wall features	Laminations and inclusions minimum diameter	10 mm
Deformations	Dents and blisters	Yes
Weld detection	Girth-, spiral-, and longitudinal weld	Yes
Installations	Minimum diameter	25 mm
Bends	Bend radius <5D/90°	Yes
Repair areas (welded)	Sleeves, patches, attachements	Yes

[2]Subject to tool speed and depth of internal metal loss.
[3]Higher velocity on request.
[4]Higher temperature on request.
[5]1.5D on request.
[6]Depending on velocity, tool settings, and data storage limitations due to pipeline condition
[7]Larger distance on request
[8]POD = probability of detection.
[9]Can potentially be reduced in tight-radius hot bends.
[10]Lower threshold, in example for thin walls.

Table 3.4: Location accuracy

Distance	$\pm$ 10 cm from nearest GW
Orientation	$\pm$ 5° for diameter $\geq$ 20", else $\pm$ 10°

3.1.7 Evaluation Procedure

Data from the ultrasonic inspection tool is stored on a recording medium during the inspection run and subsequently checked locally for quality and completeness. After extraction at the receiver, the data is extracted from the recording medium and preprocessed for analysis. The following procedure can be divided into three major steps:

1. Identification of reference points based on girth welds in addition to supplied information from the client.
2. List of reference points is generated including marker points for marker transmitter system and natural reference points such as installations.
3. Analysis of recorded data.

During the third step, the analysis, anomalies are first determined by search program according to specified criteria and saved in database. From the previous database anomaly candidates are selected from a refined search based on criteria such as depth, width, length, etc. The selected anomalies are then analyzed and classified by qualified data analysts with the aid of data analysis software. After the analysis, the features are stored in a feature list and checked by senior engineer for correctness and completeness.

After the initial screening assessment and analysis, the metal loss features are assessed in accordance with the selected anomaly assessment method, e.g. DNV-RP-F101. Then the metal loss features are classified by their remaining wall thickness or depth, and the most severe features are fully assessed to ensure the integrity of the pipeline section.

The final step of the evaluation procedure is to generate the full inspection report consisting of statistics and diagrams, lists of all recorded results, detailed feature description and detailed assessment of features.

Pigs give accurate readings of the pipelines, but issues arise because large sections of the world's pipelines are not easily accessible for this inspections with pigging tools. The following section reviews the different signal processing techniques and algorithms for different types of ILI tools.

3.1.8 Performance and Application of In-line Inspection Tools

It is of utmost importance to understand the performance of the available ILI tools in order to use them correctly and achieve their full potential. The measure of performance can be subdivided into four measures:

- **Detection:** The capability to detect a feature. Probability of detection (PoD) should be above 90% for all ILI tools.

- **Identification:** The capability to successfully classify and report the correct feature type after detection. Probability of identification (PoI) increases with increasing size of feature. Incorrect classification of feature will have significant impact on accuracy of feature growth prediction.

- **Accuracy:** The accuracy for sizing of feature is the most significant measure to assess performance of the tool. Accuracy has a big impact on integrity management, and with increasing accuracy, unnecessary inspections will be reduced. Increased accuracy will result in improved selection of essential features and failure pressure will be predicted with increased accuracy.

- **Localization:** The capability to accurately locate anomaly. With high location accuracy, features can be compared with previous performed inspections and the growth history can be used to adjust models for e.g. future corrosion growth. This measure does also have a significant impact on the maintenance and repair activities.

Previous studies have introduced algorithms for classification, detection, sizing, etc. in order to assess and improve ILI tool performance. Caleyo et al. (2007) proposed a criteria for assessing performance of ILI tools. The proposed methodology allows for determining errors associated with estimation of true defect depths, and the proposed criteria were tasted in a case study using Mont Carlo simulations in addition to a real-life test study in order to present the application. Hrncir et al. (2010) presents a case study for improving the confidence level of feature information reported by utilizing revised sizing algorithms. Wang et al. (2015) presents a methodology to estimate the true depth of corrosion features based on detection theory accounting the soil property variation by combining a Bayesian inferential framework with cluster analysis. The case study presents three significant types of uncertainties in ILI tools which affect performance. Systematic errors of ILI tools, measurement noise, and random errors from the ILI tool and changes in surface roughness. McNealy et al. (2010) studies the effects on performance from combined measurement errors associated with the current technology employed for ILI of metal loss features. The errors and subsequent uncertainty of ILI data is handled in (Mora et al., 2008) which introduced a case study that identifies uncertainty effects of ILI accuracy for criticality assessment of metal loss features.

ILI data enables the assessment and prediction of conditions of pipelines, aiding the planning of integrity activities. Examples

3.1.9 Predicting Growth Rate of Defects

For predicting the growth rates for defects in pipelines there are 2 types of methodologies: Data-based and model-based methods. This section will cover the prediction algorithms, types of anomalies an defects that a pipeline can experience, in addition to methodologies for assessing the specific defects.

3.1.9.1 Data-based

Data-based methods use ILI data or test data in order to study the defect propagation stage. The application of ILI data in order to assess anomalies was discussed in the previous section, and the data-based approach gives essential information for predicting in example corrosion growth. Schneider et al. (2001) predicted defect growth rate and useful remaining life of a pipeline system using ILI data. The challenges for this study was the inaccessibility of certain pipeline sections which was handled by fitting statistical distributions for minimum wall thickness to the sample data. The extreme values from these distributions were then used in order to derive theoretically the corresponding distributions of the "unpiggable" sections of the pipeline system. The research highlighted the need for improved understanding of corrosion rates in selected sections of the pipeline were the solution for improved predictive power was to mount ultrasonic transducers which yields detailed information on the rate of corrosion.

3.1.9.2 Model-based

Model-based methods involve the application of physical models in order to perform defect and anomaly predictions using tools such as finite element method, and the remaining useful life of a pipeline can be predicted using physical models such as pipeline degradation models based on the failure probability. Liu et al. (2013) presents a case study using Bayesian networks to determine final probabilities of damage to subsea pipelines where the analysis aids the decisions related to risk-ranking and risk-reduction. There are a wide variety of models and methodologies for assessing and predicting the rate of growth for defects and estimate when a failure is likely to occur. The deciding factor for selection of methodology is the type of defect which the pipeline is subjected to at the specific section. Hence, the importance of classifying anomalies correctly as discussed in previous sections. In the following section, prediction algorithms, models and methodologies for assessing anomalies will be discussed.

3.1.10 Metal loss

Metal loss is the major threat for pipelines, can result in catastrophic consequences for the surrounding environment if subject to a rupture or collapse, and is primarily caused by either corrosion or erosion. The methodologies for discussion these to deteriorating mechanisms are discussed in the following sections.

3.1.10.1 Corrosion

Corrosion is a common mechanism for degradation of pipelines and is heavily affected by the surrounding environment. Corrosion is a natural mechanism where the pipeline materials react with the environment, and especially on the inside due to the internal working environment usually consists of a multiphase flow of hydrocarbons and water. We distinguish between two categories for pipeline corrosion: External and internal. According to Wang et al. (2015), the factors having the greatest impact on corrosion initiation and growth rate are the concentration of CO_3^{2-}, HCO_3, Cl^-, SO_4^{2-}, soil resistivity, soil moisture, half-cell potential, pH level, and the distance between the corrosion defect to the nearest cathodic protection installation. Alamilla et al. (2009) goes into further detail about the environmental parameters and effects influencing the propagation of localized corrosion damage and developed a mathematical model for corrosion damage propagation.

Models for prediction of corrosion growth rates can be further improved by implementing the corrosion damage data from ILI inspections by assessing the corrosion history of a pipeline resulting in a clearer picture of the actual development of corrosion which can be used to adjust the predicted growth rates. This assumes that the different inspection runs are performed with similar technology and with tools with similar uncertainties, etc. If a series ILIs are performed over the operational life of a pipeline, there is a great chance of the ILI service provider being changed from time to time together with the tool type and technology. Combining that with the development in accuracy of tools could potentially result in the different runs recognizing features as different types and at different sizes. It would be natural to expect greater depth of features of a run from 2010 versus a 2006 run, but this is not necessarily the case if the 2006 tool have greater uncertainty for depth measurement. Hence, studies of run comparisons have showed how certain pipelines were actually of higher levels of integrity than previously anticipated due to measurement errors. This is further discussed in (Ricker, 2007) which analyses corrosion data from pipelines from 1922 to 1940.

Failure of pipelines caused by corrosion can occur due to either the failure pressure becoming lower than the operation pressure, or because the depth of defects have reached the critical threshold for remaining wall thickness. The industry standard for wall thickness assumed failed is 80%. From the size and shape data recorded of corrosion defects in addition to pipeline geometry and materials, we can express a function for the failure stress of corrosion defects. The effects of corrosion defects on pipeline integrity and subsequent burst pressure and capacity equations has been studied to a great extent. Netto et al. (2005) presents an approach for estimating the burst pressure of corroded pipelines due to corrosion, taking into account localized pits with various depths and irregular shapes on the internal and external surface. The proposed model was used to determine burst pressure as a function of geometric and material parameters of different pipelines and corrosion defects.

Methods for assessing the integrity of pipelines has been studied and developed to a great extent. In the oil and gas industry, there are several code-based deterministic methods which are widely used to assess pipelines:

- DNV-RP-F101 (Veritas, 2015)
- ASME B31G (Committee et al.)
- ASME B31G Modified (Kiefner and Vieth, 1989b)
- RSTRENG (Kiefner and Vieth, 1989a)
- SHELL92 (Ritchie and Last, 1995)
- SAFE (Wang et al., 1998)
- PCORRC (Leis et al., 1997; Stephens and Leis, 2000)

All methods have similar equations and methodology as for how to assess detected anomalies, but differs primarily in how they handle the defect shape factor and bulging factor. Cosham et al. (2007) presents the best practices for assessing corroded pipelines and discusses the best techniques currently available for assessing a wide variety of pipeline defects. The most used standard in Norway is DNV-RP-F101 which is also primarily used for the assessment performed on the ILI data used as the basis for the client case of this thesis. The methods provide predicted remaining useful life and integrity level by determining burst pressure from equations taking into account defect information regarding shape and size in addition to the physical properties of the specific pipeline. The main factors which affect the equations for burst pressure are properties of the pipeline such as remaining wall thickness, diameter and ultimate strength. Therefore, by determining the failure criteria we can estimate the remaining useful life by creating physics-based models which take into account pressures and defect sizes versus time.

The main methods to be combined with deterministic methods for computing the probability of failure for corrosion defects are the Monte Carlo method, first-order reliability method (FORM) and first-order Taylor series expansion of limit state functions which are further elaborated in (Melchers and Beck, 2018). The application of Monte Carlo simulation was performed by Larin et al. (2016) which proposed a method for estimating reliability of a corroded pipe in operation. The application of FORM to assess probability of failure aided Teixeira et al. (2008) in predicting the remaining useful life of a corroded pipeline.

The single most important thing when assessing integrity is to calculate and predict the growth rate of corrosion defects. The growth rate of corrosion can be estimated by either physics-based corrosion models or by using ILI data. Shaik (2015) argue that using ILI data give more accurate results for cases where several inspection-runs of data are available as this enhances the prediction by the use of the corrosion history which can be used to adjust the physics-based corrosion growth models. Kiefner et al. (1973) presents a model for predicting and determining corrosion growth rate based on ILI data. It is important to note that there are great uncertainties associated with the estimation of historical corrosion growth rate due to different ILI tools and technology, issues with relating features from previous inspection runs, etc. This comparison of previous inspection runs were performed by Spencer et al. (2010) which presents a comparison of successive ILI runs for reducing bias when using the same service provider versus using different service providers. As previously mentioned, soil properties can also affect the external corrosion growth if the pipe is buried. This is presented in Wang et al. (2015) which presents a methodology which estimates external corrosion depth based on ILI data by combining cluster analysiswith a Bayesian inferential framework. Corrosion will be further discussed in section 3.2.

3.1.10.2 Erosion

Erosion is primarily caused by sand particles and particulates in the production medium which impacts and deteriorates the inner surface of the pipe wall. Zdravecká et al. (2014) presents the process and consequences of erosion failure induced by sand particles which are heavily featured in the production medium. Parsi et al. (2014) gives a comprehensive review of the modeling of erosion of pipelines including predictive models. Erosion prediction models can be based on the following three categories:

- **Computer-aided fluid dynamics (CFD):** CFD models have the capabilities to simulate a wide variety of scenarios for erosion taking into account different parameters on erosion rates, etc. CFD tools can give great results if deployed correctly, but they are time consuming in addition to being simulation-based which can potentially result in results which are not realistic.

- **Experimental methods:** Physical experimental methods aids in achieving more realistic predictive models. Physical experiments can be conducted resulting in high quality and realistic data, but this comes at the cost of potentially being time consuming and expensive.

- **Mechanistic methods:** Mechanistic modeling is an analytically approach to predicting erosion. Mechanistic models are fast to implement and relatively cost effective, but suffer from potential over-simplification of the case-setup resulting in limited accuracy towards the prediction of growth behavior of erosion.

As a result of the limitations for the previously mentioned models, several studies have been conducted, proposing approaches for combining the models in order to increase the predictive power of the models. Ukpai et al. (2013) achieved great results by utilized acoustic emission (AE) monitoring techniques in order to investigate the effect of impacting sand particles on the degradation mechanism of X65 carbon steel pipelines. The study utilized CFD in conjunction with particle tracking to model and predict the velocity and impact angle distribution in order to predict the kinetic energy and resulting erosion damage in an erosion-corrosion environment. Tang et al. (2009) presents an approach for predicting the remaining useful life of pipelines subjected to multiphase flow induced erosion-corrosion, mainly focusing on the interactions between the multiphase flow and the pipeline.

3.1.10.3 Cracking

Cracking is another defect type which can occur. Cracking is a time-dependent threat with potentially catastrophic outcome, and can be divided into two primary types:

- **Stress corrosion cracking (SCC):** SCC is the creation an growth of cracks aided by the surrounding and internal environmental conditions enveloping the pipeline structure.
- **Fatigue crack propagation:** Fatigue crack propagation is caused by a variation in pressure, either cyclic loading or random depending on how the pipe is operated, which continuously weakens the pipe for each cycle.

The following section will primarily focus on SCC. SCC consists of three primary stages:

1. **Initiation:** Low growth but highly affected by the surrounding environment.
2. **Propagation:** Stable growth rate of the crack.
3. **Failure/rupture:** Occurs quickly and cannot be accurately modeled.

Stage 1 and 2 is the primary focus as this is where integrity management strategies can aid the prevention and mitigation of SCC resulting in failure. The remaining life of a pipeline subjected to corrosion can be defined as:

$$N_f = N_i + N_p \tag{3.2}$$

where N_f is the fatigue life, N_i is the number of cycles until crack initiation, and N_p is the number of cycles until propagation. Hence, the essential aspect here is the number of remaining cycles of useful life for the pipeline. Cracking will not be discussed further as the thesis focuses on features caused by primarily corrosion. Hence, the next sections will elaborate further on the corrosion of structures in offshore environments.

3.2 Corrosion

This section will focus on corrosion uncertainty modeling for steel structures, reviewing fundamental effects of corrosion, factors affecting the severity of corrosion, potential consequences, and how to mitigate and prevent corrosion.

3.2.1 Corrosion of Steel

Corrosion is the natural process in which a refined metal converts into a chemically more stable state such as its oxide, sulfide, or hydroxide. The process is either electrochemical or chemical, and results in the gradual destruction and breakdown of the metal as a result of it reaching with its environment. For corrosion of steel to occur, the general prerequisite is the creation of a corrosion element. There exist three types of corrosion elements (Hamann and Lampe, 2018):

- **Local element corrosion:** Corrosion is occurring i.e. due to local deformations/strains in hull of a structure, high temperature, etc.
- **Contact element corrosion:** Occurring due to different metals with different electro-potential being in contact, either through a weld, rivets, etc.
- **Concentration corrosion:** Occurring due to the metal being subjected to an electrolyte which results in pitting corrosion, crevice corrosion, uniform corrosion, etc. This is the most common type of corrosion.

Corrosion occurring with the presence of mechanical loading have relatable properties to the first and third corrosion elements.

Marine environments are considered to be highly corrosive for mild and low alloy steels, which of the use is widespread in the offshore industry as it has a low investment cost compared to high-strength steel alloys. Research show that up to 90% of failures for ships are related to corrosion and corrosion fatigue, and a significant number of environmental disasters are related to insufficient maintenance and resulting corrosion (Melchers, 1999).

Preventive measures such as cathodic protection and coatings combined with adequate maintenance should mitigate corrosion, but history show that existing maintenance procedures are not always sufficient as requirements for components and structures changes. There will always be areas where corrosion is inevitable as preventive measures are difficult to impose. Hence, the probabilistic modeling for expected corrosion is essential knowledge. The assessment of corrosive damage on existing structures could potentially have severe economic consequences if in example significantly conservative estimates of corrosion rates are considered, resulting in premature condemnation of the structure as the wrong residual life is estimated (Melchers, 1999).

3.2.2 Subsea Pipeline Corrosion Mechanism

The resistance of pipelines against internal and external forces becomes weakened as the pipeline is subjected to corrosion - the leading factor causing integrity loss (Yang et al., 2017). The following two sections will cover the essential categories of pipeline corrosion, internal and external corrosion, which of both are electrochemical processes.

3.2.3 Internal Corrosion

Internal corrosion is primarily caused by chemical agents, solids , and fluid flow, contributing to the electrochemical processes resulting from the presence of contaminants such as carbon dioxide (CO_2), hydrogen sulfide (H_2S), and microbiological growth. CO_2 dissolves in water, dissociates to bicarbonate anion creating hydrogen ions, and acts as oxidizers in the pipeline resulting in internal corrosion. The corrosion caused by CO_2 develops slow, and the growth rate of corrosion is increased with increases in concentration of CO_2, temperature and pressure. H_2S-induced corrosion has 4 major forms resulting from 4 different states of environmental conditions (Sulaiman and Tan, 2014):

1. Hydrogen-Induced Cracking (HIC) resulting from atomic hydrogen diffusion on the pipeline.
2. Sulfide Stress Cracking (SSC) resulting fro the joint effects of stress and corrosion.
3. Pitting attacks at cracking area of sulfide film formed on surface of pipeline.
4. Sulfide pitting corrosion from deposition of solid sulfide formed by reaction of ferrous ions and hydrogen sulfide.

Microbiological induced corrosion (MIC) is caused by biological growth and resulting Sulfate-Reducing Bacteria (SRB) which generates $CO2$, water and sulfide as they oxidized fatty acids leading to similar corrosive behaviour to pitting corrosion for sour environments (Sulaiman and Tan, 2014). If the pipeline is subjected to high flow rates and velocity, the flow may cause accidental damage to protective layers causing accelerated corrosion growth rates. In a situation with high velocity, high occurrence solids such as sand and particulates are especially damaging to protective layers (Ilman and Kusmono, 2013).

3.2.4 External corrosion

As a predictive measure against external corrosion, pipelines are coated with a protective layer in addition to often being buried below the sea bed to protect against currents, fishing equipment (Sulaiman and Tan, 2014). Types of pipelines which are not buried and fully exposed to seawater are e.g. risers, deep-water pipelines, tie-in spools, and up-crossing segments. The deciding factor for determining which type of corrosion mechanism that occur is the surrounding environment. If the surrounding environment is consisting of seawater, the dissolved O_2 will act as an oxidizer at the cathode resulting in corrosion of not adequately protected. Similar to the case of internal corrosion, increases in concentration of O_2 on the surface results in increased potential corrosion rate. To reduce the potential for external corrosion, coatings are of high importance. If this protective layer is broken, e.g. through mechanical damage from third parties, the rate of corrosion will increase significantly. As stated, the major oxidizing agent is dissolved oxygen in combination with potential damage to the coating. Hence, pipelines buried below sediment as a preventive measure will result in the pipelines being subjected to significantly less dissolved oxygen concentration in addition to be less susceptible to damage. This reduced the risk of anaerobic, but at the cost of greater potential for MIC (Sulaiman and Tan, 2014).

