A Generalized Machine Learning-Based Classifier Considering Cost-Effective Features for Automated Fault Detection and Diagnosis (AFDD) of Packaged Rooftop Units

Md Rasel Uddin
University of Nebraska-Lincoln, rasel07me@gmail.com

A GENERALIZED MACHINE LEARNING-BASED CLASSIFIER CONSIDERING

COST-EFFECTIVE FEATURES FOR AUTOMATED FAULT DETECTION AND

DIAGNOSIS (AFDD) OF PACKAGED ROOFTOP UNITS

by

Md Rasel Uddin

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Mechanical Engineering and Applied Mechanics

Under the Supervision of Professor Robert Williams

Lincoln, Nebraska

December, 2023

A GENERALIZED MACHINE LEARNING-BASED CLASSIFIER CONSIDERING COST-EFFECTIVE FEATURES FOR AUTOMATED FAULT DETECTION AND DIAGNOSIS (AFDD) OF PACKAGED ROOFTOP UNITS

Md Rasel Uddin, Ph.D.

University of Nebraska, 2023

Advisor: Robert Williams

Packaged rooftop units (RTUs) are widely used for space conditioning in commercial buildings and manufacturing facilities. The typical soft faults related to RTUs degrade the system's performance in terms of cooling capacity, power consumption, and Coefficient of Performance (COP), detrimentally affecting both the equipment and energy consumption and the environment. Previous research in soft fault detection for rooftop units lacked classifier validation using lab and field data, developing a generalizable algorithm, and analyzing its performance across varying fault intensities.

Using a simulated data library for multiple rooftop units, this study proposes a machine-learning classifier with a reduced set of 9 features (8 quantitative and one qualitative) to detect and diagnose typical soft faults in packaged rooftop units equipped with fixed orifice metering devices. An existing lab testing set consisting of the same training systems was utilized to validate the presented data-driven approach, showing significantly better performance than the existing fault detection and diagnosis protocols. In addition, the analyzed classifier's predicting performance improves with increasing fault severity.

In addition to the above lab validation, a manufacturing facility in Omaha, Nebraska, was chosen for field validation of the developed machine-learning algorithm. The proposed approach accurately predicted all the refrigerant undercharge fault cases from an RTU at that facility, although the RTU significantly differs from the RTUs with which the classifier was trained. The lab and field-testing results bolster that the considered machine-learning classifier can be generalizable, with some exceptions, for detecting the common soft faults from any rooftop unit equipped with a fixed orifice metering device. The presented classifier can be used in an industrial assessment for diagnosing common soft faults from an RTU, helping to develop additional energy and cost savings measures for a facility.

Dedication

To my late maternal grandparents (Nana and Nani).

Acknowledgments

I would like to take this opportunity to express my heartfelt gratitude to several individuals and organizations whose unwavering support made this dissertation possible. First, I would like to thank my supervisor, Dr. Robert Williams, for his tremendous support and belief in me. He has given me the privilege of serving the Nebraska Industrial Assessment Center (NIAC) in its mission to enhance the energy efficiency of manufacturing facilities in the greater Nebraska area. Dr. Williams encouraged me to focus on energy optimization measures related to Heating, Ventilation, and Air-Conditioning (HVAC) systems, and I learned a lot through this experience. As a result, several unique recommendations related to HVAC systems have been written during my tenure at the NIAC at UNL. I also want to thank Dr. Bruce Dvorak for his support while I worked under the NIAC at UNL. I am also deeply grateful to UNL's Energy Engineer, Samuel Ghormley, for our endless discussions on potential energy-saving measures fueling my passion for this field.

As I started looking for potential Ph.D. topics related to HVAC systems, I came across another brilliant mind, Dr. David Yuill. Dr. Yuill welcomed me and treated me as his own student. His guidance led me to discover areas closely aligned with NIAC's mission, ultimately leading me to select the topic on fault detection and diagnosis of packaged rooftop units. His recommended building energy courses under the Architectural Engineering department have greatly aided my research progress, and I'll certainly miss

the Lincoln to Omaha commute for those classes and our meetings. I also would like to thank two of Dr. Yuill's graduate students, Ali and Ashkan, for their cooperation during the field testing at Omaha, NE.

In 2022, thanks again to the unwavering support of Dr. Williams and Dr. Yuill, I was honored with the IAC Student Research Award. This award, presented by the U.S. Department of Energy, served as a profound motivation for me to dive deeper into the above research area. I am profoundly thankful to the U.S. DOE for this recognition.

I also would like to thank Dr. K. P. Rajurkar for serving on my Ph.D. committee. His teachings from various advanced manufacturing engineering courses helped me to go forward in my research endeavor.

Speaking of my family, I am indebted to my father for his continuous motivation and relentless support for pursuing higher education. My mother has been a constant pillar of support throughout my life. Finally, my heartfelt thanks go out to my wife, Firoza, and our two boys, Arham and Irham. Their unwavering support, love, and understanding have been the driving forces that propelled me forward on this Ph.D. journey while residing abroad.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1   Motivation

The motivation for this research came after performing several industrial assessments for the Nebraska Industrial Assessment Center (NIAC) at the University of Nebraska-Lincoln (UNL). The assessments require submitting recommendations estimating energy and cost savings for an industrial facility. As an NIAC student, several recommendations related to the Heating, Ventilation, and Air-Conditioning (HVAC) system of the facilities were written in the following topics,

- Elimination of soft faults from packaged rooftop units

- Fix or replace faulty economizers of the packaged rooftop units

- Proper thermostat setpoints

- Optimum chiller operation

- Replace electric resistance heating with other sources

- Use cogged v-belt in the fans of supply, return, and exhaust air systems

- Heat recovery from the air compressors

Each of the above recommendations provides significant energy and cost savings opportunities for the manufacturing facilities, and thus the motivation to perform a research project related to the HVAC system of a facility came afterward.

Another motivation came after looking for the answers to the question: how to improve the productivity of a manufacturing facility? One direct solution can be by developing novel manufacturing technologies like Abrasive Flow Machining (AFM) that can

effectively automate the finishing process to build precision components (Williams,

1993). The indirect answer to the above question is related to the HVAC system of a

facility. A better work environment can improve employees' health and productivity

through proper ventilation, air quality, thermal comfort, visual comfort, and acoustic

comfort (Mujan et al., 2019; Seppänen and Fisk, 2006).

1.2   Background: Fault Detection and Diagnosis (FDD)

In this section, a basic overview of the common soft faults in rooftop units and the Fault

Detection and Diagnosis will be provided.

1.2.1   Terminology

In this section, some basic definitions of the common terms used in this dissertation will

be provided. More detailed explanations of the common terms will be provided in the

related sections. The basic definitions for some common terms are shown in Table 1.

Table 1: Basic terminologies

Term	Definition
System	A set of interacting elements.
Fault Detection	Identifying the presence of an abnormality (fault) in a system.
Fault Diagnosis	Identifying the exact fault type and location in a system.
Fault Detection and Diagnosis (FDD)	Identifying the presence of a fault with its type and location.
Response Variable	A dependent variable for a set of independent variables.
Machine-Learning (ML)	A process of fitting mathematical model (algorithm) using suitable data.
Supervised Machine-Learning	A computer algorithm trained with known response variables.
Machine-Learning Classifier	A computer algorithm for predicting qualitative output (class).
Generalized Classifier	A ML classifier that can be used for predicting qualitative output for any type of related systems.
Support Vector Machine (SVM)	A ML classifier based on finding an optimal decision boundary to predict the qualitative output.
Training Set	A dataset used for training a ML algorithm.
Testing Set	A dataset used for testing the performance of a ML algorithm.
Confusion Matrix (CM)	A tabular representation of a ML classifier's performance showing number of classes in both actual and predicted scenarios.
False Alarm	In the FDD study, a false alarm happens when a ML classifier predicts an unfaulty case as faulty.
Misdiagnosis	In the FDD study, a misdiagnosis happens when a ML classifier predicts the presence of a fault, but a wrong fault is predicted.
Missed Detection	In the FDD study, a missed detection happens when a ML classifier fails to detect the presence of a fault.
Vapor Compression Refrigeration Cycle (VCRC)	A set of interacting elements that uses a refrigerant to absorb and reject heat from and to a space, respectively.
Cooling Capacity	The amount of heat absorbed from a space.
Coefficient of Performance (COP)	The ratio of cooling capacity to the power consumption in a VCRC.
Energy Efficiency Ratio (EER)	At a particular operating condition, it is the ratio of heat absorbed (in BTU/hour) from a space to the power input (in Watt).
Soft Fault	A fault that degrades the system performance without shutting off the associated equipment.

Term	Definition
Fault Impact Ratio (FIR)	A representation of the severity of a fault from a VCRC. In terms of COP, this is the ratio of the faulty COP to the unfaulty COP.
Feature Selection	Selection of suitable inputs for fitting a ML classifier.
Filter Feature Selection	Selection of features without using a ML classifier.
Wrapper Feature Selection	Selection of features using an underlying ML classifier.
Fixed Orifice (FXO)	A metering device, used in VCRC, that creates constant resistance to refrigerant flow, lowering the refrigerant pressure.
Thermostatic Expansion Device (TXV)	A metering device, used in VCRC, that can create variable resistance to refrigerant flow, lowering the refrigerant pressure.

1.2.2 Vapor Compression Refrigeration Cycle (VCRC) and Rooftop Unit (RTU)

A system is a set of several interacting elements. A Vapor Compression Refrigeration Cycle (VCRC) is a part of an HVAC system found in packaged rooftop units (RTUs), split systems, chillers, and other systems. A schematic diagram of the VCRC is shown in Figure 1.
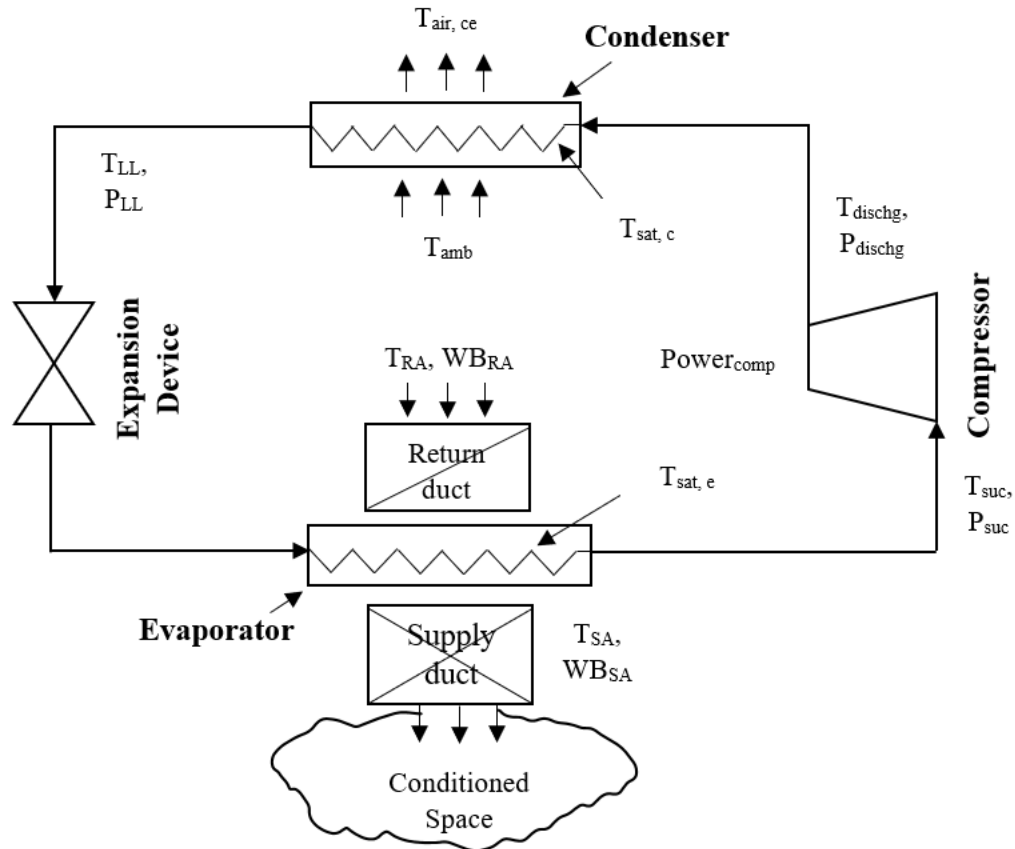
Figure 1: Schematic of Vapor Compression Refrigeration Cycle (VCRC)

As noticed from Figure 1, the four major interacting elements in VCRC are the compressor, condenser, expansion device, and evaporator. The compressor pressurizes the vapor refrigerant to an extent higher than the saturation pressure corresponding to ambient temperature. This allows the pressurized vapor to undergo a phase change to liquid by rejecting heat to the ambient in the condenser. After condensation, the liquid refrigerant goes through the expansion device to decrease its pressure below the saturation pressure corresponding to the space temperature (i.e., return air temperature). This allows the refrigerant to undergo a phase change from liquid to vapor in the evaporator by absorbing heat from the space.

Based on the ability to control the refrigerant flow, there are two expansion devices: fixed area and adjustable area. An example of a fixed area device is a Fixed Orifice (FXO) device that allows a fixed refrigerant flow resistance irrespective of cooling load from the space. The expansion device of the VCRC in this study is FXO. In contrast to the FXO, an adjustable area device (e.g., Thermostatic Expansion Valve, TXV) allows a variable amount of refrigerant useful for variable cooling load demand and thus mitigating the effect of possible faults (described in the next section) in the system.

Rooftop Unit (RTU) is a type of HVAC system, typically found in commercial buildings and manufacturing facilities, that incorporates a VCRC. This is a type of unitary system which comes in a compact package containing all the components of a VCRC. RTU uses direct expansion (DX) coils in the evaporator to directly exchange heat between the refrigerant and incoming air. Unlike the chilled water coil, DX coil does not require the chiller to serve cold water across the evaporator coils. The RTU serves conditioned air in single or multi zoned spaces.

In a typical rooftop unit, there are mainly four air streams: outside air (OA), return air (RA), mixed air (MA), and supply air (SA). The OA takes the ambient air, and the RA takes the space air. The space air is drawn into the RTU package using a RA fan. The OA and RA are then mixed with the required proportion, and the MA first goes through a filter followed by a cooling coil and heating coil. The cooling coil is same as the evaporator coil (mentioned in VCRC) that absorbs heat from the MA if there is a cooling demand from space. The heating coil is either electric or gas fired to heat the MA. The conditioned air from the cooling/heating coils is called the SA that is served to different zones using a blower fan.

1.2.3 FDD Study in VCRC

There are two types of faults related to VCRC: soft and hard (Breuker and Braun, 1998).

Soft faults are those that degrade the system performance or reduce the lifetime of the

equipment, allowing continued system operation. These faults generally go unnoticed by

the system operators until significant equipment failure or hard faults occur (Yuill and

Braun, 2013). As a result, soft faults are challenging to diagnose (Rogers et al., 2019).

This study is focused on the soft faults related to packaged rooftop units (RTUs). Typical

soft faults related to the RTUs are (Yuill et al., 2014b),

- Condenser Airflow Reduction (CA)

- Evaporator Airflow Reduction (EA)

- Liquid Line Restriction (LL)

- Non-Condensable Gases (NC)

- Refrigerant Overcharge (OC)

- Refrigerant Undercharge (UC)

- Compressor Valve Leakage (VL)

Proper Fault Detection and Dignosis (FDD) methodology to address the above soft faults

is necessary as far as the system performance and economics are concerned. Automated

Fault Detection and Diagnosis (AFDD) works on detecting and diagnosing faults based

on related fault signatures (Y. Li et al., 2014). Fault signatures are those thermodynamic

(pressures and temperatures) and electrical parameters related to VCRC that show certain

changes for a specific fault type. Fault detection involves identifying the presence of

faults, and fault diagnosis includes determining the fault type, location, and intensity of the faults (Yuill and Braun, 2013). For instance, to detect the presence of faults in unitary air conditioning systems, Rossi (2004) considered some specified tolerance limits for evaporating temperature, condensing temperature over ambient, and suction superheat. In the fault diagnosis step, he performed some collection of diagnostic algorithms for recommending corrective action (e.g., adding refrigerant charge). In contrast to this two-step FDD approach, this study will perform fault detection and diagnosis simultaneously in a single step.

Three main approaches in AFDD are rule-based, physical model-based, and process history-based (Behfar et al., 2019). The rule-based approach follows setting up a fixed set of rules to detect and diagnose fault types (Chen and Braun, 2001). Physical model-based systems deals with formulating complex models associated with a particular system. In contrast to these two approaches, the process history-based approach deals with modeling statistical learning algorithms that take input data related to several thermodynamic and electrical parameters (Ebrahimifakhar et al., 2020) and has gained increased popularity nowadays. This study is based on a process history-based approach (also called the data-driven approach) to develop an AFDD algorithm for predicting the common soft fault types in RTUs.

1.3   Background of Study Data

To predict the faults based on process history/data-driven approach, it first requires plenty of input data to train the statistical learning model. However, obtaining experimental data is costly and sometimes challenging. A solution to this problem could

be to use reliable simulated data. Cheung and Braun (2013a, 2013b) developed gray-box

models for several unitary VCRC systems (split systems and RTUs) to obtain steady-state

simulated data in cooling mode. These models were validated by Yuill et al. (2014) for

their intended purpose: evaluating the performance of an FDD protocol. Since obtaining

plenty of experimental data with several operating conditions and faults is scarce and

challenging, the simulated data from the above gray-box models for packaged rooftop

units (RTU) can be used for developing a generalized statistical learning model for RTUs

in this study. A gray-box model for a particular VCRC is based on physical law and

expert knowledge to quantify several output variables for a set of input conditions. As

shown in Figure 2, the gray-box model has 4 inputs and 12 thermodynamic outputs

(pressures and temperatures), for each steady-state case. The simulated data consists of

different combinations of the 4 inputs.
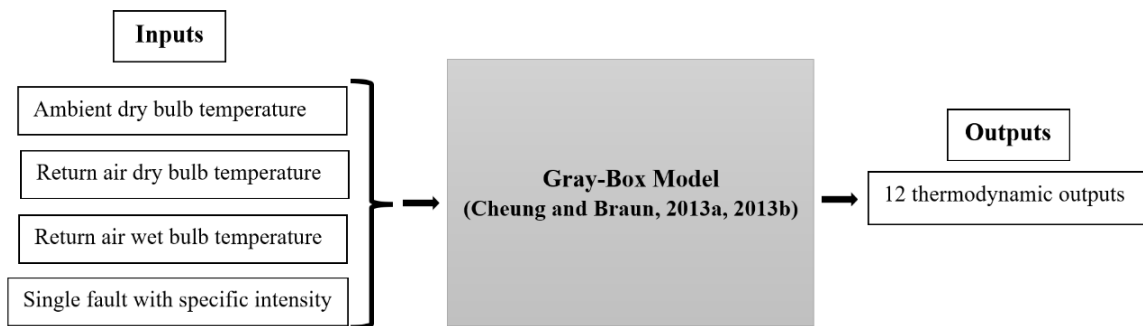
Figure 2: Simplified gray-box model from Cheung and Braun (2013a, 2013b)

In addition to the 12 major outputs, the above gray-box model also estimates other

secondary outputs like capacity, fault impact ratios, suction superheat, subcooling, and so

on.