3.2.5 Potential Consequences of Corrosion

The consequences of corrosion resulting in failure and the severity of the subsequent failure depends on several factors. If the pipeline is located onshore, the result of a leakage can be be fluid leakage, fire, explosions, etc. depending on the surrounding environment. If the content transported is e.g. natural gas and the leakage occur in the close proximity to electrical equipment, there would be a high probability of an explosion occurring. If the failure occurs in e.g. a natural habitat, the environmental damage could be catastrophic. There is also the chance of a leak occurring offshore and not immediately being detected, either due to a lower pressure drop than what would be noticed or just shear negligence. In addition to the potential loss of human life and environmental damage, we have the potential economic consequences and ramifications. A pipeline leakage can result in loss of production for a significant time period due to long and challenging maintenance work. Hence, the integrity management of pipeline systems is of high importance.

4 Machine Learning

4.1 Introduction

In the previous section, the thesis covered the theoretical background for integrity management of pipeline systems. In this section, the objective is to further study the application of machine learning in the oil and gas industry towards improving in-line inspections, condition assessments, identification and categorization of structural integrity of pipelines. The results from the literature study will be taken into account for the following chapters as to decide which parameters are selected for different algorithms in the methodology section in addition to feature selection in the data preprocessing section. The data will also be subjected to exploratory analysis with the purpose of identifying underlying patterns and feature importance.

4.2 Machine Learning - Current Status and Development

The oil and gas industry has traditionally been conservative in embracing new technological developments rapidly, but due to the economic downturn and subsequent cost reduction over the last 5 years, the shift towards the digitalization of workflows using ML and advanced analytic tools have been on the agenda for the industry. The methodologies are not new, but the implementation in production processes and integrity management is relatively new. Companies do not necessarily struggle with understanding the methodology but struggle with the implementation and deployment of models in existing production systems and in operational processes. Hajizadeh (2019) performs a review of recent developments and practices of machine learning, offering a SWOT analysis aimed towards increasing business value by aiding strategic management in order to implement the technology. The study shows that the annual number of papers towards the implementation and development of machine learning has gone up exponentially over the last two decades, giving a clear indication of how the industry is trying to adapt new technology in order to increase effectiveness of operational procedures, production processes, integrity management, etc. Liu et al. (2019) gives a comprehensive demonstration of how ILI inspection data can be utilized for corrosion characterization. The paper proposes an automated method to match multiple ILI inspections and their recorded feature data in order to perform matching of the corrosion features. The matching problem is treated as a classification problem and classified as either matched or unmatched. The features selected for the model are absolute distance, circumferential position, defect length width and depth percentage. In other words, these are the features to keep an eye on in the client case where we will explore the feature importance for predicting the class label. The algorithms used are support vector machines, decision tree, random forest, and ensemble learning. The results look promising as the predictive models managed to accurately classify and match the corrosion features with an accuracy of just above

90%. The result shows great promise as the proposed method enables the automatic matching and identification of features, reducing the amount of manual labour required.

As of today, the industry have available enormous amounts of data related to operation of production facilities, heavy asset management systems, integrity management systems, etc. One of the big issues the industry is facing is in the availability of the data. If we consider a production system for oil and gas, it consists of a wide variety of equipment, systems, and sensors. In addition, the systems can potentially be operated by different companies contributing to the challenge of accessing all relevant data.

The primary issue the industry are facing is extracting ***value*** from data.

4.3 Fundamental Steps in Extracting Value from Data

The issue with value of data is related to the form it is presented and its availability. Sensor data could e.g. be of different signal types, coming from both old and new systems and equipment with decades different technology, and not necessarily easily connected to find patterns and predict production variables. Hence, we say that the data is structured in different silos, resulting an difficulties in extracting and accessing relevant data. The three fundamental steps in extracting value from data are the foundation of the horizontal data platform.

1. **Data sources:** Data liberation from source system - remove the data silo. Evergreen data available instantly anywhere.
2. **Ingest, normalization and contextualization:** Continuous optimization and contextualization of often incomplete data. Common data model enables cross domain analytics and visualization.
3. **Value capture:** Unique tools and open APIs to ease value capture and speed of operationalization accross all assets.

This is the basic structure of an horizontal data platform which is the ideal way of structuring assets, increasing the value-add for the end user which could be e.g. a production engineer trying to implement models to increase the predictive power of a system which tries to predict operational status for a production system the next 6 hours taking into account variables such as pressure and temperature sensors at different parts of the system, weather data, historical data, flow assurance modeling, operational anomaly history, etc. The objective would be to make the relevant data readily available for the end user, reducing the necessary expert knowledge for applying data analytics tools for understanding and processing data, resulting in a faster deployment of new techniques.

4.4 Establishing a Solid Foundation for the Deployment of Machine Learning

To be effective with machine learning, we need a solid foundation for our data. We start with data collection. This involves connecting instrumentation, logging, sensors data, sensor metadata, external data, user generated content, 3D models, maintenance logs, etc. into the system. The second step would involve moving and storing data by establishing reliable data flows, necessary infrastructure and data pipelines, and handle structured and unstructured data. The third part is the exploration and transformation of data by cleaning, perform anomaly detection, and prepping the data for further analysis. The forth step is aggregating values and labels to the data. This involves performing data analytics, creating aggregate values, evaluating feature importance, creating training data, etc. The fifth step towards having a machine learning model prototype is the learn and optimize-phase. Here, the algorithms are applied and experimented towards solving the desired problem. When the model is trained, it can be applied to test data and be ready for further development or deployment in an existing system.

4.5 Machine Learning

Machine learning is the most exciting field of computer science at the current moment due to the high abundance of statistical data which are recorded on a daily basis in our society. By using self-learning algorithms from machine learning, we can convert the data into knowledge which can be used to increase efficiency in production, detect features, and numerous other fields of application. In this chapter, we will focus on the main concepts and different variants of machine learning which are readily available through the many powerful open source machine libraries. The chapter we will cover the general concepts, the three types of machine learning, and the essential steps involved in building successful models.

The objective of this section is to increase the understanding of statistical methodology in order to leverage the machine learning algorithms in Python's scikit-learn library, contributing to the knowledge of how to solve machine learning use cases.

The foundation of machine learning algorithms is statistics. To apply machine learning towards problem solving, it is important to understand the underlying statistics. One of the main goals of ML is to find statistical dependencies in the data at interest. The available data could be used to in example check blood pressure against age of or to recognize handwritten text.

When investigating the potential use of machine learning approaches for solving a problem, it is important to understand properly what machine earning is in this context. Literature often refer to artificial intelligence (AI) techniques and it is important to understand the differences and how they correlate. The term artificial intelligence is a commonly known term often portrayed in movies such as The Matrix and The Terminator to portray thinking machines capable of operating independently of the inputs of an operator. Both term are often used interchangeably, resulting in unclear differences. The following section will explain what the term mean and their differences in the context of the thesis.

The concept that AI involves machines that are characteristic of human intelligence was first coined by the founding father of artificial intelligence, John McCarthy (Morgenstern and McIlraith, 2011). The broad term can be narrowed down to the ability of performing tasks such as planning, understanding of language, recognizing sound, object and feature recognition, learning, and solving problems. AI technology can be differentiated into two main categories: General and narrow. General includes all the aforementioned abilities, while also resembling all characteristics of human intelligence. Narrow AI on the other hand revolves facets of human intelligence, is only able to solve those facets, but can perform those facets exceptionally well. In the context of the thesis such facets would be the ability to classification and feature recognition.

In other words, machine learning is an approach of achieving AI, and AI could be defined as the ability to learn without being explicitly programmed. Hence, AI techniques for classification and feature recognition could be achieved without machine learning, but without the usage of machine learning the following approach for AI would require building an extraordinary amount of lines of code taking into account all potential variables with complex rule-sets and decision-trees. Hence, the usage of machine learning could be compared to "training" the AI algorithm to adjust, improve and learn the different features and create suitable classification models which would become increasingly better for each iteration. This is way more efficient than coding software-routines for every complex outcome of integrity of a pipeline system, and more data fed to the system would result in improvement of the models.

4.6 Categories of Machine Learning

Machine learning can be divided into three categories: Supervised, unsupervised, and reinforcement learning. Supervised learning uses labeled data, direct feedback, and is used to predict outcomes. Unsupervised learning involves no labels, no feedback, and is about finding hidden structure in data. Reinforcement learning is a decision process with a reward system used to guide the learning models on how to achieve the desired solution.

The main goal in supervised learning is to learn a model from labeled training data which enables the prediction of unseen or future data. The term supervised refers to the labels that are already known which the models uses to learn from. A great example of supervised learning is the spam filter on your email. By labeling emails as spam, it learns the patterns and features of spam mail and subsequent filters it out after sufficient training. This is called a classification problem similar to the client case of this thesis. Another category of supervised learning is regression where the model predicts a continuous output value based on predictor values which can be correlating features.

Classification is a part of supervised learning where the objective is to predict an output based on previously processed labeled data. After sufficient training data is supplied, the supervised learning algorithm can accurately predict the outcome or label of the new unseen dataset. The classification algorithms are not limited to only binary classification, but can also be used for multiclass classification tasks such as identifying handwritten labels. In supervised learning we already know the class label. In reinforcement learning we uses a defined measure of reward related to specific actions by the model to increase the converging towards the ideal solution. In unsupervised learning on the other hand, we have unlabeled and unstructured datasets. Thus, it is necessary to perform an exploratory data analysis of the data in order to identify underlying patterns and extract meaningful features with information contributing to a solution. This can be performed by cluster algorithms which enables the binning of samples based on their similar attributes. In unsupervised learning we do not know the structure or clusters which each feature is related to, but the technique is great for identifying new relationships and patterns in the data.

In addition to clustering, we have another subtype of unsupervised learning models called dimensionality reduction. This involves principal component analysis (PCA) which takes high-dimensional data and reduced the number of dimensions in order to make the data easier to handle and easier to interpret. High-dimensional datasets can be difficult to interpret and identify underlying patterns, but by performing a PCA there is a chance that the dimensionality reduction results in useful visual information as the feature is now projected into a 3D or 2D plot showing patterns previously unfamiliar. After this is performed, the data can be labeled and used to train other machine learning models in order to perform e.g. supervised classification on similar high-dimensional datasets of unseen values.

4.7 Popular Machine Learning Algorithms

Here, some of the most popular types of machine learning models are presented.

Linear Regression

Linear regression is considered on of the simplest and easiest models to implement, and can be quite useful when we desire to predict a continuous value or a value as a result of two predictors in contrast with Classification algorithms where the output is of categoric nature. Hence, if the desired output is to predict future values for a production process in real time, regression algorithms would be highly suited. It should be noted that linear regression models do have some issues related to the stability when there is redundant features present in the dataset. Common use cases for regression models are predicting production volume for next month, predicting time for transportation between two points, estimating blood pressure value based on age, etc.

Logistic Regression

Logistic regression algorithms performs binary classification problems and is gives a binary output of either 0 or 1. To achieve the binary output, the algorithm takes linear combinations between features and applies a non-linear function such as the Sigmoid function in order to resemble a basic neural network (Raschka and Mirjalili, 2017). The use of logistic regression models is great due to the possible regularization it provides of features. Hence, reducing the possible effects of having redundant features with high collinearity. In comparison, Naïve Bayes models are highly dependent on not having correlated features as they do not learn the interacting functions between features. This will be displayed in the results where the Naïve Bayes models perform relatively bad due to the high feature correlation in the dataset. Logistic regression enables easy interpretation of outputs, and the models can easily be updates with new data such as ILI inspection data with different levels of corrosion compared to the one previously trained on. This enhances the predictive power of the model. Logistic regression is especially useful when creating a probabilistic framework and for when new data will be readily available to update the model. In addition, logistic regression also aid in the process of understanding underlying patterns and contributing factors behind the predicted output. Hence, one of the big selling points of logistic regression is the fact that is simply not a "black box" model where the underlying algorithms are quite difficult to understand, resulting in outputs which are hard to interpret. By understanding the process, we also understand to greater extent the potential uncertainties of the model based on the available dataset. Use cases range from predicting the amount of customers that will visit a shop to measuring effectiveness of improvement initiatives in process industries.

Decision Trees

Decision trees are used in ensembles to achieve efficient algorithms such as Random Forest and Gradient Tree Boosting, allowing for exceptional handling of interactions between features with high correlation in addition to being non-parametric. Hence, decision trees are great for handling datasets containing outliers or having values that are not linearly separable. The downside to using decision tree algorithms is that they do not support online learning, which basically means that for a model being integrated in a production system it has to be taken out and completely rebuild in order to learn from new data. In other words, it does not perform well in theory if trained on one set of pipeline systems and then being fed data from a pipeline system with a different level of corrosion. The major disadvantage of decision trees is the potential for overfitting as the depth of the tree can be to deep resulting in a complex model which fits great on the training data but does not generalize well to unseen data. Hence, when using decision trees the preferable solution is to use an ensemble of methods such as random forest which negates the negative aspects with decision trees, reducing the chance of overfitting. Another downside with decision trees with high depth is that they require large amounts of storage for data. Hence, if the RAM of the computer system is insufficient, the hard drive will be required to store the training data which results in dramatically increased computational time. Typical use cases are classification problems and decision problems like investment decisions.

K-means

If we are dealing with unlabeled data, the objective will be to assign labels based on the features of the data. This is called clustering, and clustering algorithms are quite useful for predicting labels when we have large datasets with many different features which we want to distinguish and cluster based on certain attributes. Clustering algorithms are suitable for organizing and grouping data, all depending on the initial problem statement and the features in the data. K-means work by setting an amount of desired clusters and then giving the clusters an initial guess which is then per iteration approximated closer to the actual center of each cluster. This is an efficient technique of identifying groupings of data which at first glance can look chaotic, but this is also the big disadvantage with K-means - the fact that it require knowledge about the amount of clusters in the data. Hence, if we do not know the amount of clusters, we will have to guess the best K number of clusters.

Principal Component Analysis

Principal component analysis (PCA) is an analytics tool which is used to perform dimensionality reduction of data sets with a great number of correlating features. This reduces redundancies and the risk of overfitting. The PCA is a low-dimensional representation of a set of variables which can make underlying patterns of data visible and can provide patterns in the data that are characteristic to certain groups of samples, aiding in e.g. clustering analysis. For further info on PCA, Raschka and Mirjalili (2017) provides a thorough presentation of how to implement PCA in machine learning applications.

Support Vector Machines

Support vector machines (SVMs) are a supervised machine learning technique which is widely used in pattern recognition and classification problems. It is especially suitable for binary classification problems. SVMs creates a decision region based on the samples closest to the decision region as to achieve the optimal decision region without underfitting or overfitting. SVMs have high accuracy and works well with data which are not linearly separable in the primary feature space, e.g. 2-dimensional. SVMs are great for high-dimensional space problems, but this comes at the cost of SVMs being memory intensive in addition to being hard to interpret. Typical use cases ranges from recognizing hand written letters to detecting persons which would likely suffer from diabetes.

Naïve Bayes

Naïve Bayes are a machine learning technique used for classification problems and is based on Bayes' Theorem (Zhang, 2004). The model is fairly easy to build, implement and deploy, and is suited to data sets of great sizes. They are known to outperform highly advanced classification methods relative to the effort involved in implementing them. Hence, this is one of the algorithms which will be deployed in the analysis. As it is quite simple in nature, it does not require as much CPU and memory power in order to run efficiently. This increases the possibility of implementing these models in e.g. existing production systems where the computational power may be limited. The models work by converging for each iteration towards the solution if the NB conditional independence assumption holds. The model has a tendency to converge faster than e.g. logistic regression, making it less dependent on large sets of training data. In short, NB models are great for fast deployment and relatively accurate results, but the disadvantage as stated earlier is the missing feature interaction learning. In other words, it will not necessarily recognize the relationship between depth percentage of the corrosion features in the pipeline and the remaining wall thickness. Typical use cases are recommendation systems such as Netflix and the shows you are promoted which are based on your viewing history, sentiment analysis, and face recognition. These applications the NB models do all require the models to perform efficiently without requiring heavy computational power, again underlining the suitability of NB models for these use cases.

Random Forest The random forest classifier is an ensemble of decision trees and is highly suited to classification problems. It performs well on large sets of data and is especially useful for predicting specific variables from thousands of input variables due to it taking into account the feature correlations and easily scales to any number of dimensions, resulting in great performance levels. The downside to random forest is the potentially slow learning rate for each iteration. This depends on the parameterization, and trained models are not able to improve and learn without being rebuilt. Random forest is great for predicting failure of production equipment in manufacturing, or predicting class labels of samples.

Neural Networks

Neural networks models are based on the human mind, and takes into account the weights of connections between neurons in hidden layers between the input and output. The neural network works by balancing the weights of the different neurons for each iteration and balancing the predictive weights based on the predicted and true output. Hence, when all weights are trained, the neural network can be efficiently utilized to predict e.g. class labels. Neural networks are models of high complexity, making them suitable in creating taylormade complex solutions for solving problems with high accuracy, but the extreme complexity results in the models assuming the "black box" designation which means that the model and its result can be hard to interpret. One of the big advantages is the reduced need for feature engineering and preprocessing of data by normalization or standardization before training the models. In addition, neural networks can be utilized in Deep Learning which are highly complex models which enables the creation of models taking into account several different ML models, resulting in previously unthinkable problems being solved. Neural networks are suited to both supervised and unsupervised learning, but the downside to neural networks is the high complexity which makes them hard to use and deploy. In addition, neural networks require extreme computational power and can be a significant toll on CPU and memory. Hence, they perform best by being run on GPUs and within the memory constraint of the computer RAM.

It is difficult to predict which algorithms will perform best and worst on the ILI inspection data.

The "no free lunch" theorem states that there is no single algorithm working best for all cases. Hence, the process of building machine learning models involve working iteratively towards the better performing ones by trying out different algorithms against each other and cross-validate the performance against each other. In order to achieve the optimal solution for a problem statement, is is important to understand the demand of the client, the rules and regulations, and take into account expert domain knowledge - all contributing factors toward the best performing machine learning model.

5 Data

5.1 Data Collection

This chapter will cover the in-line inspection details and data collection process. This involves operational details from the inspection selected inspection run, scope of the inspection, operational details, summary of reported findings, tool specifications and accuracies, and the data evaluation process which was performed.

5.1.1 Inspection Reports

The ILI data available for the given pipeline is supplied by 4 different inspections performed over the course of the operation history. The selection of the pipeline is based on the availability of necessary data from the client and is listed in table 5.1 Solutions (1998). The objective of the ILI is to detect, size and localize metal loss features such as internal and external corrosion, in addition to mid-wall features such as laminations and inclusions. The thesis will primarily use data from the 2010 inspection run for the client case. This is due to previous tools having greater uncertainty in depth measurements in addition to the data not being readily available in suitable formats. Hence, the earlier inspection reports were disregarded for the clientcase.

Table 5.1: In-line Inspections for Ula to Ekofisk Pipeline.

Inspection Reports				
Inspection Year	Inspection Tool	Inspection Length [m]	Pipeline	Operator
1997	UT	71141.10	Ula to Ekofisk 2/4R	PII Pipeline Solutions
2004	UT	72923.80	Ula to Ekofisk 2/4J	PII Pipeline Solutions
2006	UT	72750.33	Ula to Ekofisk J	NDT Systems and Services AG
2010	UT	72875.09	Ula to Ekofisk J	NDT Systems and Services AG

The table shows different lengths for the different inspections. The reason for the large variance between 1997 and the subsequent years is different end points. Ekofisk 2/4R was shut down and disconnected from the pipeline in 1998, and decommissioned in 2013. Hence, different end points for the inspections.

5.1.2 2010 Inspection Report

For the 2010 inspection run, the quality of the available inspection data is good and fully meets the tool and feature specifications. The inspection data were recorded by all sensors covering the full circumference of the pipeline in addition to the entire length of the pipeline of approximately 72,875 meters. The operational details of the run is listed in table 5.2. From the report, some observations are important to take into account when considering the quality and accuracy of the recorded data. For the first 0.5 km of pipeline, the inspection data is affected by a case of multiphase flow where two different media, crude oil and water, was mixed due to the following listed circumstances:

- Initial pressure test of launcher was performed by filling it with water in addition to pressurization by nitrogen injection. The procedure then requires the launcher to be drained of water before being replaced by crude oil. The draining procedure was unsuccessful, resulting in the tool being surrounded primarily by water at the start of the inspection.
- As the inspection was started, crude oil entered the top of the pipe resulting in the increase of crude oil content with the subsequent decrease of water content for the bottom part of the pipe.
- The heterogeneous oil-water interphase results in echo loss of between 30 mm and 200 mm for the first 250 m of pipeline, illustrated in figure 5.1. The vertical axis is the circumferential position of the pipeline from 0 to 360 degrees.

The number of detected metal loss features was 771, which of approximately 75% were classified as internal features (537 of 771).

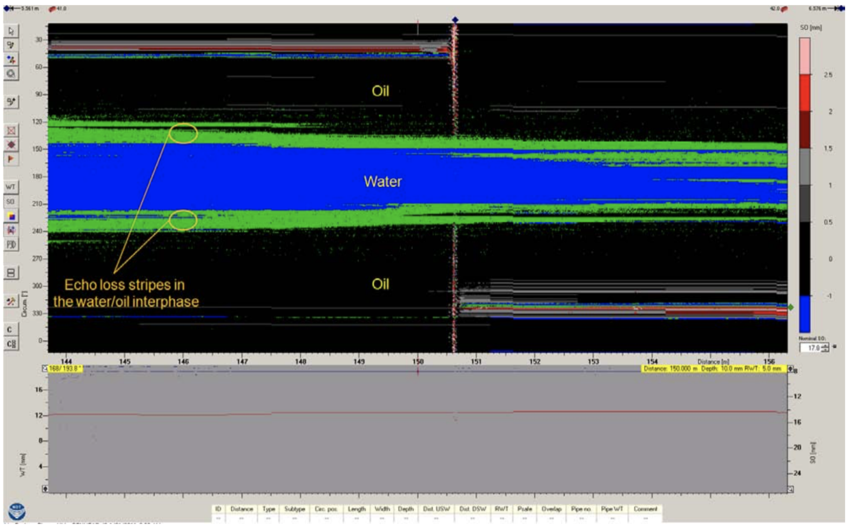


Figure 5.1: UT echo loss.

5.1.3 Operational Details

The operational details of the inspection run is listed in table 5.2.

Table 5.2: Pipeline And Medium Information 2010 Inspection Run

Inspection section	Ula - Ekofisk J
Length of inspected pipeline	73 km
Pipeline diameter	20"
Wall thickness range	12.7 mm – 20.5 mm (predominantly 14.3 mm)
Highest occurring weld type	Long seam
Minimum bend radius	5D/90°
Steel grade	API 5L X60/X65
MAOP	139.7 bar
Design factor	0.72
Conditions during normal operation:	
Pressure	Launcher: 29 bar, receiver: 20 bar
Temperature	Launcher: 72°C, receiver: 36°C
Medium	Oil
Medium flow velocity	0.5 m/s
Conditions during inspection:	
Pressure	Approx. 30 bar
Temperature	Launcher: 43°C, receiver: 8°C
Coupling medium	Crude oil
Sound velocity in coupling medium	1310 m/s at 20C
Attenuation in coupling medium	6 dB
Inspection speed	Approx. 0.45 m/s

Table 5.3: Reporting Thresholds.

Feature	**Description**	**Value**
Metal loss	Depth	$\geq$ 1.5 mm
	Length	$\geq$ 20 mm
Lamination	Length	$\geq$ 300 mm
	Width	$\geq$ 2 sensors
Wall thickness variation	Wall thickness reduction	$\geq$ 20%
	Length	$\geq$ 300 mm
Dent	Sensor carrier lift-off	$\geq$ 3 mm
Installation	All	

5.1.4 Data Thresholds

For a feature to be registered, the criteria in table 5.3 must be applied. The UT tool registers all pipe joints as well as detailed information about all features fulfilling the criteria which is identified during data analysis. The potential features recorded is listed in the pipeline register and presented in table 5.4. In figure 5.2, the pipeline register file structure is illustrated showing all the variables for each reported feature ID.

5.2 Assessment of Recorded Critical Anomalies

From the 2010 inspection report, there are three anomalies which does not fulfill the DNV-RP-F101 criterion for single defect method's conservative estimates. The highest calculated ERF

Table 5.4: Registered features.

Feature	**Description**
Feature ID	Unique identifier of the feature
GW No.	Girth weld number
Absolute distance [m]	Distance value of beginning of feature to reference distance 0 (start point of inspection)
Dist. to USW [m]	Distance value of beginning of feature to upstream girth weld
Circ. pos. [$\circ$]	Circumferential position of feature (0 degrees at 12 o'clock)
Ref. wt [mm]	Reference wall thickness in vicinity of feature
Rem. wt [mm]	Remaining wall thickness of feature
Depth [mm]	Depth of feature
Depth [%]	Depth of feature in percent of Ref. wt
Joint length [m]	Length of pipe joint
Length [mm]	Axial length of feature
Width [mm]	Width of feature
Psafe [bar]	Maximum safe pressure specified by client
ERF value	Estimated repair factor
Feature type	Geometry, installation, dent, inclusion, lamination, metal loss, wall thickness variation
Comment	Additional information about feature
Rel. pos.	Relative position of feature. bm = base material, alw = at long weld, agw = at girth weld, asw = at spiral weld, aws = at welds, ilw = in long weld, igw = in girth weld, isw = in spiral weld
Loc. in wall	Location of feature with pipe wall: Internal, external, or mid-wall

Feature ID	GW No.	Absolute distance [m]	Dist. to USW [m]	Circ. pos. [°]	Ref. wt [mm]	Rem. wt [mm]	Depth [mm]	Depth [%]	Joint length [m]	Length [mm]	Width [mm]	P_{safe} [bar]	ERF value	Feature type	Subtype	Comment	Rel. pos.	Loc. in wall	Dist. DP [m]	Circ. pos. DP [°]
900094	0.0	0.17	0.00	2	30.3					55	51			installation	fitting					8
900097	0.0	0.17	0.00	192	30.3	28.0	2.3	8		341	56			metal loss	channeling		bm	int	0.38	184
900093	0.0	0.49	0.00	310	30.3	26.3	4.0	13		37	179			metal loss			agw	int	0.52	280
GW	1.0	0.53							2.39							start of inspection				
900096	1.0	0.53	0.00	194	30.4	27.1	3.3	11		2382	102	139.7	0.38	metal loss	channeling	in pipe	bm	int	0.55	185
900099	1.0	0.54	0.00	169	30.4	28.0	2.4	8		27	379	139.7	0.34	metal loss			agw	int	0.54	150
900098	1.0	0.54	0.00	189	30.4	27.1	3.3	11		61	66			main spot		ref. box ID: 900096	agw	int	0.55	185
900095	1.0	2.62	2.09	184	30.3					284	670			installation	support					260
GW	2.0	2.92							0.77											
100001	2.0	2.92	0.00	190	30.3	27.1	3.2	11		741	69	139.7	0.37	metal loss	channeling	in pipe	bm	int	2.95	182
900385	2.0	2.92	0.00	13	30.3	27.2	3.1	10		55	97	139.7	0.34	metal loss			agw	int	2.93	359
900374	2.0	2.93	0.00	171	30.3	29.0	1.3	4		54	552	139.7	0.34	metal loss			agw	int	2.96	164
900100	2.0	2.93	0.01	187	30.3	27.1	3.2	11		61	41			main spot		ref. box ID: 100001	bm	int	2.95	182
100002	2.0	2.98	0.06	13	30.2	27.8	2.4	8		704	82	139.7	0.36	metal loss			agw	int	3.10	7
GW	3.0	3.69							0.76											
100003	3.0	3.70	0.00	197	30.3	26.2	4.1	14		106	119	139.7	0.35	metal loss	channeling		agw	int	3.70	182
900101	3.0	3.70	0.00	163	30.3	28.0	2.3	8		30	250	139.7	0.34	metal loss			agw	int	3.70	139
900102	3.0	3.84	0.15	112	30.3	27.8	2.5	8		558	307	139.7	0.36	metal loss			bm	int	4.34	90
100004	3.0	3.87	0.18	182	30.6					100	102			installation	fitting					194
100005	3.0	3.98	0.28	201	30.6	25.7	4.9	16		244	138	139.7	0.36	metal loss	channeling		bm	int	4.19	182
100006	3.0	4.11	0.41	3	30.6					54	52			installation	fitting					8
100007	3.0	4.25	0.56	181	30.6					49	52			installation	fitting					187
100008	3.0	4.28	0.59	196	30.6	26.3	4.3	14		225	132	139.7	0.35	metal loss	channeling	across girth weld	agw	int	4.45	187

Figure 5.2: Pipeline register.

values are in the range 1.01 to 1.21, located in three pipe joints: GW No. 353.0, 354.0, and 355.0. The anomalies are detected as channeling corrosion features with lengths equal to the length of the pipe spools. Hence, why the DNV method for single defects are not sufficient as the features would involve complex depth profiles. The DNV single defect method only take into account the maximum corrosion depth and total length of the feature. By using the complex defect assessment of DNV-RP-F101, the detailed depth profile of the corrosion feature would be taken into account and result in more accurate results with relatively less conservative estimates for safe operating pressure.

The datasets will be further investigated in the next chapter.

6 Machine Learning Case

6.1 Introduction

In this section, the objective is to showcase the comprehensive workflow involved when building predictive models by using machine learning tools from the Python library scikit-learn. The workflow consists of the following steps:

1. **Problem Definition:** The problem definition is important in order to know how to apply the appropriate solution to problem at hand. This involves defining clearly stating the problem and what we want to predict, aiding in identifying the suitable data with relevant value for solving the problem.

2. **Data Collection:** Collecting the necessary data of high value for the problem. This involves knowing how and where to access the necessary data, and how to get value from it in order to make it useful for the problem.

3. **Data Preprocessing:** As the data is usually incomplete, unstructured, inconsistent, and may include many errors and outliers, it is important to preprocess the data such that the data becomes cleaned, structured, and converted into suitable data types for the different algorithms.

4. **Exploratory Data Analysis (EDA):** EDA is a data analytics technique for investigating the data by using graphical statistics in order to discover patterns, identify anomalies, and check assumptions for the data.

5. **Data Modeling:** This involves the experimental phase of machine learning where we implement different learning models for our data in order to select the most suitable learning model which results in the best predictive performance based on the available data.

6. **Deployment of Model:** After comparing the different models, finding the best one, optimizing and tuning the hyperparameters, and achieving the optimal model, it is finally time for deploying the predictive model. Now we can see the results for how great the predictive power is of the model.

The client case is performed with Python and the machine learning library scikit-learn which is a great library for efficiently deploying ML models. In the following sections, the the results from the different steps in the data analysis will be summarized and explained in detail. Results are extracted from a separate Python kernel[11]. Tables from the notebook are listed in this chapter as figures due to conversion issues.

[11] A 'kernel' is a program that runs and introspects the user's code.

6.2 Problem Definition

The objective of this case is to perform an analysis of the ILI inspection data with application of machine learning techniques in order to predict classification of features recorded by the ILI tool. The classification will indicate the level of integrity of the different features of the pipeline system and be classified as either Class 0 (insignificant corrosion) or Class 1 (significant corrosion).

6.3 Data Collection

6.3.1 Importing Data

In this section the ILI data is imported to the notebook and stored in a dataframe using the pandas library in python. The data used is the reported features from the ILI inspection, and the shape of the dataset is 1791 rows and 21 columns. The rows make up each recorded pipeline feature with an individual ID tag, and the rows make up the variables for each feature. The data was split into two separate datasets - one for training and one for testing. The objective is to train the ML models on the training data, and after learning is performed the test data is used as "unseen" data in order to validate and achieve an accuracy score for the predictive models. The data was randomly split with the training set having a shape of 1000 x 21 and the test set a shape of 790 x 20. This is because the Class labeling was removed from the test set as the case is a classification problem.

6.3.2 Describing Features

Each column gives a certain feature information about each sample which can be illustrated in figure 6.1. The features can be divided into 4 categories:

- **Nominal categorical features:** Features with two or more features. No intrinsic ordering of categories.
- **Ordinal categorical features:** Features with similar characteristics as nominal categorical features, but differentiated by clear ordering of categories.
- **Integer numerical features:** Features that must take an integer value.
- **Continuous numerical features:** Features that can be infinitely many values.

The Class label is the outcome feature for the problem set and consists of binary data where 0 is high level of integrity and little corrosion, and 1 is lower integrity and significant corrosion. The class label was added to the initial raw data supplied by the ILI inspection. Due to the dataset containing generally few severe instances of corrosion, Class 1 was set to cover a lower range of values for feature depth than originally planned as the regular threshold applied when inspecting

	Description	Values	Number of unique values
FeatureID	Feature ID	[100001, 100002, 100003, 100004, 100005, 100006, 100007, 100008, 100009, 100...	1000
Class	Class label	[1.0, 0.0]	2
GWNo	Girth weld number	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 12.0, 13.0, 15.0, 16.0,...	321
AbsDist	Absolute distance from start point	[2.92, 2.98, 3.7, 3.87, 3.98, 4.11, 4.25, 4.28, 5.4, 6.54, 7.78, 8.78, 8.79,...	976
DistToUSW	Distance to upstream girth weld	[0.0, 0.06, 0.18, 0.28, 0.41, 0.56, 0.59, nan, 0.24, 0.01, 0.07, 0.34, 0.66,...	565
CircPos	Circular position of feature	[190.0, 13.0, 197.0, 182.0, 201.0, 3.0, 181.0, 196.0, nan, 284.0, 74.0, 172....	242
RefWT	Reference wall thickness	[30.3, 30.2, 30.6, nan, 26.3, 16.0, 15.7, 27.6, 19.4, 22.2, 22.4, 21.4, 19.3...	48
RemWT	Remaining wall thickness	[27.1, 27.8, 26.2, nan, 25.7, 26.3, 14.0, 13.8, 14.3, 17.6, 17.0, 17.8, 17.2...	76
Depth	Depth of feature	[3.2, 2.4, 4.1, nan, 4.9, 4.3, 2.0, 2.2, 1.4, 1.7, 1.8, 1.6, 2.8, 1.0, 1.2, ...	43
DepthPercentage	Depth percentage of feature	[10.6, 7.9, 13.5, nan, 16.0, 14.1, 12.5, 13.8, 8.9, 9.3, 12.4, 8.2, 11.3, 10...	121
Length	Length of feature	[741.0, 704.0, 106.0, 100.0, 244.0, 54.0, 49.0, 225.0, 1789.0, 1981.0, 48.0,...	319
Width	Width of feature	[69.0, 82.0, 119.0, 102.0, 138.0, 52.0, 132.0, nan, 47.0, 379.0, 333.0, 159....	184
Psafe	Safe pressure	[139.7, nan]	2
ERFvalue	Estimated repair factor value	[0.37, 0.36, 0.35, nan, 0.68, 0.55, 0.56, 0.47, 0.58, 0.59, 0.61, 0.57, 0.63...	38
FeatureType	Feature type	[metal loss, installation, main spot, geometry]	4
Subtype	Feature subtype	[channeling, nan, fitting, launcher valve, valve, T-piece, hot bend, flange,...	16
Comment	Feature comment	[in pipe, nan, across girth weld, ref. box ID: 900105, sphere-tee, echo loss...	211
RelPos	Relative position of feature	[bm, agw, nan, alw, aws]	5
LocInWall	Location in wall	[int, nan, ext]	3
DistDP	Distance to deepest point of feature	[2.95, 3.1, 3.7, nan, 4.19, 4.45, 8.8, 8.81, 8.89, 9.17, 32.37, 32.81, 32.93...	711
CircPosDP	Circular position of deepest point of feature	[182.3, 6.9, 193.9, 8.4, 187.1, 186.9, 359.7, 186.1, 0.9, 257.3, 31.2, 155.8...	376

Figure 6.1: Feature Descriptions.

corrosion features is set to 2 mm for the initial inspection assessment. This was done to achieve enough data for both classes in order to train the ML models sufficiently. The ideal situation would be to have available various pipeline ILIs and train the algorithms on both uncorroded sections and highly corroded sections where integrity is severely reduced. This data was not available for the client at the time of this thesis. Hence, the best alternative was decided after coordination with the client to set the threshold for significantly corroded features to 10 percent of the pipeline wall thickness.

6.4 Data Preprocessing

This section involves the preprocessing of data such as cleaning, structuring, and enriching the raw data from the ILI inspection into another format. The objective is to increase the value of the data by making it more appropriate for data analytics and converting it into a format which the ML algorithms can handle efficiently. The process for preparing data for training consists of 4 steps:

1. **Correcting:** Correcting values, handling outliers and removing features. The process of rectifying abnormal, accurate or non-acceptable values is of high importance as they can affect the predictive power and accuracy of the ML models.

2. **Complementing:** The process of handling missing data.

3. **Creating:** Creating new features such as aggregate values, etc.

4. **Converting:** Converting data into data types which the algorithms can efficiently handle.

The essential factors determining how well a ML algorithm can learn and perform is the quality of the available data and the relevant information that it contains (Raschka and Mirjalili, 2017). Hence, the high importance of preprocessing the data before the learning algorithm is applied. The following section will cover and discuss essential techniques for preprocessing the ILI-data which will result in the desired level of performance for the ML models.

6.4.1 Identifying Anomalies and Outliers

These values can affect the accuracy of ML models if not correctly handled. This involves investigating for potential extreme outliers, etc. To identify such features we can use the univariate distribution of each feature. We will first look at the univariate distribution of the categorical features.

6.4.1.1 Univariate Distributions of Categorical Features

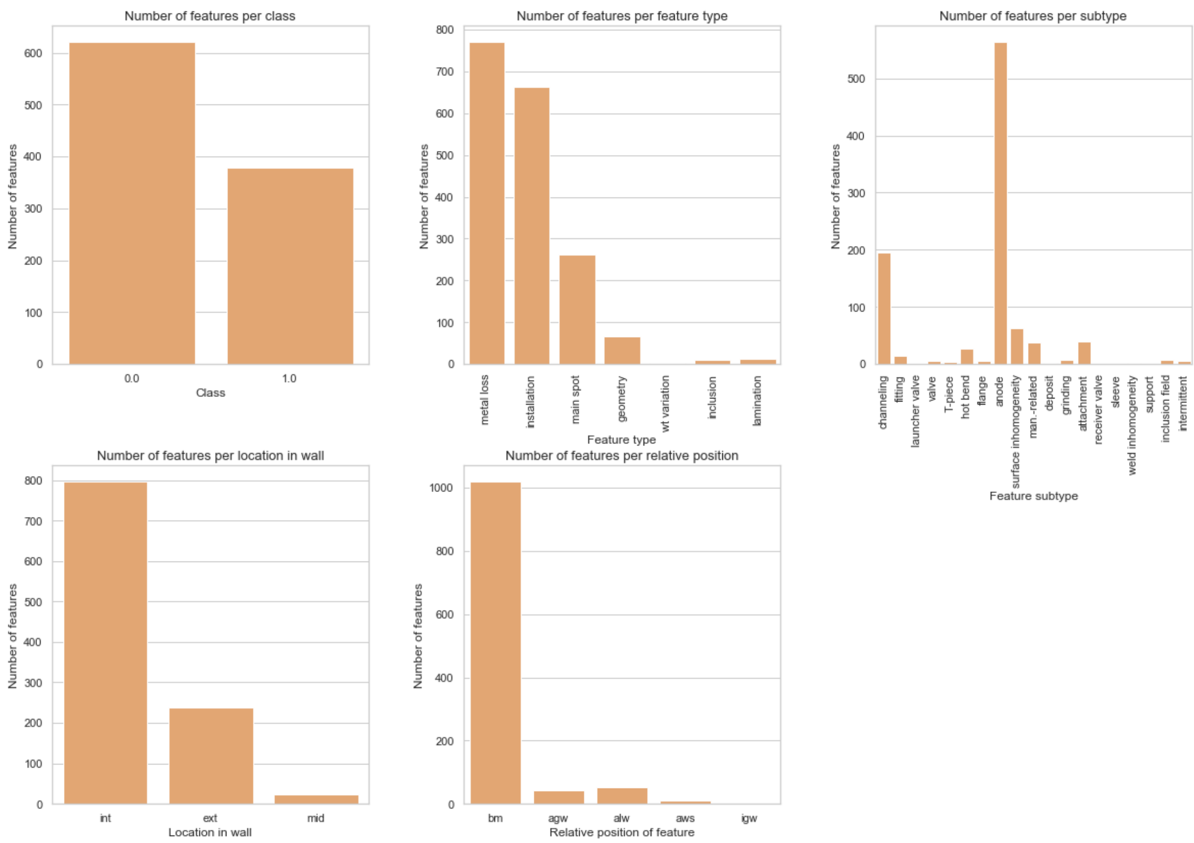


Figure 6.2: Univariate distributions of categorical features.