Simulated data for three different rooftop units running in cooling mode was considered

to train the machine learning model for classifying the soft faults in rooftop units. An

existing lab dataset for the same rooftop units (RTU) operating in steady-state cooling

mode was used as a testing set to validate the developed statistical learning model. The

information related to this lab data can be found in Yuill and Braun (2013). The

simulated and lab data have 14 thermodynamic and one electrical feature shown in Table

1. The same features are shown in abbreviated forms in Figure 1.

Table 2: Features of the system

Feature Type	Feature Name
Temperature	DB temperature of the return air (T_{RA}, ºC)
	WB temperature of the return air (WB_{RA}, ºC)
	DB temperature of the supply air (T_{SA}, ºC)
	WB temperature of the supply air (WB_{SA}, ºC)
	DB temperature of the ambient air (T_{amb}, ºC)
	Liquid-line temperature (T_{LL}, ºC)
	Suction temperature (T_{suc}, ºC)
	Discharge temperature (T_{dischg}, ºC)
	DB temperature of the outgoing air from the condenser ($T_{air,ce}$, ºC)
	Evaporator temperature ($T_{sat,e}$, ºC)
	Condenser temperature ($T_{sat,c}$, ºC)
Pressure	Liquid-line pressure (P_{LL}, kPa)
	Suction pressure (P_{suc}, kPa)
	Discharge pressure (P_{dischg}, kPa)
Electrical	Compressor power consumption ($Power_{comp}$, W)

DB: Dry-Bulb, WB: Wet-Bulb

The response variable for the above features is the fault classes of eight types (including

'no fault' or 'NoF' as a fault class). The seven types of fault classes are already described

in Section 1.2.2.

Field data collected from a rooftop unit at a facility in Omaha, NE, was used to validate

the developed machine-learning classifier. This RTU was diagnosed with severe

refrigerant undercharge fault. Thus, the field validation of the developed machine-learning classifier was performed for this undercharge fault only.

The details about the simulated, lab, and field data are shown in Section 3.2.

1.4   Research Objectives

The research objectives for this study are summarized as follows,

- Obtain two reduced feature sets using a filter and a wrapper approach.

- Find the best cost-effective features suitable for a generalized machine learning model for packaged RTU for AFDD purposes. The best cost-effective features provide a better trade-off between the number of features and machine learning accuracy using the simulated testing set. The generalized machine learning model (classifier) fitted with the best reduced features will be obtained.

- Validate the developed model using three testing sets: simulated, lab, and field datasets.

- Test how the proposed model performs with varying fault intensities using simulated and lab testing sets and analyze how the model's predicting performance can be improved.

1.5   Structure of the Thesis

Chapter 2 will provide the literature review for FDD in VCRC.

Chapter 3 will provide the methodology for data preparation, fitting machine learning models, and predicting performance using different testing sets (simulated, lab, and field data sets).

Chapter 4 will provide the details about the field testing of the proposed machine-learning classifier.

Chapter 5 will provide the predicted performance results using different testing sets and make comparisons among them.

Chapter 6 will summarize the discussion, conclusions, and future research scopes in this area.

Chapter 2. Literature Review

2.1   Energy Savings Potential from HVAC System

In most industrial facilities, HVAC energy consumption accounts for significant portion

of the total energy bills. Also, according to the U.S. Energy Information Administration

(EIA) (2021), the commercial sector used about 155 billion kWh of electricity for cooling

in 2020, equaling 12% of total commercial sector electricity consumption. Therefore,

even a small improvement in the operational efficiency of these units can lead to

significant reductions in national energy use and carbon emissions. These improvements

often lead the facility in significant cost savings with lower payback period. As of

February 2023, Table 4 shows a summary of energy efficiency improvement measures

related to space conditioning of numerous facilities across the U.S., recommended by the

Industrial Assessment Center, a program under the U.S. DOE (*IAC: Search IAC*

Recommendations, n.d.).

Table 3: Cost saving measures related to HVAC system of facilities across the U.S.

Topics	Number of Recommendations	Cost Savings ($/year)	Implementation Cost ($)	Payback (year)
Space conditioning	8,896	$83.07 million	$139.34 million	1.7

Packaged rooftop units (RTUs) are those HVAC systems primarily found in commercial

buildings and manufacturing facilities. RTUs serve more than half of the commercial

buildings in the US (Schantz, 2015). As mentioned in Yuill and Braun (2013), these

unitary systems gain increased attention because of their broad applicability without

being adequately maintained, leading to higher fault incidence over time. As a result, the

appearance of faults in an RTU often leads to higher energy consumption (Li and Braun, 2007) and equipment wear. In most cases, the equipment wear is more significant than the economic impact because of the associated faults (Rossi and Braun 1996; Yuill and Braun 2017).

2.2   Soft Faults in Vapor Compression Systems

In this section, the soft faults discussed in Section 1.2.2 will be discussed in detail. These include the typical causes of those soft faults and the impact of those faults on the system performance.

2.2.1   Typical Causes of the Soft Faults

Soft faults typically develop due to the improper maintenance or installation of the equipment. Over time, the fault severity increases if the associated faults are left unattended. In addition, the other reasons of the seven types of soft faults are discussed in Table 5. In Table 5, the typical soft faults are represented in abbreviated forms (refer to Section 1.2.3).

Table 4: Typical reasons of soft faults

Fault	Typical Reasons
CA	Accumulation of dust, grease, pollen, and plant leaves, across the surface of the condenser coil.
EA	Foreign particles deposited across the evaporator coil, clogged air filters, stuck OA and RA dampers, and broken evaporator fan.
LL	Clogged refrigerant filter/drier located in between condenser and expansion device.
NC	Unintentional addition of non-condensable gases during installation or refrigerant charging.
OC	Improper overcharging during refrigerant charging
UC	Refrigerant leaks resulting from improper installation and old components.
VL	Loose sealing in suction and discharge valves. In newer compressors (scroll type), VL happens due to improper tolerance between flanks or leakage through clearance regions.

2.2.2   Impact of Soft Faults on System Performance

The system performance of the unitary VCRC is mainly dictated by three parameters, namely, Coefficient of Performance (COP), Cooling Capacity (CC), and Sensible Heat Ratio (SHR) (Yuill and Mehrabi, 2017). Soft faults in the VCRC can change the system performance by altering different thermodynamic and electrical variables, leading to degraded system performance. The effect of individual faults on system performance (for systems equipped with FXO) is discussed in Table 6.

Table 5: Effect of individual fault on system performance

Fault	Impacts on System Performance
CA	Increased condensing pressure reduces the capacity and increases the compressor power, decreasing the COP (Li and Braun, 2007a; Mehrabi and Yuill, 2017; Kim and Lee, 2021). However, a surprising study (Mehrabi and Yuill, 2019) for split systems with field-fouled condenser coil showed that the CA and condenser fouling are different fault types, and the system performance is not significantly degraded by condenser fouling, rather in some cases, the system performance improves.
EA	The reduced airflow across the evaporator coil decreases the capacity, leading to lower suction pressure, leading to lower refrigerant suction density, leading to lower compressor power, and thus the impact on COP is uncertain (Mehrabi and Yuill, 2017). In another study (Yang et al., 2007), EA causes reduced EER due to decreased capacity and increased fan power. They also found that the selection of air filters also dictates how much EER will be decreased.
LL	The liquid-line restriction causes the refrigerant mass flow rate to drop in FXO equipped system, leading to decreased capacity and decreased COP of the system (Du et al., 2015; Mehrabi and Yuill, 2017; Hu et al., 2021). These performance degradations become more severe at higher LL fault intensities.

Fault	Impacts on System Performance
NC	The addition of non-condensable gases increases compressor power consumption and decreases the COP (Kim et al., 2009; Hu and Yuill, 2022a, 2022b). As a result, compressor lifespan decreases.
OC	Compressor power consumption increases due to increased refrigerant mass flow rate resulting from overcharging (Mehrabi and Yuill, 2017b). In addition to that, liquid refrigerant might enter the suction line, damaging and reducing the lifetime of the compressor.
UC	Decreased charge level leads to reductions in both capacity and compressor power consumption, often leading to decreased COP (Farzad and O'Neal, 1991). The lower charge level leads to increased runtime of the compressor. The most important effect is probably the increased compressor temperature, causing more metal-to-metal wear, decreasing its lifetime.
VL	The VL results in decreased discharge pressure and increased suction pressure, leading to less suction superheat (Mehrabi and Yuill, 2017). If the suction superheat is too low, compressor lifetime decreases due to handling liquid refrigerant.

2.2.3   Impact of Faults on Energy Consumption

Previous studies estimated the effect of common faults on system performance. For instance, Roth et al. (2004) found more than 100 faults in commercial buildings, and based on three performance factors (AEC_{re}, $Frequ_{fault}$, and $Degrade_{fault}$), they identified

13 key faults (including HVAC system faults) that have a significant impact on the building energy consumption. Altogether, the 13 key faults had an additional commercial building energy consumption in the range of 0.35 to 1.7 quads. According to Wiggins and Brodrick (2012), HVAC system faults introduced between 1% and 2.5% of total commercial building energy consumption. They presented several kinds of HVAC FDD devices that can reduce labor costs, maintenance costs, electrical demand, and wasted energy. Fernandez et al. (2018) estimated national energy savings in the commercial building through better control, fault elimination, and better sensing. They found 2.76 quads of primary energy savings, performed across multiple building types and climates in the US.

To meet the net-zero carbon emission goal by United Nations by 2050, there should be a 45% reduction of global greenhouse gas emissions by 2030, from 2010 levels (*Net Zero Coalition | United Nations*, n.d.). To achieve that, the commercial sector can play a major role by reducing the national energy and demand requirements. Thus, eliminating soft faults from the packaged rooftop units in U.S. commercial buildings should be an integral part as far as national energy and demand reductions are concerned.