Observations:

- The number of features that is Class 0 is significantly higher than Class 1.
- The feature types that is dominating is metal loss, installation and main spot. Metal loss and main spot are basically the same with main spot being the most significant corrosion feature in a cluster of corrosion features making up a complex shaped defect.
- The feature subtypes occurring most frequently is anode and channeling.
- Most features are located internally, but there is still a considerable amount of external features.
- Features located in the middle of the wall are usually considered as inclusions and defects from the production of the pipeline.
- The majority of features for relative position is in the base material.

6.4.1.2 Univariate Distributions of Numerical Features

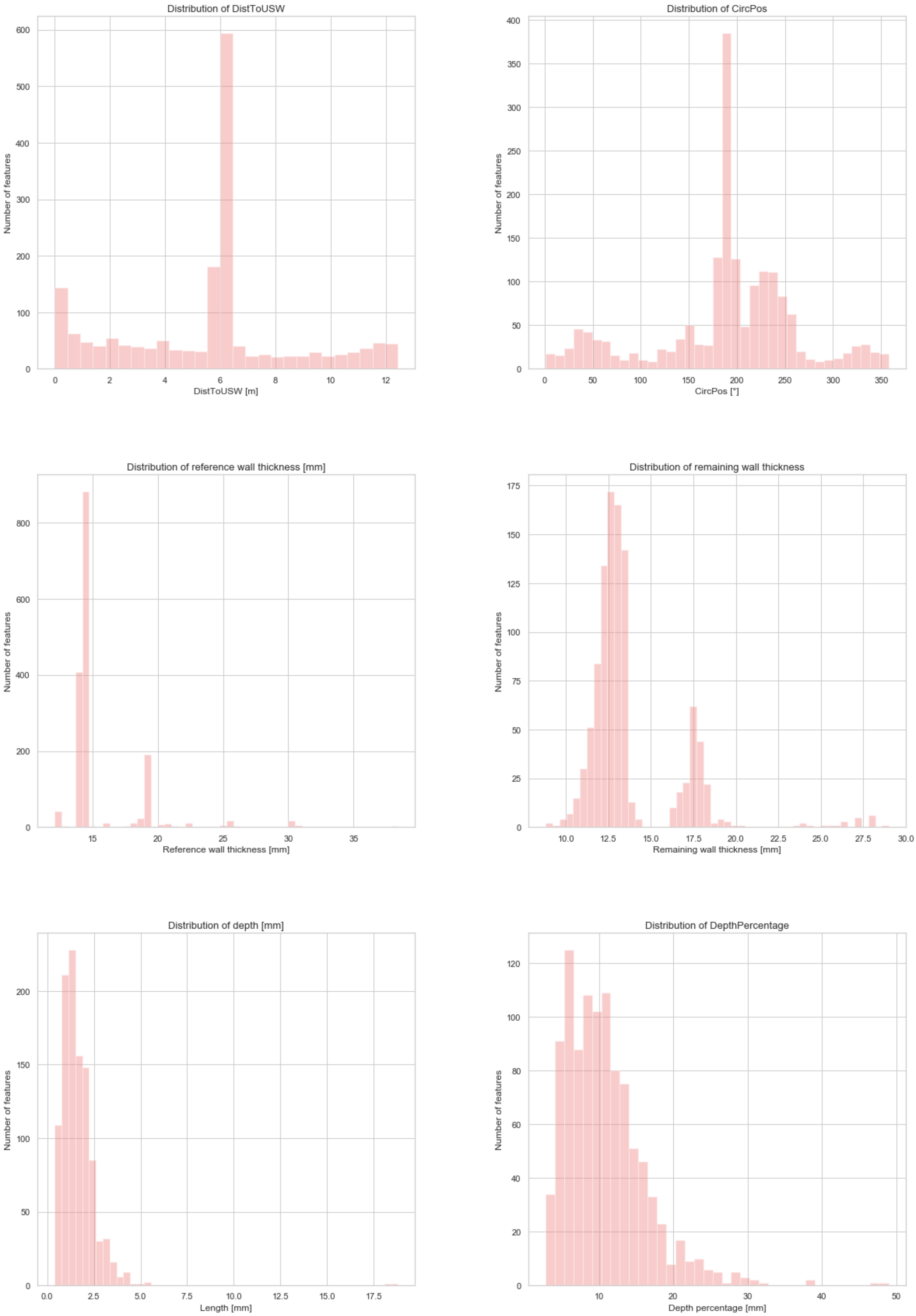


Figure 6.3: Univariate distributions of numerical features.

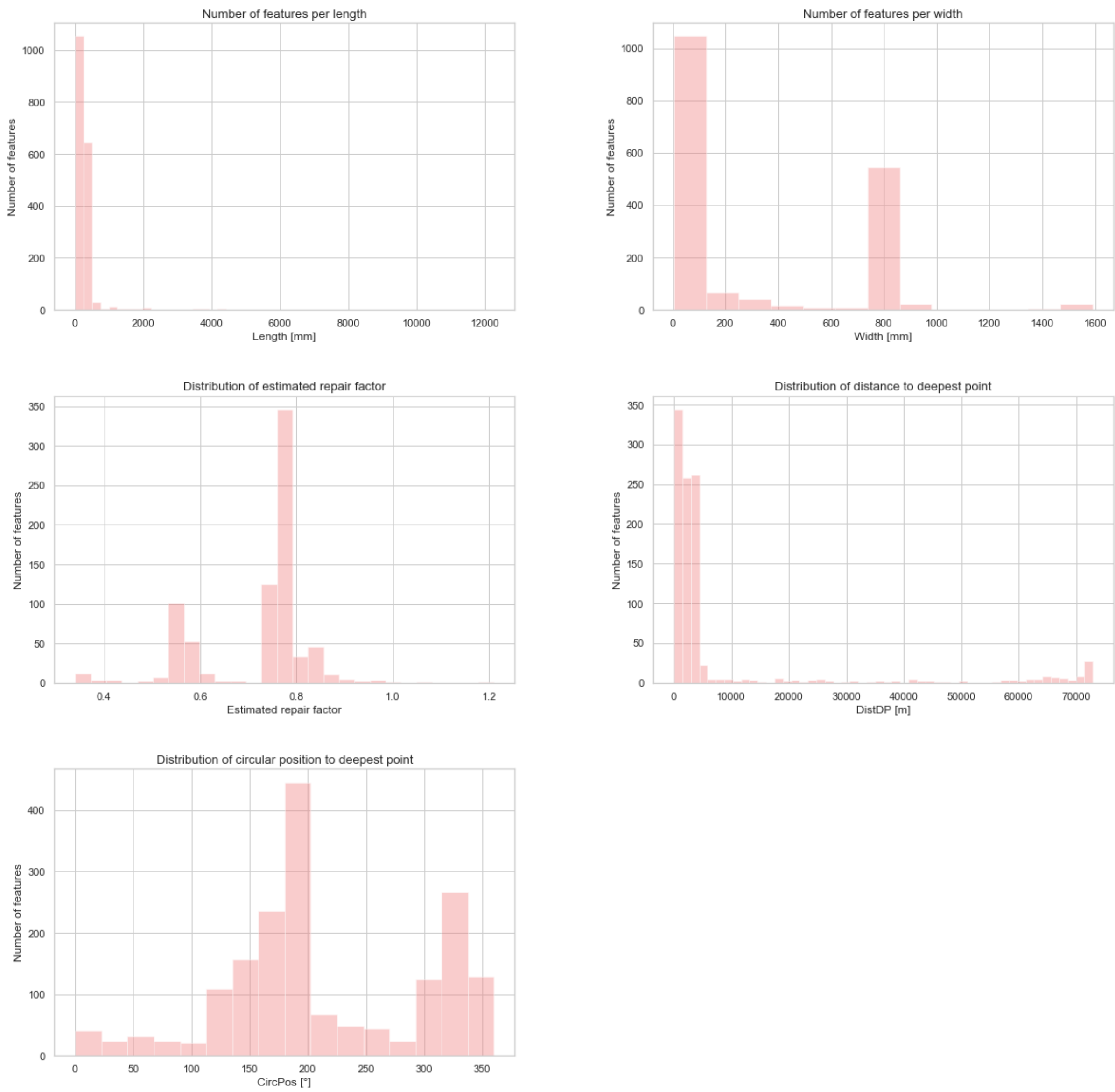


Figure 6.4: Univariate distributions of numerical features.

Observations:

- DistToUSW feature ranges from 0 to 12 meters with the majority of features occurring around 6 meters, indicating that the average pipe section is 12 meters. The majority of features occur around 6 meters, but there is a pretty uniform distribution for the rest of the range.
- The distribution of CircPos feature resembles normal distribution with the mean value being around 180 degrees. This is corresponding with the position of 6 o'clock in the pipeline where channeling occurs, often as a consequence of the multiphase flow and water assembling on the bottom part.
- RefWT feature ranges from 12.1 to 38.4 mm with the majority being centered just below 15 mm, corresponding with the majority of the pipeline type.

- RemWT feature ranges from 8.8 to 29.0 mm. The majority of features are centered at two local concentrations around 12.5 and 17.5 mm.
- Depth feature resembles a log-normal distribution with the majority of features occurring between 0 and 2.5 mm. Some outliers occur at up to 18.8 mm.
- DepthPercentage feature does also resemble a log-normal distribution with the majority being centered around 10 percent wall of the wall thickness.
- Length feature ranges from 0 to above 12000 mm in some cases. The majority of the features occur from 0 to 400 mm.
- Width feature ranges from 5 to 1591 mm, and the data is distributed around two local maximums at 50 and 800 mm width.
- ESFvalue feature resembles a normal distribution with the mean value just below 0.6.
- DistDP feature indicates that the majority of recorded deepest points occur at the first and last section of the pipeline.
- CircPosDP feature resembles a normal distribution similar to the CircPos feature, indicating that there is a high chance of collinearity between the two features, resulting in one of them being redundant. In other words, one of them is likely to be removed for the ML model.

Based on these observations, no value correction is necessary as they can be considered acceptable.

6.4.2 Outlier Data

An outlier is a point observed to be distant from the majority of the observations. Outliers can distort and affect the accuracy of ML models. In this section the box plot rule is applied. The rule states that for a given numerical variable, outliers are observations outside the 1.5 interquartile range (IQR) which is the difference between the 75th and 25th quartiles. The numerical features are plotted with outliers in 6.5.

Observations:

- **DistToUSW:** No values considered outliers.
- **CircPos:** Values below 120 and above 280 degrees is considered outliers.
- **RefWT:** values below 14 and above 15 is considered outliers.
- **RemWT:** Values below 10 and above 15 is considered outliers.
- **Depth:** Values above 3.6 is considered outliers.
- **DepthPercentage:** Values above 26 percent is considered outliers.
- **Length:** Values above 1000 is considered outliers.

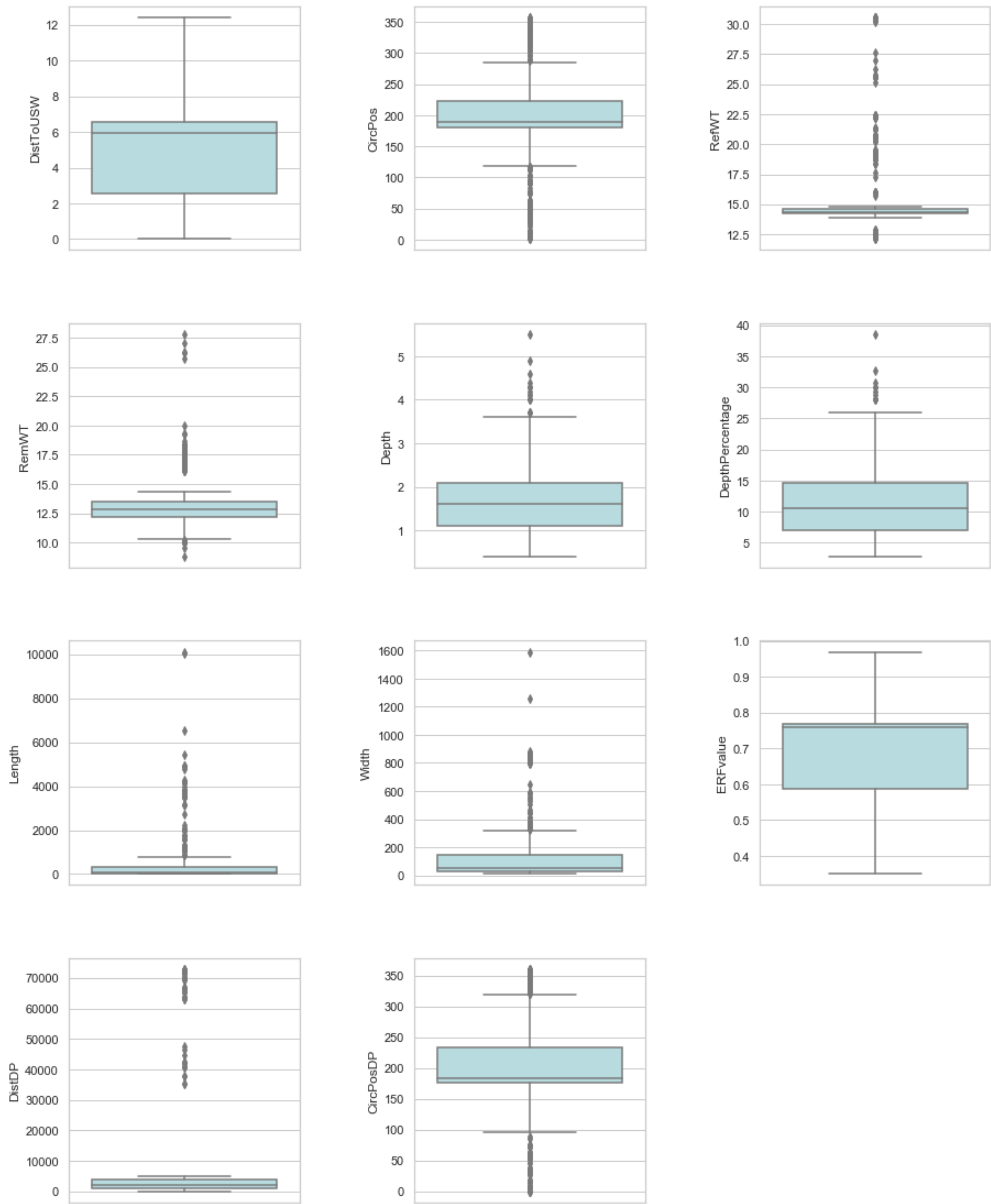


Figure 6.5: Box plots of numerical features with 1.5 IQR.

- **Width:** Values above 350 is considered outliers.
- **ERFvalue:** No values considered outliers.
- **DistDP:** Values above 5000 is considered outliers. This feature is just a modified absolute distance so it should be removed for further analysis as it would affect results by considering the majority of the pipeline to be outlier. Will be further inspected in the exploratory data analysis (EDA).
- **CircPosDP:** Values below 100 and above 320 is considered outliers. In other words, a greater range compared to CircPos feature.

Now the outliers are identified, the next step is to decide what to do with them. Hence, we will perform exploratory analysis of the features later in the thesis.

6.4.3 Missing Data

ML algorithms cannot handle missing values. Hence, it is important to identify features containing missing values in order to fix them. The number of missing values in the datasets are displayed in figure 6.6.

6.4.3.1 Identifying Missing Data in Training and Test Datasets

	Number of Missing Values (Training)	% of Missing Values (Training)	Number of Missing Values (Test)	% of Missing Values (Test)
AbsDist	0	0.0	0.0	0.00
CircPos	14	1.4	0.0	0.00
CircPosDP	0	0.0	0.0	0.00
Class	0	0.0	NaN	NaN
Comment	459	45.9	297.0	37.59
Depth	266	26.6	488.0	61.77
DepthPercentage	266	26.6	488.0	61.77
DistDP	268	26.8	492.0	62.28
DistToUSW	14	1.4	0.0	0.00
ERFvalue	516	51.6	505.0	63.92
FeatureType	0	0.0	0.0	0.00
GWNo	0	0.0	0.0	0.00
Length	0	0.0	0.0	0.00
LocInWall	266	26.6	467.0	59.11
Psafe	516	51.6	505.0	63.92
RefWT	23	2.3	108.0	13.67
RelPos	249	24.9	415.0	52.53
RemWT	266	26.6	488.0	61.77
Subtype	526	52.6	279.0	35.32
Width	14	1.4	0.0	0.00

Figure 6.6: Missing data.

Observations:

- Comment, Depth, DepthPercentage, DistDP, DistToUSW, ERFvalue, LocInWall, Psafe, RefWT, RelPos, RemWT, Subtype, and Width all have missing values.
- The different features have their natural reasons for missing values. Not all features are e.g. necessary to comment.
- As stated in the report, the first 500 meters of the pipeline is suffering from ecco-loss due to water in the pipeline. This has led to a high amount of missing values in this region.
- After the feature s with missing values are identified, we must appropriately handle the missing values to increase the value of the data.

The methodology for handling missing data is illustrated in figure 6.7.

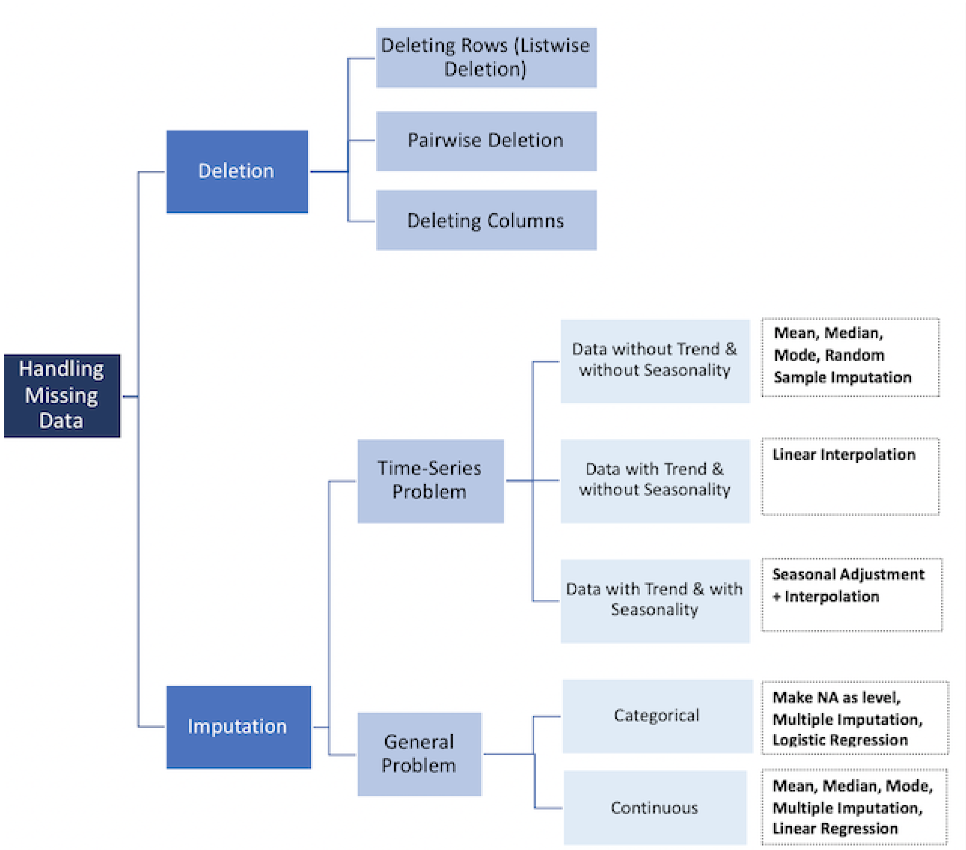


Figure 6.7: Procedure for handling missing data.

From the previous data the options are to either impute or delete the missing data. By selecting to delete data, there is a risk of loosing valuable information which the algorithms can use. If deletion is performed on a complete row, that would result in the loss of all variables for the classified feature. Hence, loosing valuable information which can aid in improving the prediction model.

Imputing values is an option for the numerical data as use of the feature correlation matrix enables the prediction of missing values. This can involve using ML techniques such as linear regression to predict a missing value by selecting two highly correlating features as predictors. Linear regression can provide good estimates assuming there is a correlation between the features. Other alternatives for imputing values are to insert the mean value of the column for the missing value.

6.4.3.2 Plotting Feature Correlation Matrix

There may be unknown and complex relationships between the features in the dataset. Hence, it is important to first and foremost identify the correlating features, but also quantify to which degree which feature correlate. This will aid the prediction of missing values, and to detect multicollinearity between features. Features can correlate for a number of reasons. The correlation between features can increase the value of the data for data analytic purposes as it can improve the understanding of the statistical relationships between features. The correlation value can range from positive to negative, indicating to which degree the features correlate.

- **Positive value:** Both features change in the same direction.
- **Neutral value:** No correlation between features.
- **Negative value:** Features change in the opposite direction.

Values close to -1 or 1 indicates high correlation between the features. Figure 6.8 shows the feature correlation matrix which can be used to identify possible features which can be used to predict the missing values for the different features.

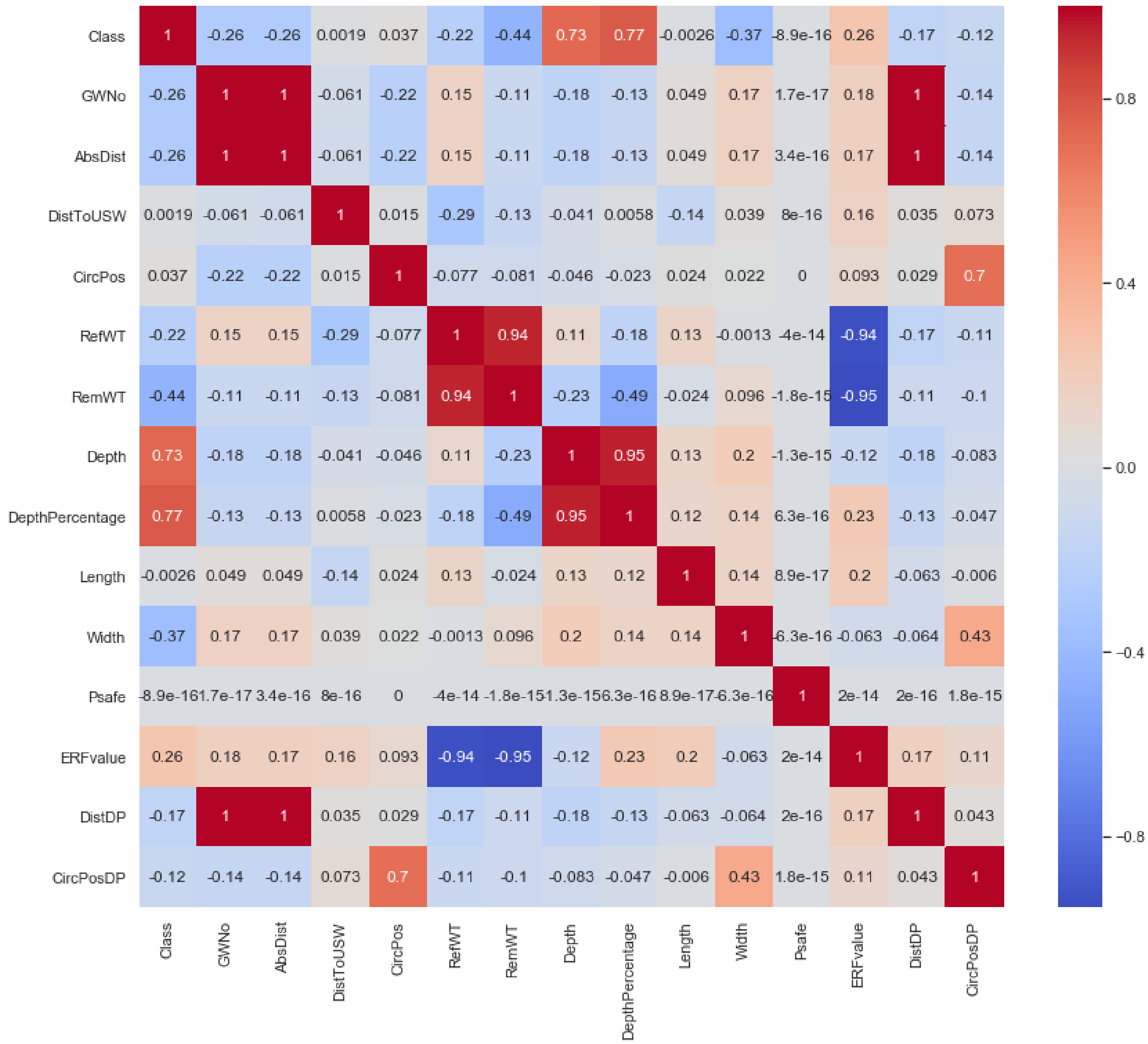


Figure 6.8: Feature correlation matrix.

Observations:

- **Class (not missing values):** Correlates highest with RemWT, Depth, DepthPercentage, Width.
- **GWNo:** Correlates highest with AbsDist, DistDP.
- **AbsDist:** Correlates highest with GWNo, DistDP.
- **DistToUSW:** Only significant correlation value is for RefWT (-0.29).
- **CircPos:** Correlates highest with CircPosDP. This is natural as the feature circumferential position of the deepest point should not vary much from the centre position of the selected feature.
- **RefWT:** Correlates highest with RemWT, ERFvalue.
- **RemWT:** Correlates highest with Class, RefWT, DepthPercentage, ERFvalue.

- **Depth:** Correlates highest with Class, DepthPercentage (less with RemWT, ERFvalue, DistDP).
- **DepthPercentage:** Correlates highest with Class, RemWT, Depth.
- **Length:** Correlates with no significant feature.
- **Width:** Correlates highest with Class, CircPosDP.
- **Psafe:** Correlates with no significant feature.
- **ERFvalue:** Correlates high with RefWT, RemWT.
- **DistDP:** Correlates highest with GWNo, AbsDist.
- **CircPosDP:** Correlates highest with CircPos, Width.

Now, we can, in additions to performing deletion, generate predictions for missing values for the features with highest correlation. This will be performed for some of the features over the coming sections.

6.4.3.3 Imputing Missing Values for CircPos by Linear Regression

Linear regression is a ML technique for predicting a value based on e.g. two predictor values that are known. The correlation matrix indicates that the CircPos feature correlates with CircPosDP and AbsDist feature. Hence, we make CircPosDP and AbsDist as the predictors in order to predict CircPos. The linear regression model is trained and then applied to the samples with missing feature data which then is filled in with a predicted value. Figure 6.9 shows the distribution of the CircPos feature after imputing the missing values. The distribution plots look OK with no negative values indicating that the linear regression model have predicted wrong. Hence, no need to investigate this further.

6.4.3.4 Imputing Missing Values for DistDP by Linear Regression

The correlation matrix indicates that DistDP correlates with GWNo and AbsDist feature. Hence, we select them as the predictors for the linear regression model. Figure 6.10 shows the distribution plot of the DistDP feature after imputation. The plots indicate that the highest number of features occur at the beginning of the pipeline, which does comply with the observations from the ILI where the first sections close to the installation were identified to experience the highest number of relatively severe corrosion features.

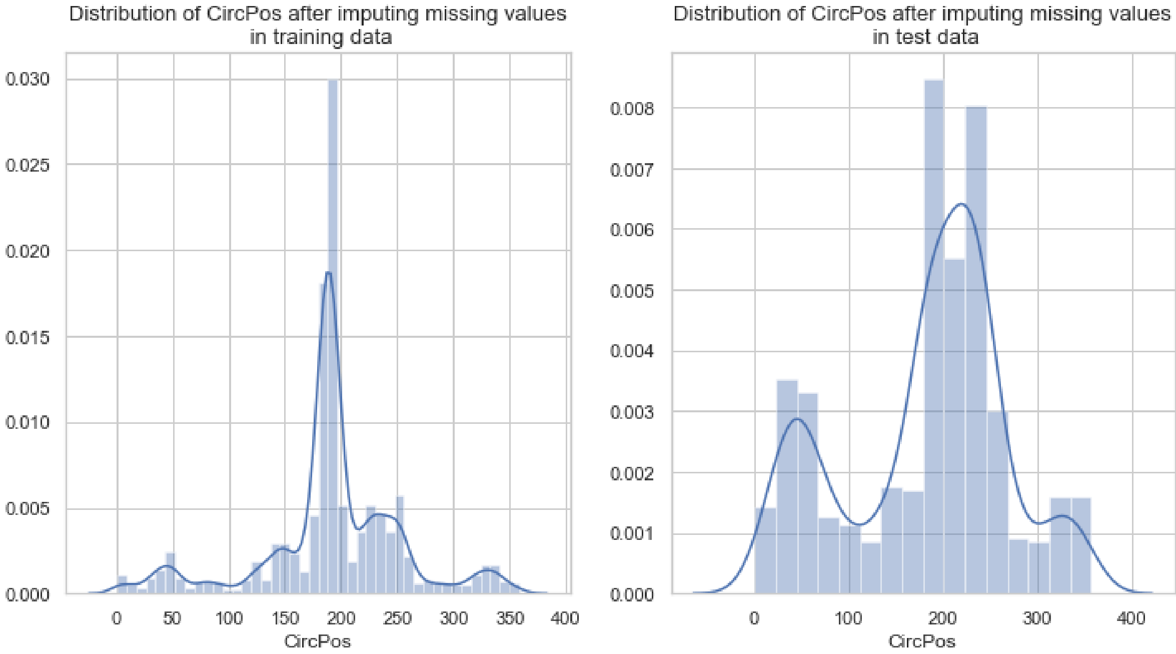


Figure 6.9: Distribution of CircPos after imputation.

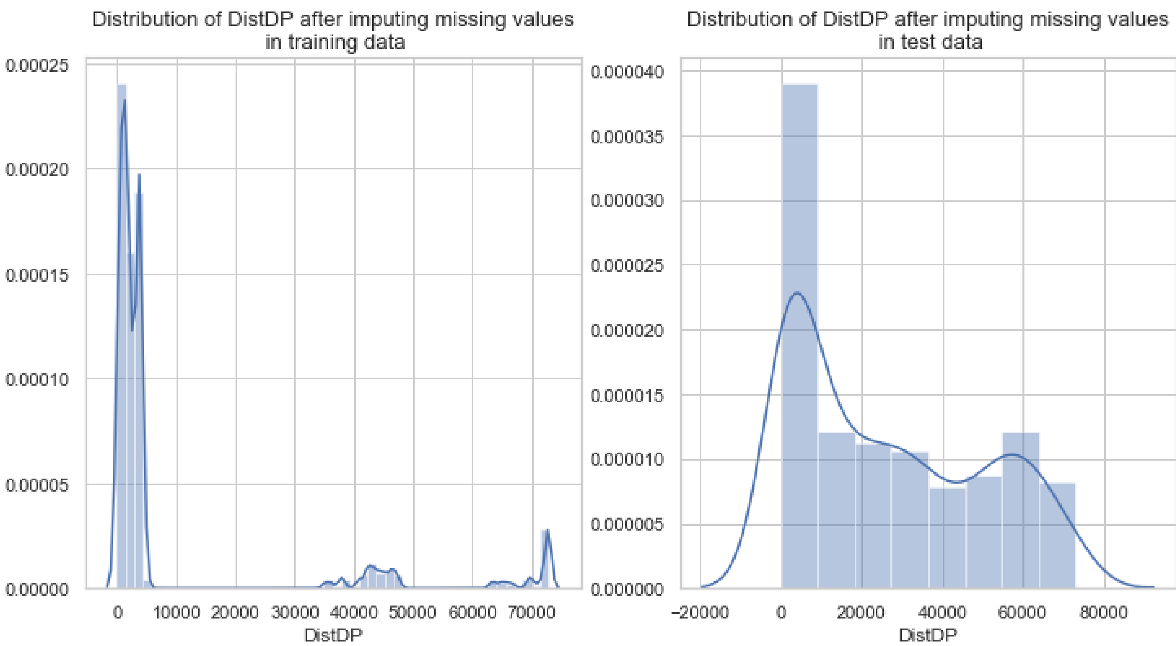


Figure 6.10: Distribution of DistDP after imputation.

	Number of Missing Values (Training)	% of Missing Values (Training)	Number of Missing Values (Test)	% of Missing Values (Test)
CircPos	0	0.0	0.0	0.00
CircPosDP	0	0.0	0.0	0.00
Class	0	0.0	NaN	NaN
Depth	0	0.0	0.0	0.00
DepthPercentage	0	0.0	0.0	0.00
DistDP	0	0.0	0.0	0.00
ERFvalue	250	25.0	17.0	2.15
FeatureType	0	0.0	0.0	0.00
Length	0	0.0	0.0	0.00
LocInWall	0	0.0	2.0	0.25
RefWT	0	0.0	0.0	0.00
RelPos	0	0.0	0.0	0.00
RemWT	0	0.0	0.0	0.00
Width	0	0.0	0.0	0.00

Figure 6.11: Remaining features.

6.4.3.5 Dropping Features of Less Importance

In the following section unnecessary features which can be argued that do not contribute to a possible solution will be removed as they are not longer necessary and only increases computational time and required power. In the next section, after the different features are removed, a threshold will be applied to remove samples with a certain amount of missing features. Features such as "installation" under "FeatureType" and "anode" under "SubType" have many missing values and can be considered to not contribute significantly towards the Class. From DNV-RP-F101 Part B, the major features affecting the Class is "Depth" and "Length" feature for single isolated defects. The selected features which are dropped are: GWNo, Psafe, Comment, AbsDist, DistToUSW, and SubType. All these features included a great amount of missing values.

For the next step a threshold of minimum 10 feature values is applied in order to remove rows containing less values. After this step, the following number of missing values are present in the datasets, illustrated in figure 6.11. This list indicates that the remaining missing value in the training set is the ERFvalue which will be estimated in the next section.

6.4.3.6 Imputing missing values for ERF

All the main spot features are missing the ERF value as the main spot is a part of a complex shaped defect. Though, we can predict the ERF value assuming we know Psafe and the MAOP. MAOP is set to be equal to the design pressure of 139.7 bar. To impute the missing values for ERF value, it is necessary to create a new feature column in our dataset called Pfail.

6.4.4 Creating New Features from Existing Features

New features can be created on the basis of existing features which can increase the value of the features towards contributing to the solution for the problem. From existing features we can create the following features:

- **Area:** By multiplying width and length of feature.
- **Volume:** By multiplying Area with depth of feature.
- **Failure pressure:** Calculation in next section.

It should be noted that this is a simplification of the feature dimensions as many of the defects in the pipeline are part of clusters making up complex shaped defects which require considerably more attention when performing a defect assessment. The simplifications are justified because the ILI indicates that the pipeline at hand is not significantly corroded. Hence, for ILI data from a heavily corroded pipeline it would be necessary to implement the methodology for assessing complex shaped defects in order to determine e.g. the failure pressure.

6.4.4.1 Creating Failure Pressure Feature

The methodology for calculating failure pressure done by performing the following series of calculations.

$$P_f = \frac{2 \cdot t \cdot UTS}{(D - t =)} \cdot \frac{\left(1 - \frac{d}{t}\right)}{\left(1 - \frac{d}{t \cdot Q}\right)} \tag{6.1}$$

where Q is

$$Q = \sqrt{1 + 0.31 \left(\frac{1}{\sqrt{Dt}}\right)^2} \tag{6.2}$$

The safe operating pressure is then calculated from the following equation:

$$P_{safe} = F \cdot P_f \tag{6.3}$$

where

$$F = F_1 + F_2 \tag{6.4}$$

From the safe operating pressure we can calculate the estimated repair factor (ERF):

$$ERF = \frac{MAOP}{P_{safe}} \tag{6.5}$$

If the client does not provide the MAOP, it is set to the design pressure which can be calculated as follows:

$$P_d = \frac{2t}{D} \cdot SMYS \cdot f_d \tag{6.6}$$

Table 6.1: Variables

MAOP	=	Maximum Operating Pressure	139.7 bar
t	=	Wall thickness	RemWT feature
d	=	Depth of metal loss feature	Depth feature
L	=	Length of metal loss feature	Length feature
D	=	Nominal outer diameter	508 mm (for the majority of the pipeline
F1	=	Modeling factor	0.9
F1	=	Operational Usage Factor	0.72
UTS	=	Ultimate Tensile Strength	

Table 6.2: UTS corrected for operational temperature.

Pipeline type	X60	X65
Standard UTS for pipeline type	413 MPa	448 MPa
Temperature corrected UTS	393 MPa	428 MPa

In this case, the MAOP is defined to be 139.7 bar as stated previously. Hence, we have our ERFvalue feature. The variables and the existing features in our dataset is listed in 6.1. For strength calculation according to DNV-RP-F101, the UTS value must be reduced to account for the design operating temperature of 78 degrees Celsius. The new UTS value is listen in table 6.2. As most of the pipeline consists of X60, we select 393 MPa as the UTS value for estimating the ERFvalue as this is the most conservative solution.

6.4.5 Converting Data Types

Most ML algorithms require all features to be converted into some numeric representations that can be understood by these algorithms. Categorical features cannot be processed directly by the algorithms and conversions on these features shall be done before subsequent data processing. Hence, by identifying the data type of each feature we can take the appropriate means to convert the data.

The dataset contains 12 features that have float data type and 3 features which have object datatype. FeatureType is considered a categorical feature as there are seven types of features, RelPos consists of 5 subtypes, and LocInWall consists of 3 types. Hence, the features cna be considered ordinal categorical as they can be set the values of 0, 1, 2, 3, and 4.

- **FeatureType:** metal loss = 1, main spot = 2.
- **LocInWall:** int = 0, ext = 1.
- **RelPos:** bm = 0, aqw = 1, alw = 2, aws = 3, iqw = 4.

6.5 Exploratory Data Analysis

Exploratory Data Analysis (EDA) is an approach for performing initial investigations on the dataset. EDA has a descriptive nature, using graphical statistics in order to recognize patterns, to detect anomalies, to test hypothesis, and check assumptions. The objective of this section is to perform EDA for several of the features in the dataset in order to find patterns and correlations between independent features and the class label. Some of the features are binned in order to create new features to increase the correlations and the predictive power of the learning algorithms.

First, we perform EDA on the categorical features vs the probability of Class. Then, we investigate the numerical features vs the probability of Class. Not all numerical features will be fully explored as earlier iterations of this Python kernel found the potential binning of features, etc. to have little impact on the final results as the algorithms easily classify the labels with high accuracy. This will be illustrated later in the thesis.