2.3   Details on Fault Detection and Diagnosis (FDD)

As a supervisory control feature, Fault Detection and Diagnosis (FDD) in HVAC systems aim to find subtle faults, determining their type, intensity, and location before they cause major equipment malfunctions (Yuill and Braun 2013; Li and Braun 2007). In addition to Fault Detection and Diagnosis, there can be an additional step called Fault Evaluation where it informs whether the detected fault needs to be eliminated immediately or not

(Isermann, 1984). Figure 3 shows a simplified Fault Detection, Diagnosis, and

Evaluation, adapted from Rossi and Braun (1997). In the fault detection step, it is

identified whether the system is faulty/unfaulty based on the residuals of measured and

expected variables. In the fault diagnosis step, upon confirming the presence of fault, the

specific type of fault is diagnosed based on the directional change of measured and

expected variables. In the fault evaluation step, it is decided whether the fault should be

repaired immediately or not. Depending on the type of FDD methods, fault detection and

diagnosis can be performed separately or simultaneously.

Figure 3: Simplified fault detection, diagnosis, and evaluation

Three kinds of FDD are used for HVAC systems (Katipamula and Brambley, 2005a,

2005b; Yang et al., 2014). These are quantitative, qualitative, and process-history based.

The first two approaches are also known as knowledge-based approaches because they

require sound knowledge of the associated systems. In the following sections knowledge based and process-history based approaches will be discussed in detail.

2.3.1 Knowledge-Based FDDs

Quantitative models are based on detailed physics and engineering principles and are thus quite complex, often requiring greater computational power and physical knowledge. Qualitative model-based FDDs need prior knowledge to develop different rules to diagnose various faults. Since both quantitative and qualitative methods require sound knowledge of the associated systems, they are also known as knowledge-driven approaches. For instance, Rossi and Braun (1997) proposes a rule-based FDD method for VCRC which is a combination of both quantitative and qualitative approaches. Their method had three major steps, pre-processing, fault detection, and fault diagnostics. In pre-processing step, they used a quantitative/steady-state model (developed by Rossi, 1995) that provides the expected (unfaulty) thermodynamic quantities by solving the mass, energy, and momentum balances for any input air conditions. In the fault detection step, they identified the presence of faults by comparing the expected and actual thermodynamic measurements. In fault diagnostics step, they pin-point the exact fault type based on the directional change in thermodynamic properties for that particular fault type. In both the later steps, they used some specific set of qualitative rules applicable to the systems similar to the system they studied, i.e., the rules were not generalizable to be used for all rooftop units.

Chen and Braun (2001) studied two simple rule-based methods for rooftop units equipped with thermostatic expansion valves. One set of simple rules is shown in Table 7. As we

can see, to detect and diagnose the faults using these fixed set of rules, it would require

measuring evaporating and condensing pressures, in order to get the saturated

temperatures (T_{evap}, T_{cond}) for those pressures. Measuring the pressures are costly and

intrusive to the system, thus it might not be suitable. In addition, the rules listed in Table

7 may not work well for all operating conditions of the RTU. As a disadvantage of the

knowledge-based FDD approaches, Behfar et al. (2019) showed that the rule-based

methods are less reliable when the performance variable cannot be kept steady under

unfaulty operations.

Table 6: Simple rules from Chen and Braun (2001)

Fault Type	Performance Indices	Sensors and Locations	No-Fault Range	Fault Range
Evaporator fouling	$(T_{ra} - T_{evap})$ will be abnormally high	T_{evap}, T_{ra}	32-33°F	>34°F
Condenser fouling	$(T_{cond} - T_{amb})$ will be abnormally high	T_{cond}, T_{amb}	19-20°F	>21°F
Liquid line restriction	ΔT_{ll} will be abnormally high	$T_{ll,in}, T_{ll,out}$	1-2°F	>3°F
Refrigerant leakage	T_{sc} will be abnormally low			<12°F
Refrigerant overcharge	T_{sc} will be abnormally high	T_{sc}	13-14°F	>15°F
Noncondensable gas	T_{sc} will be abnormally high			>15°F
Compressor leakage	$(T_{ra} - T_{evap})$ will be abnormally low	T_{evap}, T_{ra}	32-33°F	<31°F

While validating the gray-box models for unitary air-conditioning systems proposed by

Cheung and Braun (2013a, 2013b), Yuill et al. (2014b) showed the fault prediction

performance of the existing commercial FDD protocols in terms of different frequency of

occurrence rates, like False Alarm Rate (FAR), Misdiagnosis Rate (MR), and Missed

Detection Rate (MDR). FAR is the percentage of actual unfaulty cases incorrectly

predicted as faulty; MR is the percentage of total faulty cases incorrectly predicted as a

fault class different from the actual fault class, and MDR is the percentage of total faulty

cases incorrectly predicted as unfaulty. The lower values of these frequency of

occurrence rates represent better performance from an FDD protocols. It is unclear if

those protocols followed quantitative/qualitative/process-history based FDD approaches.

The results (Figure 4) from Yuill et al. (2014b) showed unsatisfactory performance in terms of FAR in all except one FDD (FDD E) protocol. The MR and MDR also showed variable performance across different FDD protocols. Figure 4 (b,c) shows MR and MDR, with respect to different fault severity cases. In these Figures, FDD D shows lower MR and MDR values than the other protocols. However, FDD D was a detection only protocol (cannot diagnose fault type, hence MR = 0), and it could not pin-point which fault is present in the system. Also, although FDD D had a lower MDR value, it showed very high FAR values in Figure 4 (c). In addition, some of the protocols are not suitable for specific scenarios, marked by predicting "no response" or "no diagnosis".

(a) FAR

(b) MR

(c) MDR
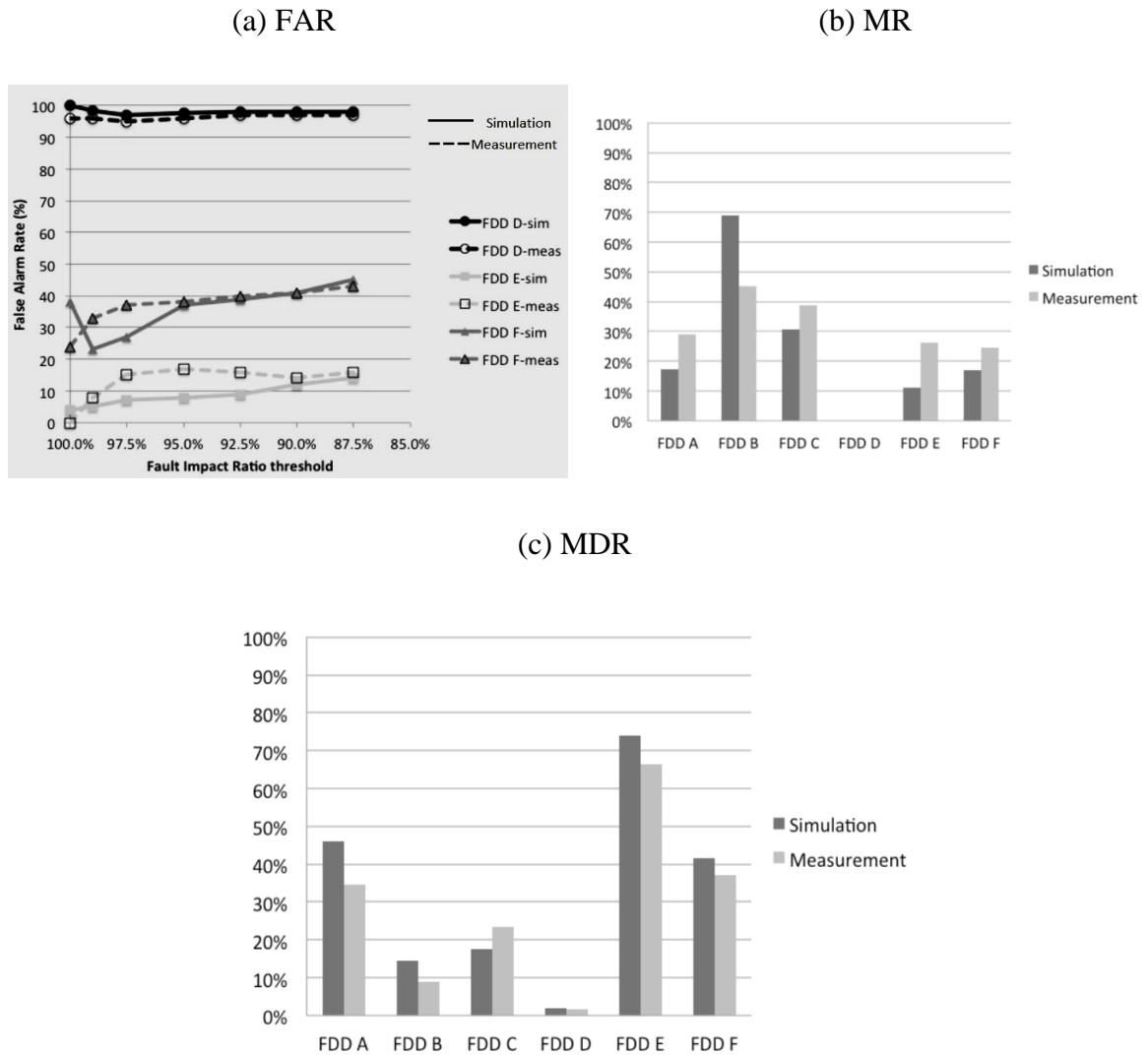
Figure 4: Performance results from 6 FDD protocols (Yuill et al., 2014b)

2.3.2   Process History Based FDDs

In contrast to the two knowledge-based approaches, the process-history-based approach

requires collecting data with relevant features for different operating conditions to fit

either a black box or gray box model that can predict faults when testing with a new data

set. The model from the process-history-based approach is developed based on pattern

recognition and intra-attribute inherent correlations (Iyengar et al., 2018). This method requires plenty of useful data and is thus also called the data-driven method. Examples of data-driven FDD studies are Ebrahimifakhar et al. (2020) and Chen et al. (2022), who developed machine learning models using full available features and reduced set of features, respectively (details in next paragraph). According to Kim and Katipamula (2018), the data-driven approach is best-suited for a complicated physical system where the system's performance cannot be fully explained using theoretical knowledge. In addition, the former two approaches (quantitative and qualitative) might be suitable for detecting some specific faults. In contrast, the data-driven method can detect and diagnose most of the common faults if relevant data are available. The potential of data-driven FDD approach was reviewed in details by Mirnaghi and Haghighat (2020). They reviewed previous data-driven studies related to supervised, semi-supervised, unsupervised, and hybrid algorithms. They found that the data-driven approaches are a more promising FDD process for large-scale HVAC systems than the quantitative and qualitative approaches.