It is important to note that the EDA is performed on the training dataset as it has the largest amount of features. If the decision is taken to create new features by binning certain ranges of values, this will be performed for both training and test dataset.

6.5.1 Feature Type EDA

For FeatureType feature, we visualize the following:

- FeatureType vs Class
- FeatureType and LocInWall vs Class
- FeatureType and RelPos vs Class

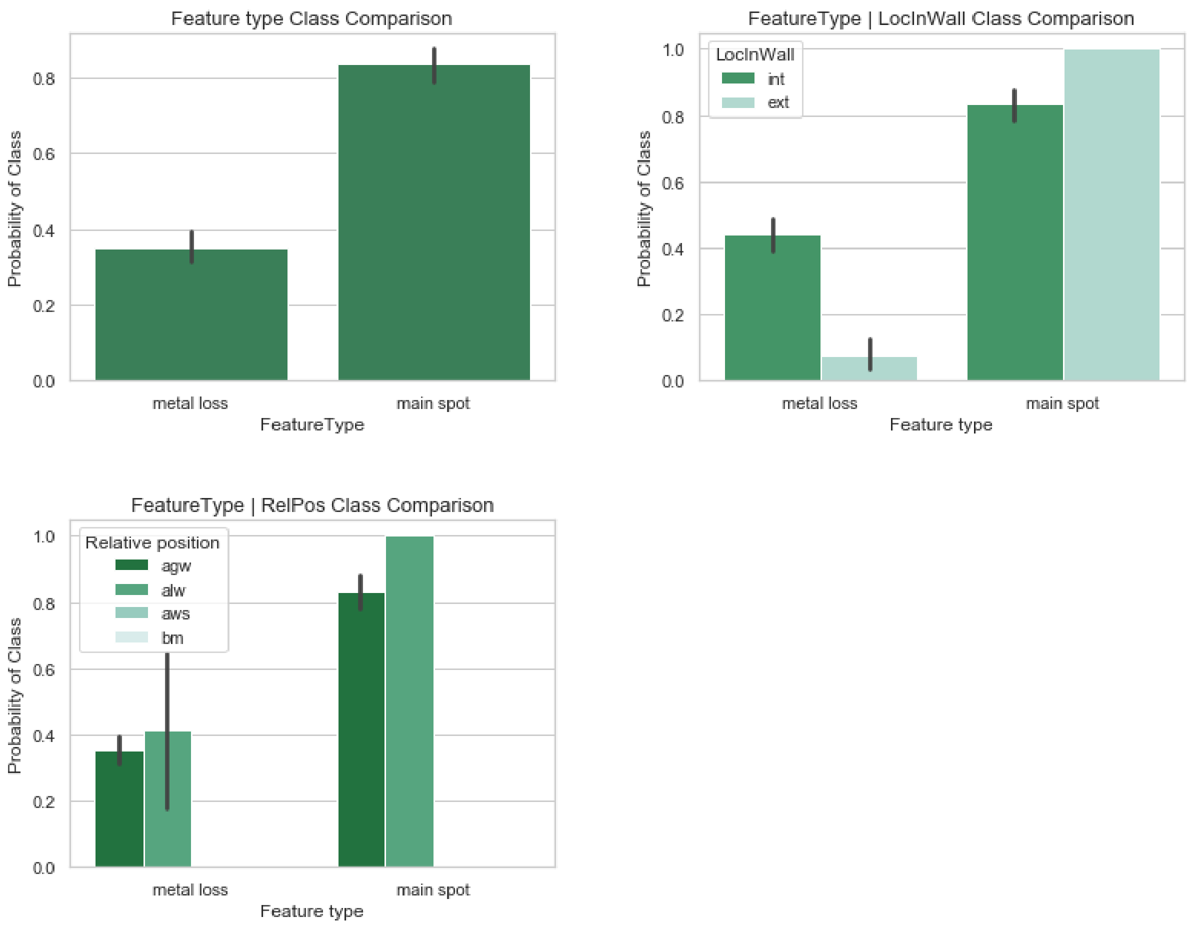


Figure 6.12: Feature type EDA

Observations:

- Main spot features have a higher chance of being classified as Class 1. This makes sense as main spot indicates the point of greatest depth of a complex shaped corrosion feature, which again is synonymous with a significantly corroded section of the pipeline.
- External features has a significantly higher chance of being classified as Class 1 if being a main spot. This does also make sense as there are 4 recorded external main spots which of all is of significant size. Hence, the high probability of being Class 1.
- For RelPos feature, main spot has a higher chance of being class 1 compared with metal

loss.

- RelPos 'bm' and 'aws' is has 0 percent chance of being classified as class 1 as indicated by the plot.

Based on the observations, we can conclude that FeatureType is an important feature for predicting Class.

6.5.2 Location in Wall EDA

For LocInWall feature, we visualize the following:

- LocInWall vs Class
- LocInWall and FeatureType vs Class
- LocInWall and RelPos vs Class

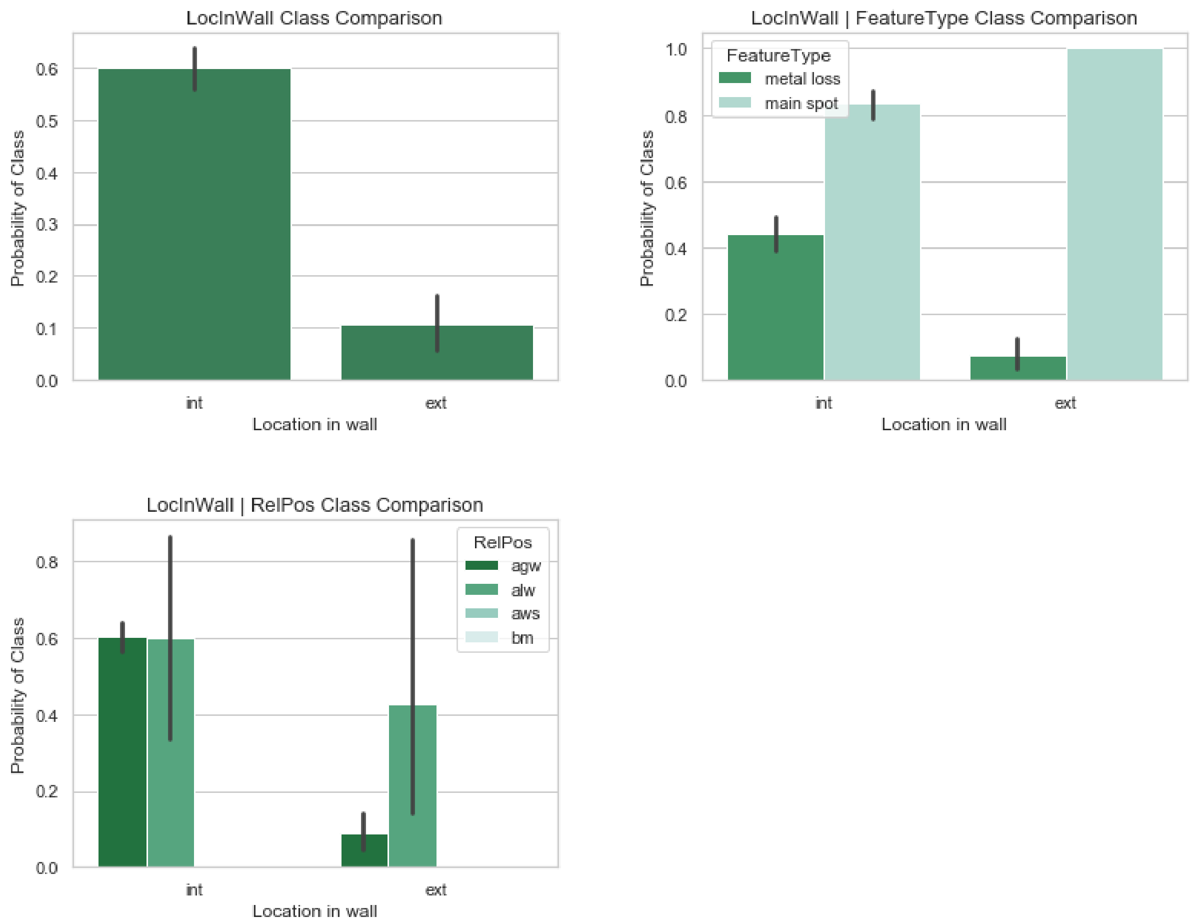


Figure 6.13: Location in wall EDA

Observations:

- Internal features have a significantly higher probability of being classified as Class 1.

- If an external feature is main spot, it is almost a guarantied Class 1 feature.
- External metal loss features have a very low probability of being Class 1 feature.

6.5.3 Relative Position EDA

For RelPos feature, we visualize the following:

- RelPos vs Class
- RelPos and FeatureType vs Class
- RelPos amd LocInWall vs Class

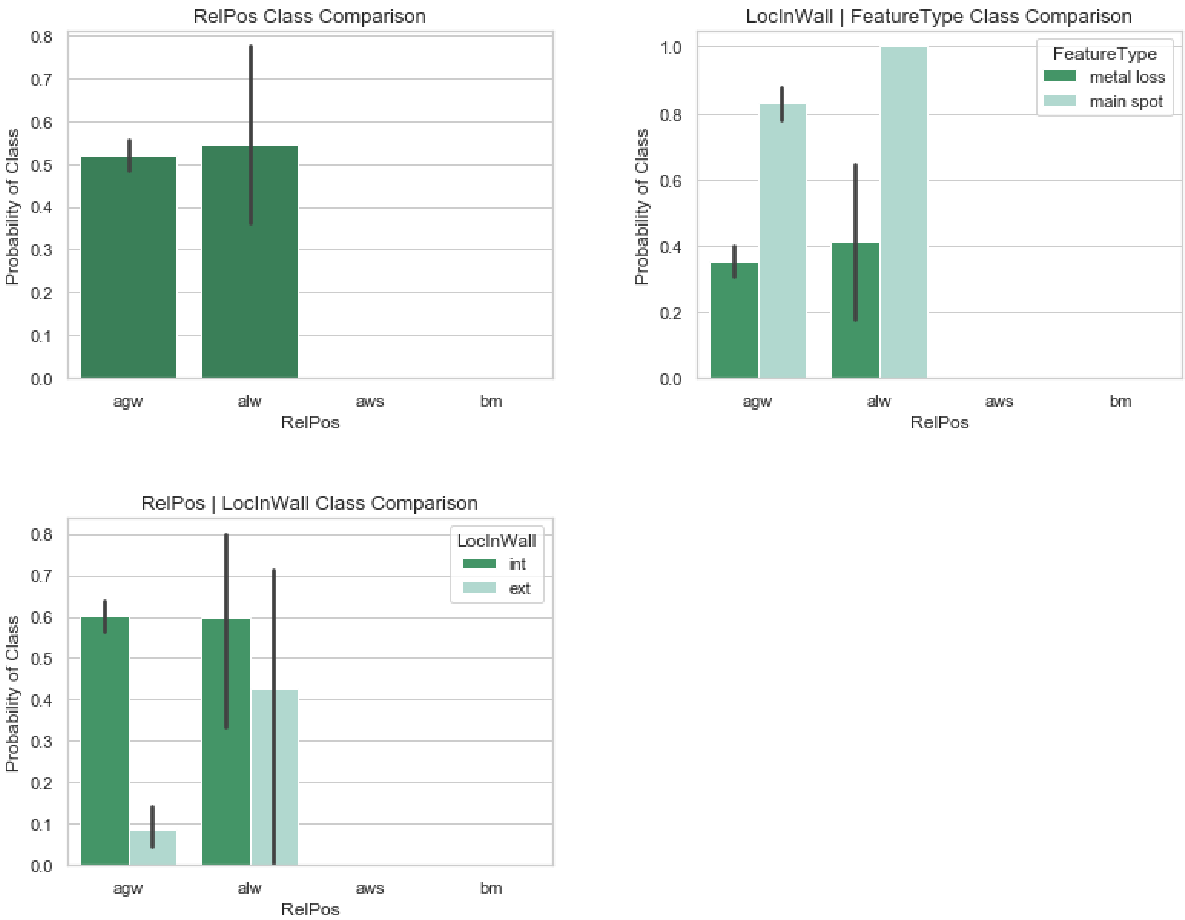


Figure 6.14: Relative position EDA

Observations:

- 50 percent chance for agw and alw to be Class 1 and 0.
- FeatureType has a significant impact on the type Class.
- LockInWall does also have a significant impact on Class as internal features have a significantly higher probability.

6.5.4 Circumferential Position EDA

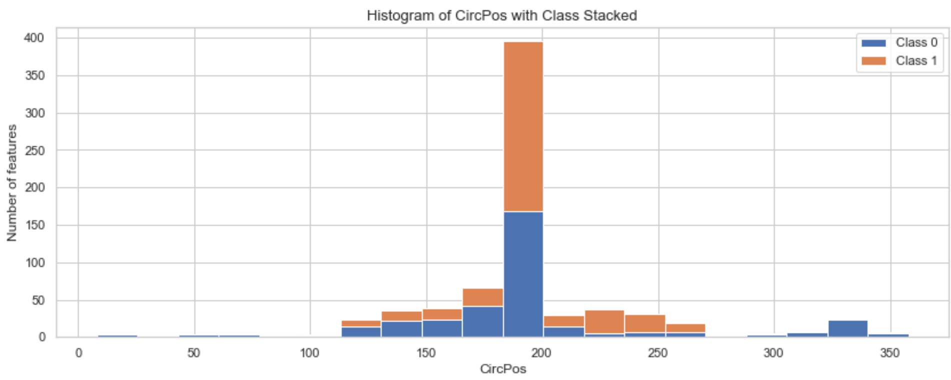


Figure 6.15: Circumferential position EDA.

Observations:

- CircPos resembles a normal distribution.
- Values are distributed with the mean value close to 180 degrees.

In order to increase the predictive power of the CircPos feature, we bin the feature into a new feature named CircPos_binned. The binning is performed into 5 groups illustrated in figure 6.16.

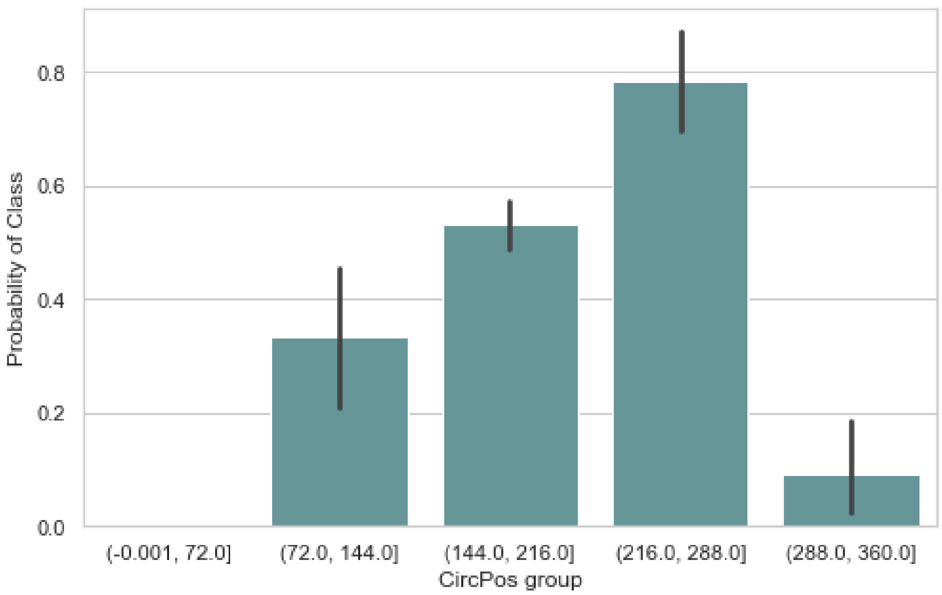


Figure 6.16: Circumferential position binned.

There seems to be a clear correlation between the CircPos feature and Class probability after the feature has been sampled. A consequence of binning numerical features are the subsequent creation of categorical features. Hence, it is important to remember to convert all binned features into integers.

6.5.5 Depth EDA

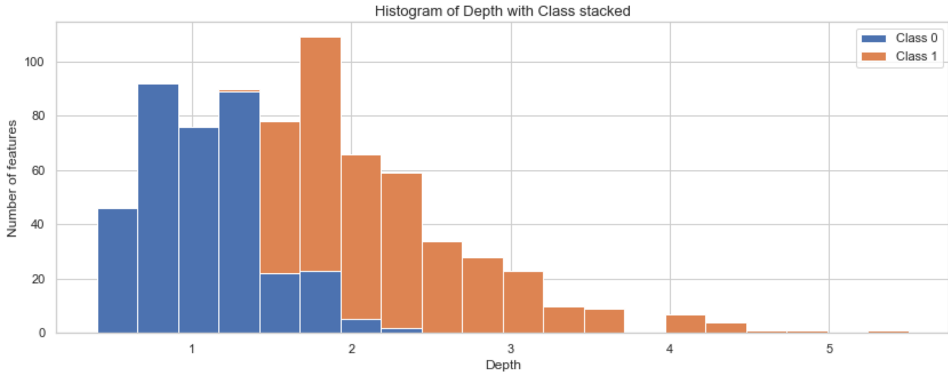


Figure 6.17: Depth EDA

Observations:

- The majority of depth data is located between 0.5 mm and 3.0 mm with some outliers.
- Sharp change at approximately 1.5 mm from class Class 0 to Class 1 features.
- Histogram resembles a normal distribution with the mean number of features occurring between 1.5 mm and 2.0 mm.

The decision is taken to group the data into a new feature, Depth_binned, consisting of 5 groups illustrated in figure 6.18.

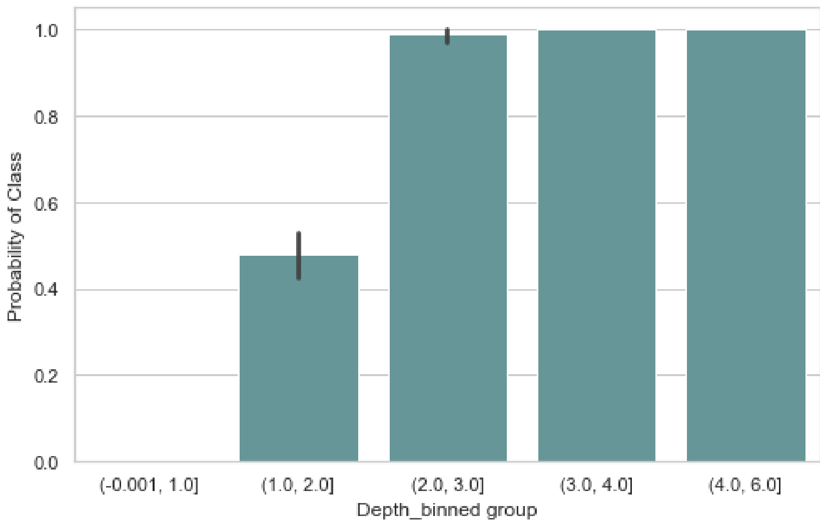


Figure 6.18: Depth binned.

After data type conversion into integer values, we now have 5 distinctive values for the binned data. From the graph, all features up to 1.0 mm in bin 0 are most likely Class 0. Features in bin 1 are have a 50 percent chance of being either Class 0 or Class 1. This can be explained by the sudden change at 1.5 mm from Class 0 to 1. Hence, all features binned in group 2, 3, or 4 have a high probability of being Class 1. It can be concluded that the feature is important for predicting Class label. Next, the DepthPercentage feature will be explored in order to uncover potential hidden patterns as it is a measure of remaining wall thickness.