Several FDD studies have used data-driven approaches to unitary air conditioning systems. For instance, Ebrahimifakhar et al. (2020) compared the predicted performance of typical soft faults by fitting several supervised machine learning models using all the 15 features of simulated data for an RTU (equipped with FXO) from Cheung and Braun (2013a, 2013b). Using the simulated testing set, they found that the radial kernel-based SVM outperformed all the other machine-learning classifiers. However, the predicting performance regarding false alarm rate (FAR) was unsatisfactory (100%) since their trained model had few unfaulty cases. Even after increasing their unfaulty cases using an

oversampling technique, 12 of the 17 unfaulty cases were still flagged as false alarms (refer to the 'NF' row in Figure 5). In practical AFDD applications, false alarms are the costliest error.
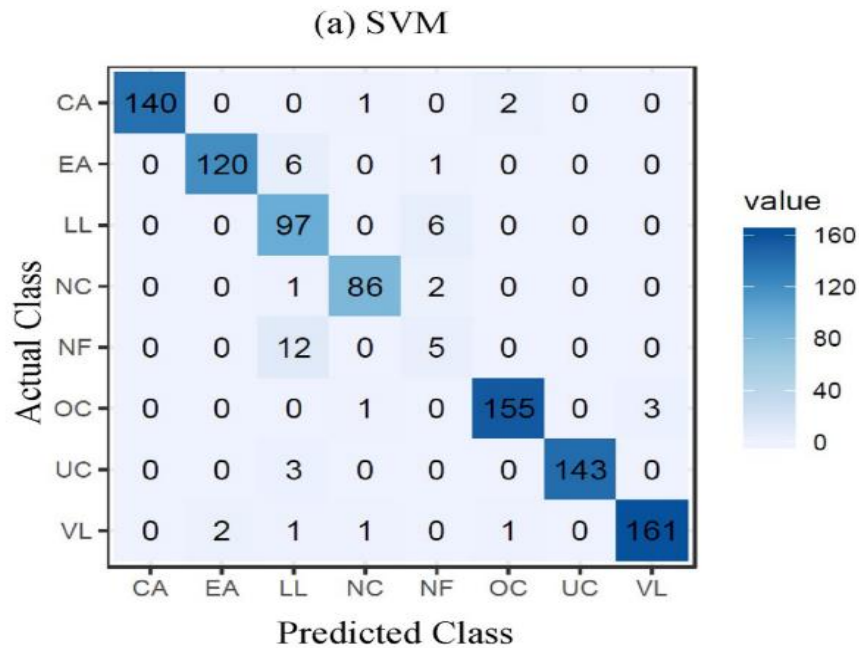
Figure 5: Confusion matrix from Ebrahimifakhar et al. (2020)

To find how the reduction of features affected the predicting performance of SVM classifiers fitted using the simulated data for a split system (equipped with TXV), Chen et al. (2022) analyzed two separate classifiers using two sets of 10 features, one from sequential backward selection (SBS), another from manual selection. They have found that, when tested with the same system, the reduction of 5 features has little impact on the overall predicting performance of the classifiers. As a continuation of their work, to find how the SVM model is generalizable to be used when the model is trained with the simulated data of a split air-conditioner but tested with the simulated data from different

split systems, Chen *et al.* (2023) worked with the same two sets of 10 features (using

Sequential Backward Selection (SBS) and manual selection for classifiers C1 and C2,

respectively) in addition with the full 15 features (for classifier C3). The list of variables

in those two reduced sets of 10 features is shown in Table 7.

Table 7: Feature selection from Chen et al. (2022, 2023)

Features	Selected variables
SBS* features for C1	T_{SA}, WB_{SA}, P_{LL}, T_{suc}, P_{dischg}, T_{dischg}, $T_{air,ce}$, $T_{sat,e}$, WB_{RA}, T_{amb}
Manually selected features for C2	T_{SA}, T_{LL}, T_{suc}, T_{dischg}, $T_{air,ce}$, $T_{sat,e}$, $T_{sat,c}$, Power, T_{RA}, T_{amb}

*Sequential Backward Selection

From the above study of Chen *et al.* (2023), although the systems were similar (R410A

refrigerant with TXV), the classifiers performed badly when tested with a different

system (using simulated data) than the system with which the classifier is trained. Figure

6 demonstrates one such confusion matrix from classifier C2, showing heavily biased

towards predicting CA (refer to column CA in Figure 6). Thus, future steps from their

study proposed training the classifier with multiple similar systems with increased

unfaulty conditions. Also, the obtained feature selection results are only for split system

air conditioners, requiring similar studies for packaged rooftop units. In addition to that,

the validation of the machine-learning classifiers using lab/field data needs to be

performed before using the classifier for practical applications. Also, it is necessary to

know how the performance of the classifier varies with the fault intensity.

Figure 6: Confusion matrix using Classifier C2 from Chen et al. (2023)

Albayati et al. (2022) studied three different fault classification methods for predicting five fault classes (4 soft faults and one unfaulty type) for a field-tested RTU with FXO as an expansion device, taking 24 selected features. Their studies considered both single fault and simultaneous faults occurring at a single time. The first method (capable of detecting single/simultaneous faults) proposed an SVM-supervised ML model with an average accuracy of 93.5%, with unfaulty prediction accuracy being the lowest (80%) among all five classes. This low accuracy for unfaulty classes happened due to minority unfaulty classes in their trained model. The second method (for single/simultaneous faults) proposed a semi-supervised ML model with an average accuracy of 94.9% and an increased unfaulty prediction accuracy of 91.8%. The third method (for single faults) was also a semi-supervised model that had an accuracy of up to around 98%, having a False Alarm Rate (FAR) of 13.21% (Figure 7). Finally, they proposed a trade-off method that

took either Method 1 or 2 based on the individual highest-class accuracy, which provided an average accuracy of 95.7%. However, their proposed models are not generalized to be applied for detecting faults for all RTUs and thus require further studies for a generalizable ML model.
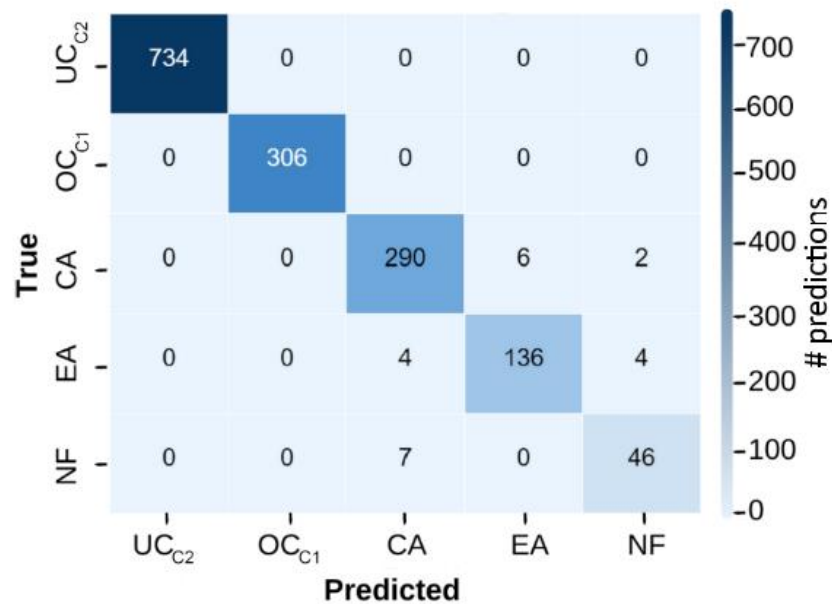
Figure 7: Confusion matrix using Method 3 from Albayati et al. (2022)

2.4   Feature Selection Study

Feature Selection means finding the suitable set of inputs for fitting the machine-learning classifier. In this section, the details about the Feature Selection will be provided.

2.4.1   Need for Feature Selection

Commercial sectors often have access to a large amount of data related to the HVAC system thanks to the developed Building Energy Management System (BEMS)

technology. Data from Building Automation Systems can be used in FDD and thus has significant potential to improve the energy efficiency in the HVAC system (Granderson et al., 2020). However, access to more data means it is required to perform rigorous data preprocessing before the data is suitable for fitting machine learning classifiers for FDD purposes. The field collected data from BEMS often has both relevant and irrelevant features with the presence of outliers, making it more likely to overfit the machine learning classifier if accurate data preprocessing is not done (Jung and Sundstrom, 2019). The first step in any feature selection process involves discarding noisy, non-informative, and highly collinear features (i.e., redundant features) from the raw dataset (Llobet et al., 2006; Chandrashekar and Sahin, 2014). Previous FDD studies in HVAC systems also practiced finding the minimum number of features (or, sensors) for detecting faults. For example, Rossi et al. (1995) developed a model-based FDD system for detecting refrigerant leakage from an experimental RTU, and their sensitivity studies found three sensors (or, features) measuring superheat, subcooling, and compressor discharge temperatures, sufficient for detecting the refrigerant leakage. In addition, since features for the data-driven FDD approach in HVAC systems mainly indicate a list of associated sensors, incorporating more features while developing a data-driven FDD classifier means that the cost of sensors will also be increased (Kim and Braun, 2020). Sometimes, it can be challenging for the manufacturer/facility to bear the increased sensor costs. So, it is necessary to perform proper feature selection before fitting the machine learning classifier for FDD purposes.

2.4.2   Classification of Feature Selection

There are mainly two kinds of feature selection approach, filter, and wrapper (Kohavi and

John, 1997; Yan et al., 2018). As shown in Figure 7, Filter approach can be further

subdivided into ANOVA and Pearson correlation. The filter approach using Pearson

correlation (used in this study) is independent of machine learning algorithms and

depends on the correlation among the features to find the effective set of features. The

Pearson correlation approach filters out the most uncorrelated set of features based on a

correlation cut-off value (detailed discussion in Section 3.9). On the other hand, the

wrapper approach can be further subdivided into forward, backward, and stepwise

eliminations (Figure 8). In contrast to filter approach, feature selection using the wrapper

approach is based on a training classifier with a specific machine-learning algorithm and

is thus computationally expensive (Langley, 1994). For example, backward elimination

starts with full features and recursively eliminates the less important features. It does that

based on a machine-learning algorithm informing which features are adding less

knowledge to the trained algorithm, marked by accuracy.
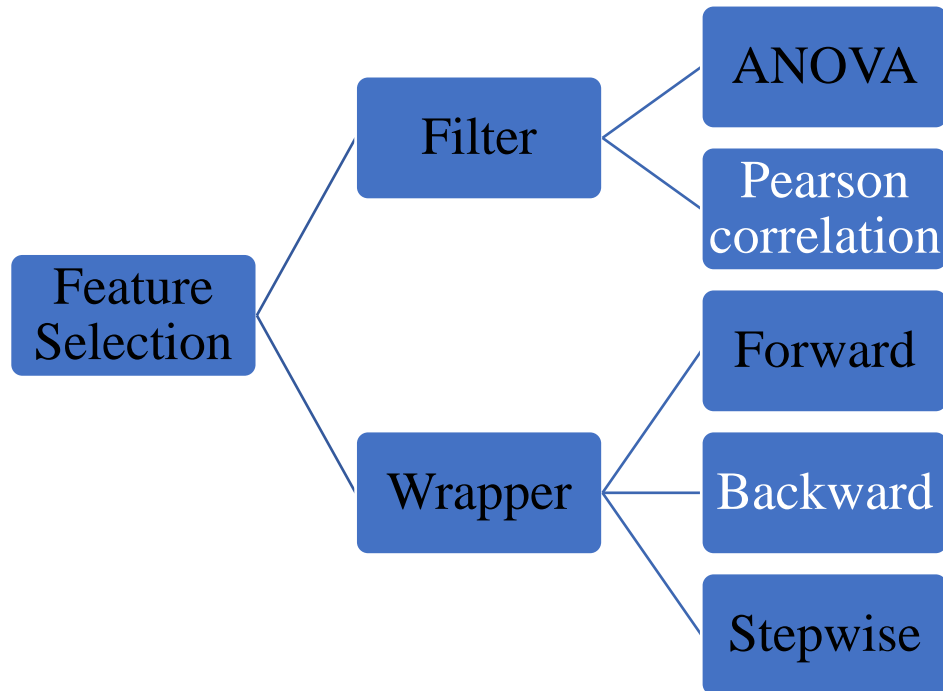
Figure 8: Feature selection classification

In this study, the filter approach using Pearson correlation was found to be effective to get the reduced set of features. To validate the reduced set of features using the filter approach, a separate set of features obtained from a wrapper approach (backward elimination) was considered.

2.5   Focus of this Study

Any successful data-driven classifier for fault prediction in an HVAC system requires input data consisting of all possible combinations of operating conditions, fault types, and fault intensities (Yuill and Braun, 2016). Also, control over this wide range of combinations is necessary so that the FDD classifier can be evaluated under a wide range

of conditions (Yuill et al., 2014a). However, obtaining experimental data consisting of all the required combinations is challenging since it is labor-intensive, time-consuming, and requires technical expertise. One solution can be to perform experiments with labeled and unlabeled faults and, using a semi-supervised approach afterward, labeling the unlabeled faults can be performed based on the labeled data (Fan et al., 2021; Li et al., 2021). However, this approach still requires performing experiments to collect the data. Another solution to get the necessary inputs for the machine learning AFDD classifier is using simulated data from gray-box models (Cheung and Braun, 2013a, 2013b). This approach can simulate the required data and eliminate the need for fault labeling, as the gray-box model can generate the inputs for a data-driven FDD approach based on the operating conditions, fault type, and fault intensity. If the latter solution is performed, then this will be a combination of knowledge-based and data-driven FDD methods. A review study conducted by Zhao et al. (2019) concluded that new methods combining both knowledge-based and data-driven approaches are needed to have added advantage of both of these approaches.

None of the above studies considered investigating the cost-effective set of features to develop a generalized machine classifier for packaged RTUs. To develop a generalized machine learning classifier with a reduced set of features for packaged rooftop units, this study considered three separate simulated data from three RTUs from Cheung and Braun (2013a, 2013b). The developed classifier was first tested with unseen simulated data. Then, the developed classifier was validated using an existing lab dataset. Finally, the developed classifier was validated with field measurement data. The possible

comparisons among different predicting performances of the simulated, lab, and field-

measured testing sets were performed, and appropriate conclusions were made.

Chapter 3. Methodology

In this chapter, the methodology related to data pre-processing, feature selection, and fitting a machine learning classifier will be discussed in detail.

3.1   Machine Learning (ML) Workflow

A machine-learning algorithm is a mathematical model fitted using suitable data. It can be either classification or regression type. In a classification-based ML, the response variable has categorical types (classes), whereas, in a regression-based ML, the response variable has continuous/quantitative variables. The ML algorithm in this study is of classification type, thus the corresponding ML algorithm is termed as ML classifier.

A typical machine-learning workflow is shown in Figure 9.
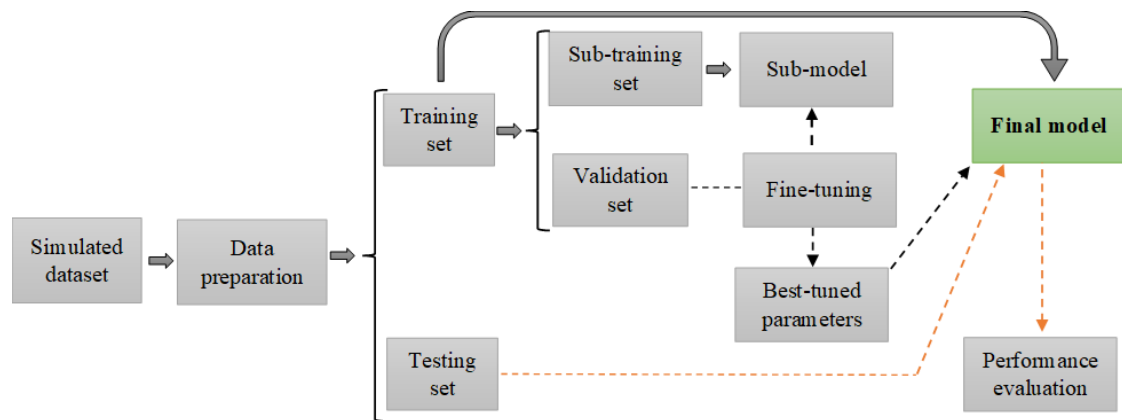
Figure 9: Typical machine learning workflow

As shown in Figure 9, the simulated dataset first goes through a data preparation stage where relevant features (columns) and appropriate rows are selected. The data

preparation also includes shuffling the entire dataset and randomly splitting it into two subsets: training and testing sets. Next, depending on the validation approach, the training set is divided into sub-training and validation sets (Figure 9). For a machine learning algorithm, using suitable sub-training and validation sets, different sub-models corresponding to different combinations of tuning parameters are checked (grid search discussed later) before finding the best combination (that gives maximum accuracy) of the tuning parameters for the final model. Finally, the performance of the final model is analyzed using the testing set. In this study, 80% of the simulated data was randomly selected for training the final model, and the rest 20% was used for testing the model.

3.1.1   Support Vector Machine

The Support Vector Machine (SVM) is elaborately discussed in Brereton and Lloyd (2010) with examples. This section will discuss the basic methodology of SVM based on their work.

Support Vector Machine is a machine-learning algorithm that can be explained using a binary classification problem. In a binary classification algorithm, the observations in the training data are separated into two classes. Before learning how the two classes are separated using SVM, it is necessary to know two additional classifiers called Maximum Margin Classifier (MMC) and Support Vector Classifier (SVC).

Maximum Margin Classifier is a classifier that linearly separates the two classes (in the training data) using a hyperplane. A hyperplane can be considered as a threshold boundary (represented by a line in 2-D space) that separates the two classes. The best hyperplane is selected in such a way that maximizes the margin. For each class, a margin

can be defined as the perpendicular distance between the hyperplane and the closest

observation for that class. The two margins for the two classes are equal. In MMC, the

hyperplane completely separates the two classes, i.e., no misclassification occurs.