6.5.6 Depth Percentage EDA

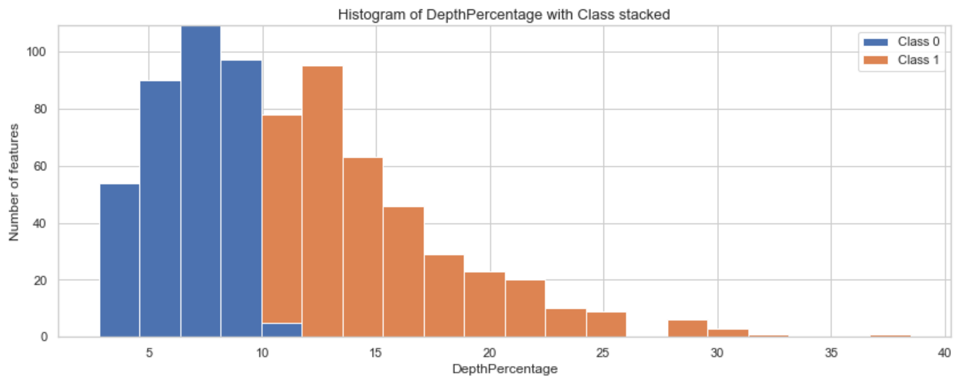


Figure 6.19: Depth percentage EDA.

Observations:

- The fraction of Class 1 increases dramatically as the depth exceed 10% of the wall thickness.
- This corresponds well with the initial raw dataset as it was sorted and classified for features below and above 10% feature depth.

It can be assumed that potential binning of features below and above 10% would result in a new feature which will be of high importance for predicting the Class label. Thus, the feature is binned in 5 groups illustrated in figure 6.20.

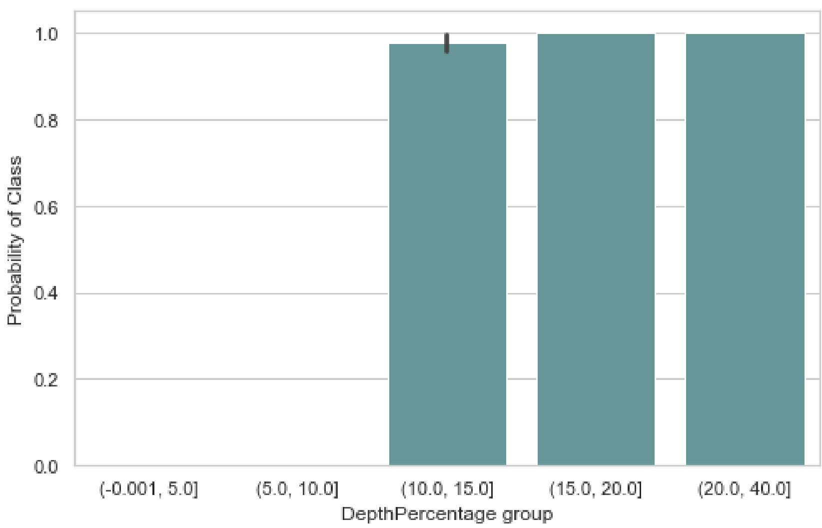


Figure 6.20: Depth percentage binned.

In comparison with Depth_binned, DepthPercentage_binned does not have a 50% chance of both classes for feature the second group. Hence, it has greatly increased the predictive power as it clearly distinguish Class based on the depth percentage. As the feature is clearly either Class 0 or Class 1, we do also risk the potential of overfitting the models, but as the failure pressure is primarily based on the remaining wall thickness at given sections we can conclude that the feature should be taken into account for the learning models.

6.5.7 Failure Pressure EDA

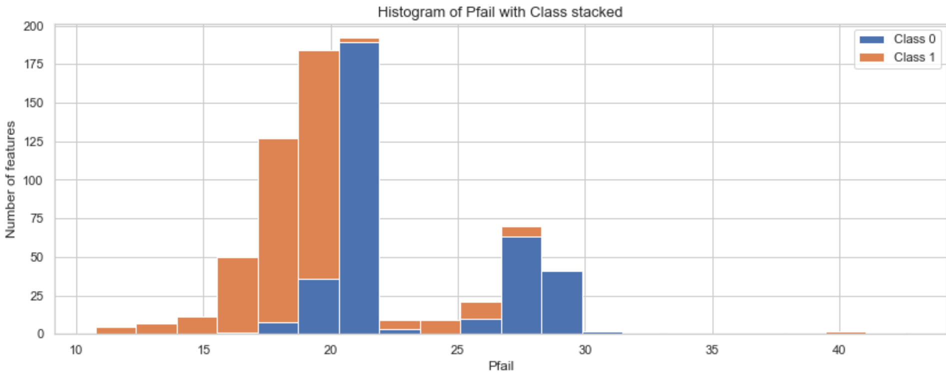


Figure 6.21: Failure pressure EDA.

Observations:

- For values below 20 N/mm^2, the majority of features are classified as Class 1.
- The majority of features are located in two local distributions with mean values of 20 N/mm^2 and 27 N/mm^2.
- First distribution has a greater amount of features, primarily Class 1, and has a lower failure pressure for the subsequent section of pipeline.
- The second distribution has a lower amount of feature, primarily Class 0, and has a higher failure pressure which makes sense.

The next step is to perform a binning of the feature in order to achieve a new feature, Pfail_binned, which will contribute to increasing the value of the dataset and thus the predictive power of the models, illustrated in figure 6.22.

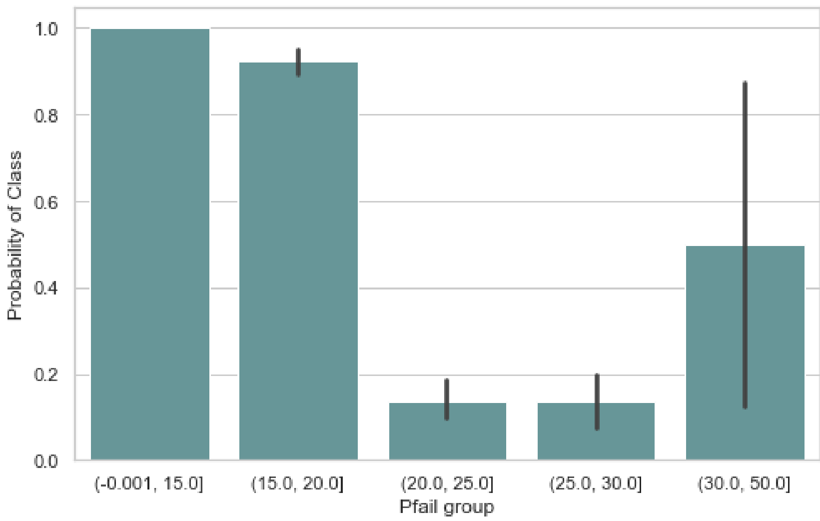


Figure 6.22: Failure pressure binned.

The result of the binning is 5 groups where the first two, covering low failure pressure features, have a high probability of being Class 1. The next two groups, covering higher failure pressure features, have a great probability of being classified as Class 0. This makes sense as a higher failure pressure would indicate higher capacity and therefore higher level of integrity. The 5th group is interesting as it consists of oulier values centered around 40 N/mm^2 in addition to being classified as Class 1 in figure 6.21, which have in figure 6.22 a 50% probability of being Class 0 and Class 1. This value will for now be left in place, but if binned feature skews results in inaccurate results for the ML models, it could be necessary to remove the outlier.

6.5.8 Feature Correlation Matrix

In figure ref, all features are plotted in the final feature correlation matrix. These features will be used in the coming sections in order to train and test the different ML algorithms and their performance of the classification problem.

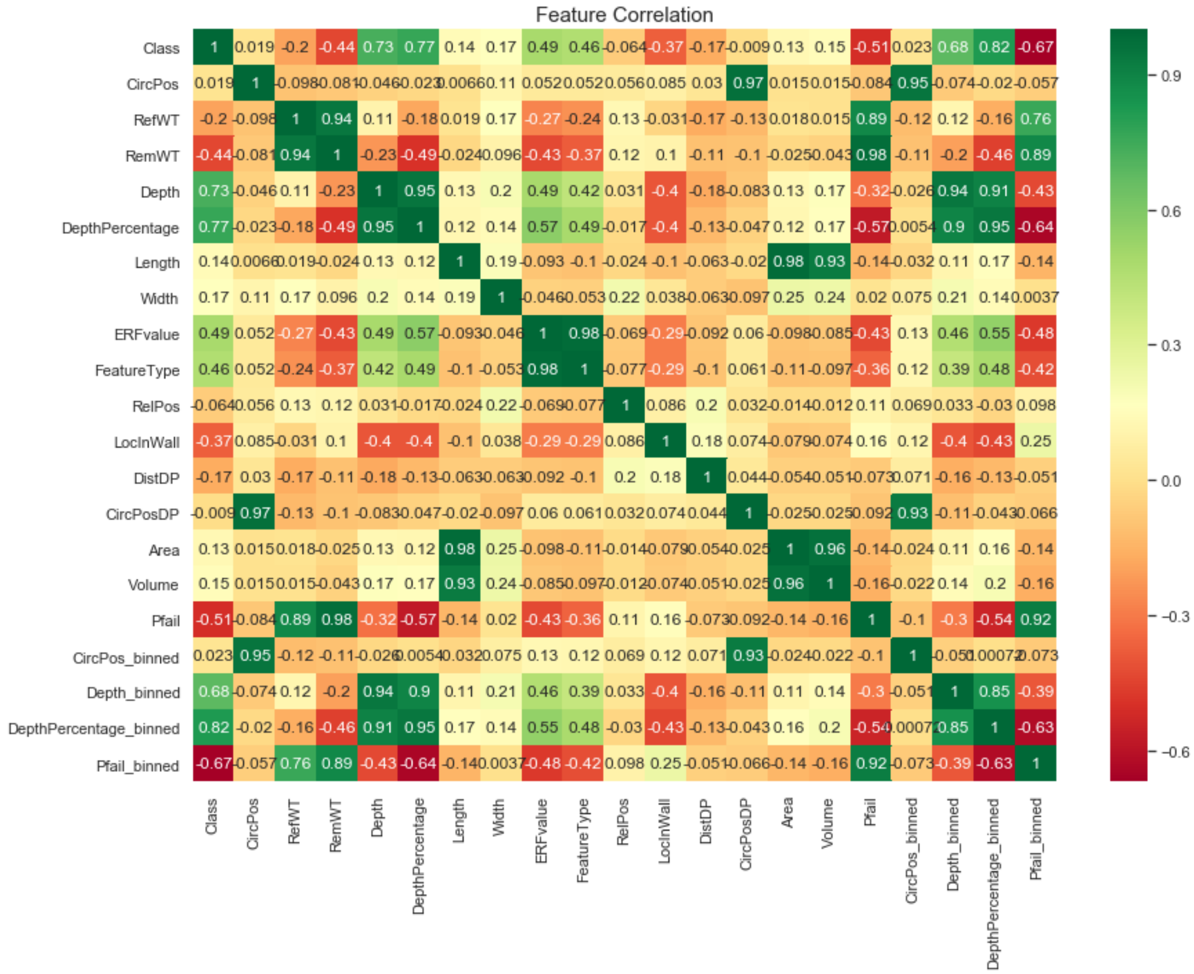


Figure 6.23: Complete feature correlation matrix.

From the feature correlation matrix, there is a high correlation between the Class label and RefWT, RemWT, Depth, DepthPercentage, ERFvalue, FeatureType, LocInWall, Pfail, Depth_binned, and DepthPercentage_binned. Multicollinearity are detected between Depth and DepthPercentage, in addition to their binned features. Multicollinear features can cause redundancies as they contain approximately the same information for the classification problem. Multicollinearity results in increased time and and computational time required to train the ML models. One should always strive to eliminate redundant features as it would increase efficiency. Will permit multicollinearities in the dataset as it is not too significant in size, but for significantly larger datasets elimination would be preferable.

6.6 Data Modeling

In this section different ML algorithms are trained on the training dataset, in order to perform a comparative analysis of which model that have the highest performance based on the available data. There is no single ML algorithm which performs best for all usecases. hence, the best approach is to generate a high performance model for the specific problem trial and error with a set of different algorithms in order to identify the best performing models, optimizing them, and then compare them given the specific problem at hand. In the following section the following steps will be performed in order to identify the best performing classification algorithm:

1. Machine learning modeling and performance comparison.
2. Hyperparameter tuning for chosen models.
3. Feature selection and optimization of best performing model.

The ML models is run on a MacBook Pro 15 2016 with 16 GB RAM and 4 CPU cores with 8 threads. The run times were acceptable but for larger datasets it would be preferable to run more complex algorithms on a server farm or by using the GPU. On a side note, if the objective were to use in example artificial neural networks (ANN) with e.g. a Python library called TensorFlow, then it would be preferable to run the training on the GPU in tandem with the RAM as storage as this would reduce the computational time required as the training can be parallelized.

6.6.1 Machine Learning Modeling and Model Performance

The machine learning models selected for the initial testing are considered to be some of the best for solving classification problems. As we have a binary classification problem, the hypothesis is that the best performing models will be the ones that manage to create a boundary decision region for based on the training set without achieving over- or underfitting of the model. In the Python kernel, all the models are first imported and listed in a list of estimators. The second step is to separate independent features and dependent features from the dataset. This means that we split the training set into two sets where the true output value is stored in a separate dataframe which then is compared with the predicted value to give a measure of accuracy. After separating the features, we generate a dataframe to compare the performance of the ML models. The next step involves generating training and validation dataset splits for cross-validation which is used to estimate the performance of the models. The results of the performance comparison is illustrated in figure 6.24.

From the results, it is clear that the best performing models are XGBoost, Bagging Classifier, Gradient Boosting Classifier, AdaBoost Classifier, Decision Tree Classifier, and Random Forest Classifier. The weakest performing models were SVC and Nu SVC closely followed by the different versions of Naive Bayes. The initial performance of the models without tuning performed exceptionally well and it is interesting to see the wide span in performance between the best and worst performing models. Should note that this does not mean the models are generally worse, but that they are rather performing less on the available dataset and the way it is structured.

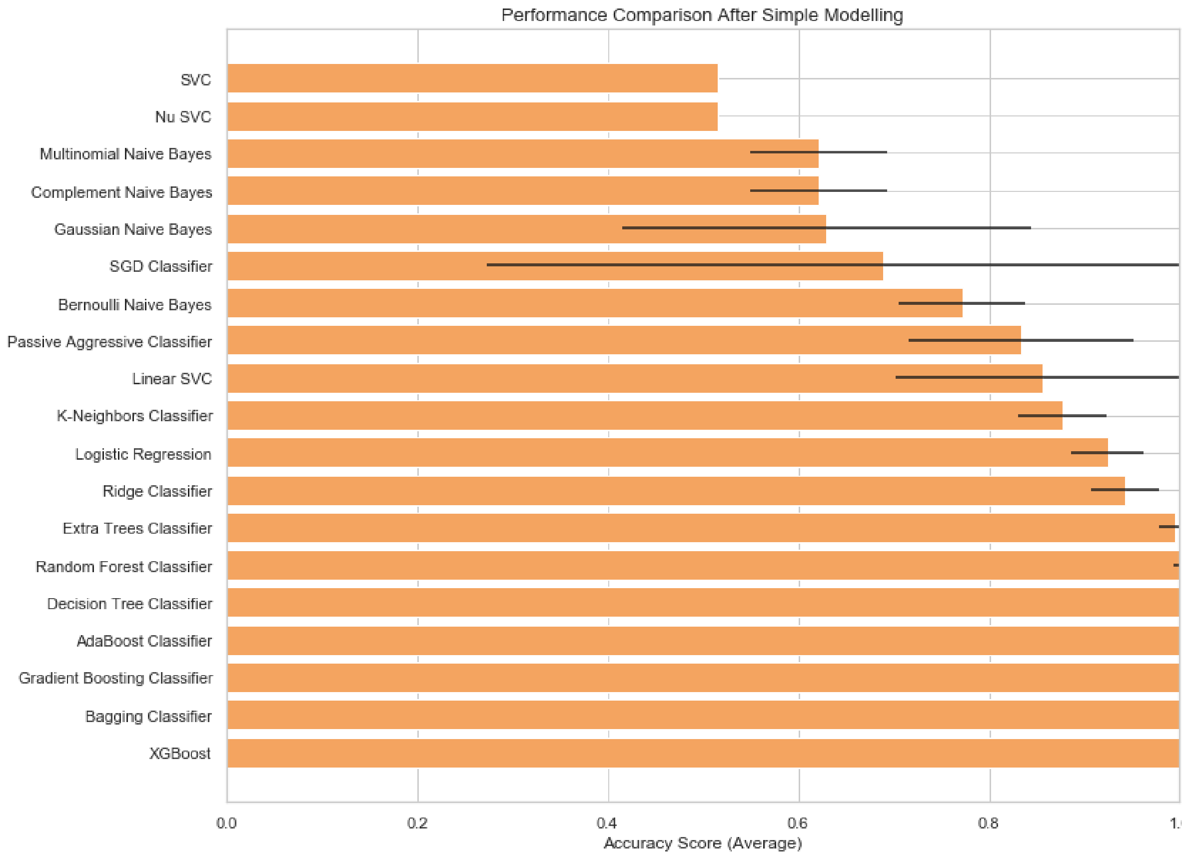


Figure 6.24: Performance comparison of machine learning models.

6.6.2 Hyperparameter Tuning

Hyperparameters are parameters with a set value before we initiate the learning process. In other words, the hyperparameters cannot be optimized through each iteration of the learning process as it is set. The process of hyperparameter optimization is about selecting values which increases the accuracy and predictive power of the model with the objective of achieving the maximized accuracy for the selected models. In the following section, some of the best and worst performing algorithms from the previous section are selected to be further optimized in order to see if there are any room for potential increases in performance. The method which will be used for optimizing the hyperparameters is the grid-search cross-validation method (Chang and Lin, 2011; Krstajic et al., 2014). The results of the hyperparameter tuning is listed in figure 6.25. As expected, the top performing models cannot be improved as they have a 100 accuracy on the test dataset. The worst performing models, SVC and Nu SVC have not improved in performance which is quite interesting. In addition, SVC and Nu SVC together with a series of algorithms listed in figure 6.24 have not improved significantly by hyperparameter tuning.

	Optimized Hyperparameters	Accuracy
SVC	[{'C': 1, 'gamma': 0.005, 'random_state': 0}]	0.515837
Nu SVC	[{'gamma': 0.001, 'nu': 0.5, 'random_state': 0}]	0.521267
Linear SVC	[{'C': 1, 'random_state': 0}]	0.852036
Ridge Classifier	[{'alpha': 0.001, 'random_state': 0}]	0.942986
Logistic Regression	[{'C': 4.75, 'random_state': 0, 'solver': 'newton-cg'}]	0.985973
Random Forest Classifier	[{'criterion': 'gini', 'max_depth': 1, 'n_estimators': 200, 'random_state': 0}]	1
AdaBoost Classifier	[{'learning_rate': 0.05, 'n_estimators': 50, 'random_state': 0}]	1
Gradient Boosting Classifier	[{'learning_rate': 0.001, 'max_depth': 1, 'n_estimators': 300, 'random_state...	1
Bagging Classifier	[{'n_estimators': 200, 'random_state': 0}]	1
XGBoost	[{'learning_rate': 0.001, 'max_depth': 1, 'n_estimators': 300, 'random_state...	1

Figure 6.25: Optimized hyperparameters.

6.6.3 Feature Selection and Optimization of Model

From the previous section 4 models achieved a 100 percent accuracy for predicting the Class label. This was also achieved by the simple models. Hence, we can choose the model we want from these and investigate how few features the models needs to achieve this accuracy. As the computational time was insignificant, we can conclude that the reduction of features is not necessary for this size of dataset. If the size increases significantly though combined with a more severely corroded pipeline, there may be necessary to remove features to save time.

Feature selection is the process of selecting features based on their value towards solving the problem, contributing highest to the output feature which is Class in this problem. If the dataset consists of features with values that do not contribute, they are considered irrelevant and can decrease accuracy of the model by making it learn on said features. A technique for identifying features with high importance is recursive feature elimination with cross validation (RFECV). RFECV will in the following section be applied to XGBoost model to investigate feature importance.

6.7 Model Evaluation

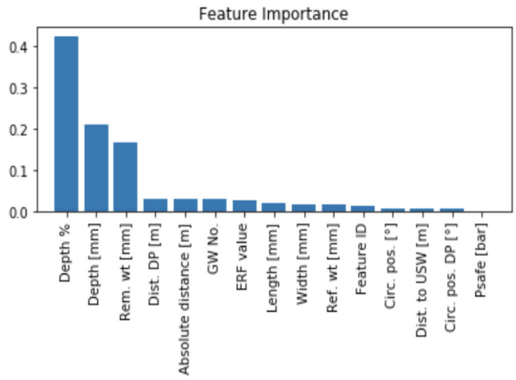
From the feature ranking, the features of highest importance for the Class label is DepthPercentage, Pfail_binned, DepthPercentage_binned, Depth_binned, CircPos_binned, and Pfail. These features have as previously stated multicollinearities between them as illustrated in figure 6.23. This is not an issue for the classification problem in this case, but it is fair to assume it has contributed to longer computational time. As the numbering indicates, the top features are the most important features for the selected algorithm.

By selecting just the top four ranked features, the XGBoost model achieves a 100% accuracy

	Features	Ranking
4	DepthPercentage	1
19	Pfail_binned	2
18	DepthPercentage_binned	3
17	Depth_binned	4
16	CircPos_binned	5
15	Pfail	6
14	Volume	7
13	Area	8
12	CircPosDP	9
11	DistDP	10
10	LocInWall	11
9	RelPos	12
8	FeatureType	13
7	ERFvalue	14
6	Width	15
5	Length	16
3	Depth	17
2	RemWT	18
1	RefWT	19
0	CircPos	20

(a) Feature ranking.

```
 1) Depth %                          0.423697
 2) Depth [mm]                       0.210334
 3) Rem. wt [mm]                     0.165497
 4) Dist. DP [m]                     0.031118
 5) Absolute distance [m]            0.029822
 6) GW No.                           0.028332
 7) ERF value                        0.026422
 8) Length [mm]                      0.019028
 9) Width [mm]                       0.017715
10) Ref. wt [mm]                     0.016763
11) Feature ID                       0.012758
12) Circ. pos. [°]                   0.006550
13) Dist. to USW [m]                 0.006447
14) Circ. pos. DP [°]                0.005169
15) Psafe [bar]                      0.000348
```



(b) Feature importance.

Figure 6.26: Feature ranking and original feature importance.

with the selected features. Hence, we can build a model with significantly fewer features which performs at the exact same accuracy. This is the power of feature selection in machine learning. This complies with DNV-RP-F101 and the assessment methodology for the allowable stress approach and the features it takes into account when assessing the integrity level of anomalies. In regard to the feature importance listed in figure 6.26a, this corresponds quite well with the initial estimations of feature importance performed on the original dataset as illustrated in figure 6.26b.

7 Discussion

Corrosion in pipelines is regarded as a major issue which cannot be neglected. It could have severe consequences for the integrity of the pipeline, and potentially result in leakage if not properly managed. The consequences of a leakage and the subsequent flow of hydrocarbons into the surrounding environment is a frightening thought. Such a situation can cause damage to infrastructure, people, the environment, in addition to having significant economic and operational consequences for the operator of the pipeline. In order to prevent corrosion induced failures from occurring, there are various types of mitigation and prevention strategies. Measures can be taken to impede the electrochemical process of corrosion such as using cathodic protection, use inhibitors, etc. To ensure that the integrity management strategy is adequate for the specific pipeline, it is important to support the integrity management strategy with accurate measurements and classification of existing corrosion features in addition to predicted growth rates for the existing corrosion features.

7.1 The potential of Machine Learning

We have in this thesis proved that supervised machine learning models have the capabilities to accurately predict the class label of features. The hypothesis was that the most important features for the classification algorithms would be the depth features. This is based on the allowable stress approach and how it focuses on feature depth and remaining wall thickness to calculate the failure pressure, in addition to the classified labels actually being labeled according to the depth percentage feature.

Based on conversations with industry domain experts on both machine learning and integrity management in addition to findings in the literature review, the current development is towards creating predictive models for estimating corrosion growth rates for different sections of pipelines. This can be performed by taking into account flow assurance modeling, pipeline topography, production and inspection history in addition to ILI data. Flow assurance modeling involves analyzing the contents of the pipeline and here we have traditional models for predicting corrosion growth rates which can be used as a basis for new predictive models corrected by machine learning. To successfully develop and deploy such models, there are certain requirements that should be met. First, the data should be covering detailed production history, inspection history, etc. In addition, it is necessary to get sufficient amounts of data on pipelines that are severely corroded. This was one of the major drawback of the data used for this thesis. It was simply not enough corrosion features to train the models on correctly classified labels as the labeling was downgraded to 10% depth of wall thickness.

Gains from automation

Machine learning automates important parts of the model building that previously required deep domain expertise (Chollet, 2017). Thus, it lowers the barriers to developing an artificial intelligence for a given domain. The largest challenges are associated with the building of the models, but after the model is fully developed, it can be used to automate processes which previously was extremely time consuming and repetitive analyses performed by human domain experts. It can be assumed that for operational monitoring, the engineers are primarily working and monitoring work during office hours. Machine learning have the potential for automating weekly monitoring of production values as it can perform it live at all times continuously. For the specific case of classifying class labels and assessing pipeline integrity, it can be argued that the effectiveness of the presented models demonstrate the potential for using machine learning in previously manual processes.

7.2 Practical Barriers of Machine Learning

As the literature study found, the major barriers for the application of machine learning is related to knowledge about the development and deployment process combined with the knowledge for how to assemble teams with the necessary domain expertise in their respective sciences. Despite having large amounts of raw data available and expert knowledge, the industry suffers from slowly adapting new technological developments and usually gets stuck in the development phase of the models. In other words, it is of high importance when performing projects related to machine learning to have a project leader which has the necessary knowledge of how to successfully develop and deploy the models. This means to assembly a team with experts in their respective fields such as corrosion and machine learning, in addition to having experts with background in computer science which can take into account the knowledge of the corrosion experts in order to develop models which takes into account all necessary aspects of the problem which they want to solve. It is also important to take into account the existing systems that the specific company uses as projects related to machine learning most likely is initiated with the objective of increasing the effectiveness and efficiency of existing procedures, etc. Hence, the importance of involving people with expert knowledge of the existing systems and methods.

7.3 Theoretical Limitations and Avenues for Further Research

The major limitation to the work performed in this thesis is related to the available data on defects. As the majority of defects in the selected pipeline could be considered as single defects, the result was that the models were not trained on data of complex shaped defects. For future work, the next step would be to have available large quantities of data on pipelines suffering from severe corrosion. This would enable us to use the results from the traditional time-consuming assessment methodologies in order to label the data. This would result in increased value for the data as we now will have the ability to train models on severe corrosion features.

8 Conclusion

The thesis has demonstrated the usage and accuracy of machine learning towards solving classification problems by utilizing exploratory data analysis on feature importance and feature correlation, and gives clear recommendations for future work on defect assessment and integrity management of pipeline systems. The best performing models achieved a 100% accuracy in predicting the class label which is above expectations. Based on the results and the literature review, we can conclude the following.

Pipelines are suffering from corrosion due to a number of reasons and the mitigation and prevention technologies cannot prevent the creation of corrosion features. At best, the pipeline lasts its entire lifetime without maintenance, but most likely it would be detected e.g. higher oxygen levels in the flow assurance modeling resulting in potential higher corrosion growth rates and the subsequent need for performing ILI inspections. This can be challenging on pipeline systems which are not designed for receiving pigging tools. Many pipelines are unpiggable due to a number of reasons, increasing the need for new predictive models. Current research is focusing on using flow assurance modeling together with ILI data and historical data in complex models in order to predict corrosion severity in such unpiggable regions. Predicting corrosion growth without the use of data analytic tools have severe limitations due to the complex nature of corrosion. Machine learning can also be used as a tool for initial screening of features or for live monitoring of operational values. One of the major findings from the literature review is the potential for using machine learning on a series of ILI inspections of the same pipeline in order to identify local corrosion growth rates based on the historical corrosion data. This works by matching features' historical data and training the models to identify patterns in the corrosion features, which from history, indicating the potential for high corrosion rate in the future.

While we find the machine learning models to be exceptional for predicting class labels, we can argue that this thesis demonstrates the wide potential for machine learning in the oil and gas industry. There are many types of machine learning - not to mention of artificial intelligence - for a range of AI problems relevant to the integrity management of pipeline systems, predicting corrosion feature growth rates, perform anomaly detection on live production data, etc. What they all have in common is the need for comprehensive, reliable and high quality data, in addition to domain expertise and the computational resources for deploying complex models. The development phase is the most challenging where the need for computational power is at its highest, but one should be aware of potential up-front investments and costs related to having the computational power ready for deployment. If the company does not perform machine learning or AI related training on a regular basis, cloud computing is a great alternative. It should also be noted that the up-front investments and related costs could be significantly lower in the long-term if the alternative is to rely on manual work performed by domain experts. If the work is of repetitive nature, the expert knowledge can be of higher need in other projects. Machine learning is on the rise in may industries being the major buzzword these days, and while it is of the utmost importance to have realistic expectations and be aware of the potential pitfalls when initiating machine learning projects in a conservative and slowly evolving industry, there should be no concern and reason to assume that the oil and gas market are any different.

References

Afangide, E., Sinha, J. K., and Katnam, K. (2018). Quantified approach to pipeline health and integrity management. *Journal of Loss Prevention in the Process Industries*, 54:28 – 36.

Ahammed, M. and Melchers, R. (1996). Reliability estimation of pressurised pipelines subject to localised corrosion defects. *International Journal of Pressure Vessels and Piping*, 69(3):267 – 272.

Alamilla, J., Espinosa-Medina, M., and Sosa, E. (2009). Modelling steel corrosion damage in soil environment. *Corrosion Science*, 51(11):2628 – 2638.

Antaki, G. (2003). *Piping and Pipeline Engineering: Design, Construction, Maintenance, Integrity, and Repair*. Mechanical Engineering. CRC Press.

Caleyo, F., Alfonso, L., H Espina-Hernández, J., and M Hallen, J. (2007). Criteria for performance assessment and calibration of in-line inspections of oil and gas pipelines. *Measurement Science and Technology*, 18:1787.

Chang, C.-C. and Lin, C.-J. (2011). Libsvm: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):27.

Choi, J., Goo, B., Kim, J., Kim, Y., and Kim, W. (2003). Development of limit load solutions for corroded gas pipelines. *International Journal of Pressure Vessels and Piping*, 80(2):121 – 128.

Chollet, F. (2017). *Deep Learning with Python*. Manning Publications Co., Greenwich, CT, USA, 1st edition.

Committee, A. B. et al. Asme b31g-2012 (r2017): Manual for determining the remaining strength of corroded pipelines: Supplement to b31 code for pressure piping.

Committee, A. B. et al. (2009). Asme b31g-2009: Manual for determining the remaining strength of corroded pipelines. *American Society of Mechanical Engineers*.

Cosham, A., Hopkins, P., and Macdonald, K. (2007). Best practice for the assessment of defects in pipelines – corrosion. *Engineering Failure Analysis*, 14(7):1245 – 1265.

da Cunha, S. B. (2016). A review of quantitative risk assessment of onshore pipelines. *Journal of Loss Prevention in the Process Industries*, 44:282 – 298.

Hajizadeh, Y. (2019). Machine learning in oil and gas; a swot analysis approach. *Journal of Petroleum Science and Engineering*, 176:661–663.

Hamann, R. and Lampe, J. (2018). Probabilistic model for corrosion degradation of tanker and bulk carrier. *Marine Structures*, 61:309–325.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer series in statistics. Springer.

Hrncir, T., Turner, S., Polasik, S. J., Vieth, P., Allen, D., Lachtchouk, I., Senf, P., and Foreman, G. (2010). A case study of the crack sizing performance of the ultrasonic phased array combined crack and wall loss inspection tool on the centennial pipeline, the defect evaluation, including the defect evaluation, field feature verification and tool performance validation (performed by marathon oil, dnv and ge oil & gas). (44205):137–143.

Ilman, M. and Kusmono, K. (2013). Analysis of internal corrosion in subsea oil pipeline. *Case Studies in Engineering Failure Analysis*, 2.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Khodayari-Rostamabad, A., Reilly, J. P., Nikolova, N. K., Hare, J. R., and Pasha, S. (2009). Machine learning techniques for the analysis of magnetic flux leakage images in pipeline inspection. *IEEE Transactions on Magnetics*, 45(8):3073–3084.

Kiefner, J., Maxey, W., Eiber, R., and Duffy, A. (1973). Failure stress levels of flaws in pressurized cylinders. In *Progress in flaw growth and fracture toughness testing*. ASTM International.

Kiefner, J. and Vieth, P. (1989a). A modified criterion for evaluating the remaining strength of corroded pipe.

Kiefner, J. F. and Vieth, P. H. (1989b). A modified criterion for evaluating the remaining strength of corroded pipe. Technical report, Battelle Columbus Div., OH (USA).

Kishawy, H. A. and Gabbar, H. A. (2010). Review of pipeline integrity management practices. *International Journal of Pressure Vessels and Piping*, 87(7):373 – 380.

Krstajic, D., Buturovic, L. J., Leahy, D. E., and Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of cheminformatics*, 6(1):10.

Larin, O., Barkanov, E., and Vodka, O. (2016). Prediction of reliability of the corroded pipeline considering the randomness of corrosion damage and its stochastic growth. *Engineering Failure Analysis*, 66:60–71.

Layouni, M., Hamdi, M. S., and Tahar, S. (2017). Detection and sizing of metal-loss defects in oil and gas pipelines using pattern-adapted wavelets and machine learning. *Applied Soft Computing*, 52:247 – 261.

Leis, B., Stephens, D., et al. (1997). An alternative approach to assess the integrity of corroded line pipe-part i: current status. In *The Seventh International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers.

Liu, H., Liu, Z., Taylor, B., and Dong, H. (2019). Matching pipeline in-line inspection data for corrosion characterization. *NDT & E International*, 101:44–52.

Liu, Y., Hu, H., and Zhang, D. (2013). Probability analysis of damage to offshore pipeline by ship factors. *Transportation Research Record: Journal of the Transportation Research Board*, 2326:24–31.

McNealy, R., McCann, R., Van Hook, M., Stiff, A., and Kania, R. (2010). In-line inspection performance iii: Effect of in-ditch errors in determining ili performance.

Melchers, R. E. (1999). Corrosion uncertainty modelling for steel structures. *Journal of Constructional Steel Research*, 52(1):3 – 19.

Melchers, R. E. and Beck, A. T. (2018). *Structural reliability analysis and prediction*. John Wiley & Sons.

Mishra, M., Keshavarzzadeh, V., and Noshadravan, A. (2019). Reliability-based lifecycle management for corroding pipelines. *Structural Safety*, 76:1 – 14.

Mora, R. G., Murray, A., Paviglianiti, J., and Abdollahi, S. (2008). Dealing with uncertainty in pipeline integrity and rehabilitation. (48609):611–619.

Morgenstern, L. and McIlraith, S. A. (2011). John mccarthy's legacy.

Nash, I. (2011). Inspection maintenance and repair of deepwater pipelines. Deep and Ultra-deepwater Pipelines Conference.

Netto, T., Ferraz, U., and Estefen, S. (2005). The effect of corrosion defects on the burst pressure of pipelines. *Journal of Constructional Steel Research*, 61(8):1185 – 1204. Second Brazilian special issue.

Palmer, A. C. A. C., King, Roger A., ., and (Firm), P. (2004). *Subsea pipeline engineering.* Tulsa, Okla. : PennWell. Includes bibliographical references and index.

Parsi, M., Najmi, K., Najafifard, F., Hassani, S., McLaury, B. S., and Shirazi, S. A. (2014). A comprehensive review of solid particle erosion modeling for oil and gas wells and pipelines applications. *Journal of Natural Gas Science and Engineering*, 21:850–873.

R Core Team (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

R. Walker, J., Mallaburn, P., and Balmer, D. (2010). Inline inspection: Both effective data collection and interpretation needed to achieve high quality reporting results.

Rankin, L. (2004). Pipeline integrity information integration. *Materials performance*, 43:56–60.

Raschka, S. and Mirjalili, V. (2017). *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-learn, and TensorFlow, 2Nd Edition.* Packt Publishing, 2nd edition.

Ricker, R. (2007). *Analysis of Pipeline Steel Corrosion Data From NBS (NIST) Studies Conducted Between 1922-1940 and Relevance to Pipeline Management.*

Ritchie, D. and Last, S. (1995). Burst criteria of corroded pipelines-defect acceptance criteria. In *Proceedings of the EPRG/PRC 10th biennial joint technical meeting on line pipe research.*

Schneider, C., Muhammed, A., and M Sanderson, R. (2001). Predicting the remaining lifetime of in-service pipelines based on sample inspection data. *Insight - Non-Destructive Testing and Condition Monitoring*, 43:102–104.

Shaik, M. R. (2015). Pipeline integrity assessment: Methodology. (56468):V001T03A001–V001T03A001.

Solutions, P. I. I. P. (1998). Ula oil export ili report 1997. Technical report.

Spencer, K., Kariyawasam, S., Tetreault, C., and Wharf, J. (2010). A practical application to calculating corrosion growth rates by comparing successive ili runs from different ili vendors. In *2010 8th International Pipeline Conference*, pages 467–473. American Society of Mechanical Engineers.

Stephens, D. R. and Leis, B. N. (2000). Development of an alternative criterion for residual strength of corrosion defects in moderate-to high-toughness pipe. In *2000 3rd International Pipeline Conference*, pages V002T06A012–V002T06A012. American Society of Mechanical Engineers.

Sulaiman, N. and Tan, H. (2014). Third party damages of offshore pipeline. *Journal of Energy Challenges and Mechanics*, 1(1).

Systems, N. and Services (2011). Report 20ula-to-ekofiskj-e$_{d}raft.Technicalreport.$

Tang, P., Yang, J., Zheng, J., Wong, I., He, S., Ye, J., and Ou, G. (2009). Failure analysis and prediction of pipes due to the interaction between multiphase flow and structure. *Engineering Failure Analysis*, 16(5):1749–1756.

Teixeira, A., Soares, C. G., Netto, T., and Estefen, S. (2008). Reliability of pipelines with corrosion defects. *International Journal of Pressure Vessels and Piping*, 85(4):228–237.

Ukpai, J., Barker, R., Hu, X., and Neville, A. (2013). Determination of particle impacts and impact energy in the erosion of x65 carbon steel using acoustic emission technique. *Tribology International*, 65:161–170.

Varela, F., Tan, M. Y., and Forsyth, M. (2015). An overview of major methods for inspecting and monitoring external corrosion of on-shore transportation pipelines. *Corrosion Engineering, Science and Technology*, 50(3):226–235.

Veritas, D. N. (2015). Dnv-rp-f101 recommended practice for corroded pipelines.

Wang, H., Yajima, A., Liang, R. Y., and Castaneda, H. (2015). A bayesian model framework for calibrating ultrasonic in-line inspection data and estimating actual external corrosion depth in buried pipeline utilizing a clustering technique. *Structural Safety*, 54:19 – 31.

Wang, W., Q. Smith, M., H. Popelar, C., and A. Maple, J. (1998). A new rupture prediction model for corroded pipelines under combined loadings. page 563.

Xie, M. and Tian, Z. (2018). A review on pipeline integrity management utilizing in-line inspection data. *Engineering Failure Analysis*, 92:222 – 239.

Yang, Y., Khan, F., Thodi, P., and Abbassi, R. (2017). Corrosion induced failure analysis of subsea pipelines. *Reliability Engineering System Safety*, 159:214 – 222.

Zdravecká, E., Slota, J., and Tkáčová, J. (2014). Erosive failure of steel pipeline by solid pulverized particles. *Engineering Failure Analysis*, 46:18–25.

Zhang, H. (2004). The optimality of naive bayes. *AA*, 1(2):3.