Practical binary classification problems often have overlapping observations. Thus, while

training the classifier, it may not be possible to completely separate the two classes. In

such cases, linearly separating the observations into two classes would require some

observations to be misclassified. A classifier with some misclassified observations in the

training dataset is called a Support Vector Classifier (SVC). Figure 10 shows an example

of SVC for a two-featured (Variable 1 and Variable 2) classification problem. In this

figure, the blue circles represent class A, and the red diamonds represent class B.
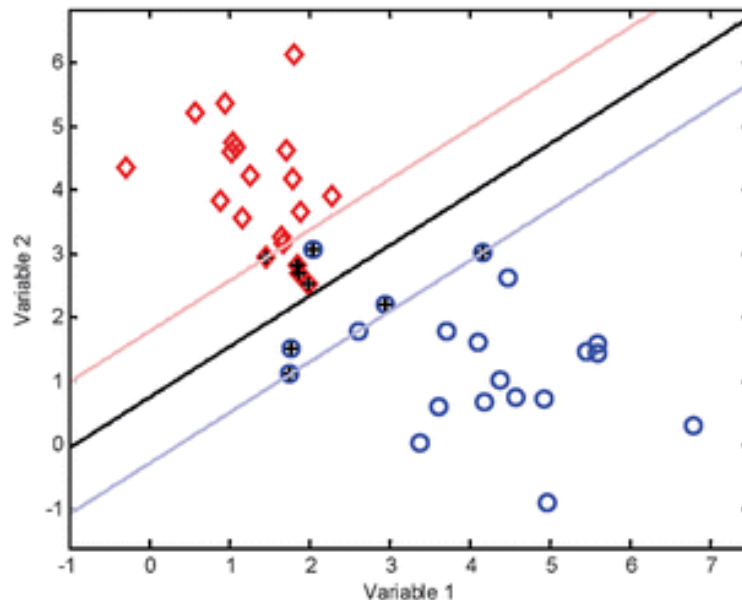
Figure 10: Understanding Support Vector Classifier (Brereton and Lloyd, 2010)

In Figure 10, the black line represents the hyperplane. Red and blue lines are the two

marginal lines, for class B and A, respectively. As we can see, since the observations

from two classes overlap, it is not possible to completely separate the two classes using a line (hyperplane). Thus, the hyperplane is chosen in such a way that misclassifies one point of A as B. Ideally, the observations for a class should be lying beyond the marginal line for that class. Support vectors are the points that lie on the two marginal lines and points lying on the wrong sides of the marginal lines. These points are marked by the shapes with a '+' inside. The hyperplane location largely depends on these support vectors. Now, when a new point is fed into the above SVC, the classifier determines which side of the hyperplane the new point falls on and predict the class accordingly.

In many practical classification problems, the two classes cannot be linearly separated using a hyperplane. Instead, it requires a non-linear boundary to separate the two classes. This is called a kernel trick to separate the two classes. The kernel can be of polynomial or radial type. The idea is to create new feature(s) from the existing features/predictors and non-linearly separate the observations into two classes at higher dimensions. When the two classes are separated by a non-linear boundary (at higher dimensions) after creating new features from the existing feature set, this is called a Support Vector Machine.

The classification problem studied in this dissertation is a higher-dimensional problem with up to 16 dimensions (for 15 thermodynamic and single categorical feature), and intuitively, it is not possible to separate the fault classes linearly using a hyperplane discussed above. In addition, previous studies showed that the SVM with radial kernel works well for classifying faults in VCRC (Ebrahimifakhar et al. (2020) and Chen et al. (2022)). Thus, this study will follow SVM with radial kernel as the underlying machine-learning algorithm for all classifiers. Since the classifiers studied in this dissertation have

significantly higher dimensions (up to 16), it is not possible to graphically show the classifiers.

The above description provides the classification as a binary problem (having two classes). However, in this study, there are 8 fault classes. Thus, the multi-class SVM classification needs to be converted to suitable binary classifiers. Multi-class problems can be solved binarily in two ways, One-Vs-One or, One-Vs-All. One-Vs-One is selected in this study because previous study by Hsu & Lin (2002) found this as more suitable for practical classification. In One-Vs-One classification, since the classifier in this study has 8 classes, there can be 8C2 or 28 pairs of faults. Thus, 28 SVM classifiers (with radial kernel) will be fitted using the appropriate training dataset. For a new testing case, the predictions from all 28 SVM classifiers were gathered. The final prediction is the class that has the majority of the predictions. In R programming software (R Core Team, 2022), the package "e1071" is used in this study to perform One-Vs-One classification. In contrast to One-Vs-One classification, One-Vs-All classification fits all the classifiers for a class Vs the remaining classes.

3.2   Study Data

As mentioned in Section 1.3, three types of datasets were used in this study. The simulated dataset, consisting of three rooftop units, was obtained from a gray-box model developed by Cheung and Braun (2013a, 2013b). An existing lab dataset consisting of the same three rooftop units was used to validate the developed machine-learning classifier. Field data collected from a rooftop unit at a facility in Omaha, NE, was used to validate the developed machine-learning classifier. While tested during the late morning hours on

July 14, 2023, circuit 1 of this RTU has shown significantly lower subcooling and higher

superheating values, indicating severe refrigerant undercharge fault in the system (details

in Chapter 4). In this study, the field validation will only be done considering this UC

fault. Since this study proposes a reduced set of 9 features (8 measurable, 1 categorical)

to detect the common soft faults in rooftop units, 8 measurements were logged at specific

intervals from the above faulty system. During the time between 14:38 to 18:13 on the

above day, based on the steady-state operating conditions (marked by steady T_{suc} and

T_{dischg}) of circuit 1 of the faulty RTU, 24 steady-state cases have been considered by

taking a 5-minute average value for each of the 8 measurements. The details about the

field testing are mentioned in Chapter 4. Table 8 shows all the RTU information for

simulated, lab, and field data. The sample simulated and lab data can be observed in

Appendix G and H, respectively.

Table 8: Summary of simulated and experimental data

RTU System	Refrigerant	Capacity (kW)	Compressor Type	Expansion Device Type	Data Type	Number of Cases								Ambient Temperature Range (°C)	RA Temperature Range/Values[2] (°C)
						CA	EA	LL	NC	NoF[1]	OC	UC	VL		
S1	R410A	10.55	Scroll	FXO	Simulated	317	275	193	251	39	422	259	357	15.56-48.89	23.89, 29.44
					Experimental	6	21	0	0	24	12	25	0	19.49-51.74	19.73-27.18
S2	R407C	17.59	Scroll	FXO	Simulated	220	149	99	126	38	106	78	112	15.56-48.89	23.89, 29.44
					Experimental	8	19	0	0	17	12	15	0	19.35-46.46	19.60-27.27
S3	R22	10.55	Reciprocating	FXO	Simulated	264	170	315	162	46	52	148	179	18.33-46.11	21.11, 25, 28.89
					Experimental	36	26	34	0	39	0	34	33	15.34-37.76	21.12-27.91
S4	R22	43.96	Scroll	TXV	Field	0	0	0	0	0	0	24	0	31.46-35.63	23.81-25.07

[1]NoF = No Fault
[2]For simulated data, unique values are shown. For lab and field data, ranges of values are shown.

The S4 mentioned above refers to circuit 1 of the field-tested RTU. Although S4 is

equipped with TXV, it can be treated as an FXO-equipped system because the TXV

could remain fully open to meet the cooling demand since it has severe UC fault present.

Table 9 shows the descriptive statistics of the 15 features for simulated and experimental data. For the field data, since this study proposes a machine learning classifier with 8 measurable and 1 categorical feature, the descriptive statistics are shown for the 8 variables, and the other 7 variables are showing NA values in Table 9.

Table 9: Descriptive statistics (mean μ, and standard deviation σ) for each system

Features	Simulated Data						Lab Data						Field Data	
	S1		S2		S3		S1		S2		S3		S4	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
T_{RA} (°C)	26.5	2.8	26.8	2.8	25.7	3.0	26.6	0.8	26.6	0.9	23.6	0.9	24.4	0.3
WB_{RA} (°C)	18.0	4.3	18.5	4.5	18.2	4.3	17.5	2.6	18.1	1.9	15.8	1.2	14.1	0.2
T_{SA} (°C)	16.1	3.5	17.0	3.3	17.9	4.0	15.3	2.5	15.0	3.0	13.5	1.3	13.0	0.2
WB_{SA} (°C)	12.9	5.1	13.5	5.0	14.4	4.9	12.4	3.2	12.7	2.8	11.3	1.4	NA	NA
T_{amb} (°C)	34.0	10.1	33.8	10.0	32.5	9.6	34.8	7.2	35.3	6.8	25.8	4.9	33.3	1.1
P_{LL} (kPa)	2876	707	2185	572	1768	402	2878	495	2143	362	1388	198	NA	NA
T_{LL} (°C)	40.5	12.0	40.3	10.4	43.1	11.5	41.1	7.4	41.1	7.5	34.5	6.4	40.7	0.5
P_{suc} (kPa)	1054	147	592	84	590	109	1007	115	585	62	456	38	NA	NA
T_{suc} (°C)	15.0	6.1	14.9	7.2	15.7	5.5	14.2	5.1	14.6	6.7	11.9	4.1	13.2	0.2
P_{dischg} (kPa)	3038	686	2503	614	1909	367	2966	486	2437	367	1529	190	NA	NA
T_{dischg} (°C)	80.7	16.9	89.5	12.4	101.5	22.7	80.2	11.6	91.2	10.3	90.2	4.7	82.2	0.8
$T_{air,ce}$ (°C)	44.7	10.6	46.3	11.0	38.5	9.2	45.3	7.1	46.7	6.7	32.2	4.9	NA	NA
$T_{sat,e}$ (°C)	10.2	5.2	8.5	4.7	4.8	6.7	7.6	4.2	7.2	3.5	5.5	2.0	NA	NA
$T_{sat,c}$ (°C)	47.8	10.0	56.8	10.5	45.6	9.2	46.9	7.2	48.8	7.3	39.4	5.2	NA	NA
$Power_{comp}$ (W)	2788	720	5917	1433	2925	643	2524	518	5192	790	2825	181	8863	77

NA: Not applicable since not measured.

3.3   Normalizing Compressor Power and Discharge Temperature

Out of the 15 measurable features described in Table 2, $Power_{comp}$ significantly varies depending on the cooling capacity of an RTU. For instance, the field-tested RTU (S4) has a higher mean $Power_{comp}$ (Table 9) than that of S1, S2, and S3, since S4 has higher cooling capacity. Thus, to make a generalizable classifier for any rooftop unit, it is

necessary to normalize the compressor power in such a way that accounts the variation of compressor power among different RTUs. To do that, this study normalizes the actual Power$_{comp}$ based on the unfaulty Power$_{comp}$ and used those normalized values as a feature instead of the actual Power$_{comp}$. Mathematically, normalized compressor power,

$$Power_{comp,norm} = \frac{Power_{comp}}{Power_{comp,unfaulty}} \quad \ldots \quad (3.1)$$

Where,

Power$_{comp}$ = actual compressor power

Power$_{comp, unfaulty}$ = unfaulty compressor power for the corresponding compressor

Thus, we need Power$_{comp, unfaulty}$ to normalize the actual compressor power based on Equation 3.1. The compressor power depends on the compressor lift (difference of suction and discharge pressures) and mass flow rate. Since measuring compressor lift and mass flow rate are challenging in practical AFDD applications, in this study, compressor performance data was utilized to get the Power$_{comp, unfaulty}$ for a specific compressor model with a specific refrigerant type.

Now, to get the unfaulty compressor power for a rooftop unit, mean power value from the compressor performance data from the specific compressor model was utilized (refer to Appendix I). The compressor manual shows the unfaulty rated compressor power values in tabular form, depending on the evaporating ($T_{sat,e}$) and condensing temperature ($T_{sat,c}$) of a refrigerant. To get the mean unfaulty compressor power value for a refrigerant, mean values of $T_{sat,e}$ and $T_{sat,c}$ are necessary. For a refrigerant, the mean unfaulty values of $T_{sat,e}$ and $T_{sat,c}$ are determined based on the corresponding mean unfaulty values from the

simulated training data. Therefore, the mean unfaulty compressor power for those specific $T_{sat,e}$ and $T_{sat,c}$, is interpolated from the compressor performance data using the specific compressor model. It is true that the mean unfaulty values for the $T_{sat,e}$ and $T_{sat,c}$ in the corresponding training data does not contain all the operating conditions, still this approach is suitable instead of measuring costly compressor lift and refrigerant mass flow rate.

Table 10 shows the mean unfaulty compressor power for the four RTU systems.

Table 10: Mean unfaulty power values

System	Refrigerant	Compressor Model	Mean $T_{sat, e}$ (ºC)	Mean $T_{sat, c}$ (ºC)	Power$_{comp, unfaulty}$ (W)
S1	R410a	Copeland ZP32K3E-PFV	11.22	49.16	2,720.35
S2	R407c	Alliance (not found), Closest model: Copeland ZRD61KCE-PFV	7.33	53.83	5,639.18
S3	R22	Copeland CRH3-0275-TFD	6.97	46.33	3,106.93
S4	R22	Copeland ZR16M3E-TWD	6.97	46.33	11,642.65

Therefore, to get the normalized compressor power, the actual compressor power is divided with the corresponding unfaulty compressor power mentioned in Table 10.

Table 9 also shows that the mean T_{dischg} significantly varies across the three refrigerant types in the simulated data (80.71°C Vs. 89.52°C Vs. 101.47°C). Therefore, to get a generalizable classifier, just like the normalized compressor power, the actual T_{dischg} needs to be normalized with the unfaulty T_{dischg}. Again, to get the unfaulty T_{dischg}, mean

discharge temperature of the simulated training set for each refrigerant type is considered.

Mathematically, normalized discharge temperature,

$$T_{dischg,norm} = \frac{T_{dischg}}{T_{dischg,unfaulty}} \qquad \dots \qquad (3.2)$$

Where,

$T_{dischg, unfaulty}$ = Mean unfaulty discharge temperature for the corresponding

refrigerant

Table 11 shows the $T_{dischg, unfaulty}$ for each of the three refrigerant types, computed using

simulated training data.

Table 11: Mean unfaulty discharge temperature

Refrigerant	$T_{dischg, unfaulty}$ (^{o}C)
R410a	81.62
R407c	89.59
R22	95.61

Therefore, to get the normalized discharge temperature, the actual discharge temperature

is divided with the corresponding refrigerant's unfaulty discharge temperature mentioned

in Table 11. The normalized discharge temperature will be used as a feature instead of the

actual discharge temperature.

3.4   Standardization of the Features

After splitting the whole data set into training, and testing sets, the normalization of

$Power_{comp}$ and T_{dischg} were done based on the methods described in Section 3.3. Then, it is

necessary to standardize the training set's features so that each feature has a mean of 0 and a standard deviation of 1. For any feature X, the standardization of each value of X is done by using the following well-known standardization formula,

$$z = \frac{(x-\mu)}{\sigma} \qquad ... \qquad (3.3)$$

Where,

z = Standardized output corresponding to a value x from Feature X

x = A value of feature X

μ = Arithmetic mean of feature X

σ = Standard deviation of feature X

Next, the testing set standardization should be performed based on the mean and standard deviation of the corresponding feature from the training set. In other words, the testing set should be standardized after splitting the whole dataset into training and testing sets. If standardization is performed on the entire dataset and then splitting is performed, then there exists a leakage of information from the testing set to the training set, which is strictly forbidden. In this study, this phenomenon is avoided in the testing set.

3.5   Validation Approach for Fine-Tuning the ML Classifier

The performance of any ML algorithm depends on the tuning parameters associated with that algorithm. As discussed in Section 3.1.1, this study will use a radial kernel-based

SVM classifier to fit the ML classifier since previous studies showed satisfactory

performance with this algorithm. Two tuning parameters related to this classifier are cost

(C) and gamma. Since both of these parameters are continuous, there can be an infinite

number of combinations possible to try before obtaining the best combination for a radial

kernel-based SVM classifier. To avoid that, a grid search method will be performed

where the set of tuning parameter values are manually selected, as shown in Table 12.

Table 12: Tuning parameters for radial SVM classifier

Tuning Parameters for Radial SVM	Range of Values	Total Number of Values
C	$10^0, 10^1, 10^2, …, 10^9$	10
gamma	$10^{-7}, 10^{-6}, 10^{-5}, …, 10^0$	8
Total tuning combinations		80

To find the best-tuned parameter among all the 80 combinations, it is necessary to follow

a validation approach for each combination of the tuning parameter. This study performs

5-fold cross-validation (CV) for each combination of tuning parameters. In a 5-fold CV

approach, the training set is divided into 5 splits, and 4 of those splits are used for model

sub-training, and the remaining fold is used for validating the sub-trained model (refer to

Figure 9). The fitted classifier can be validated based on a performance measure

discussed later in Section 3.7.

3.6 Performance Metrics

After fitting the fine-tuned classifier using the training set, its prediction performance

needs to be tested using a testing set. A confusion matrix is a visual representation of a

classifier's performance using a testing set. It consists of rows and columns, where each

row represents numbers in a predicted class, and each column represents numbers in an actual class. Several performance metrics can be utilized to observe the performance of a machine learning classifier. In this study, the performance metrics can be classified into two categories: frequency of occurrence rates (Yuill and Braun, 2012) and class performance rates (Chen et al., 2022).

In VCRC, there could be three types of frequency of occurrence rates (Yuill and Braun, 2012): False Alarm Rate (FAR), Misdiagnosis Rate (MR), and Missed Detection Rate (MDR). Lower values for these three rates indicate the better predicting performance of the machine learning classifier. For instance, for a typical three-class confusion matrix including 'NoF' as a class, Figure 11 shows the false alarms, misdiagnoses, missed-detections, and correct predictions (true positives).

Figure 11: Confusion matrix showing three frequency of occurrence rates

The descriptions of these frequency of occurrence rates are given below.

FAR: The percentage of actual unfaulty cases incorrectly predicted as faulty. A false alarm happens when an AFDD algorithm incorrectly predicts the presence of a fault when there is no fault present (red elements in Figure 11).

MR: The percentage of total faulty cases incorrectly predicted as a fault class different from the actual fault class. A misdiagnosis error happens when the AFDD algorithm correctly predicts the presence of a fault, but a wrong fault is predicted (orange elements in Figure 11).

MDR: The percentage of total faulty cases incorrectly predicted as unfaulty. A missed detection occurs when an AFDD algorithm fails to detect the presence of a fault (blue elements in Figure 11).

In practical AFDD applications, false alarms are the most serious error because costly maintenance can be performed when there is no significant fault present. Misdiagnosis is the next serious error because a wrong diagnosis of a fault can again lead to an improper maintenance, keeping the actual fault unattended. Thus, as mentioned in Yuill and Braun (2013), the order of importance (from most to least important) for the above three metrics are FAR, MR, and MDR.

There could be three class performance rates for analyzing the individual class performance from a machine learning classifier: precision, recall, and F1-score. Before knowing the definitions of these metrics, it is necessary to know several other metrics related to a particular class. For each class, these are defined as follows,

True Positive (TP): The number of cases that are correctly predicted as that class.

False Negative (FN): The number of cases that are incorrectly predicted as other fault classes.

False Positive (FP): The number of cases that are incorrectly predicted as that class.

True Negative (TN): The number of cases predicted to be any other classes.

Now, for any class, the three class performance rates can be mathematically defined as,

$$precision = \frac{TP}{(TP+FP)} * 100\% \quad ... \quad (3.4)$$

$$recall = \frac{TP}{(TP+FN)} * 100\% \quad ... \quad (3.5)$$

$$F1\ score = 2 * \frac{precision*recall}{(precision+recall)} * 100\% \quad ... \quad (3.6)$$

For a particular class, higher values for the three class performance rates mean better predicting performance of the machine learning classifier. In this study, (100-recall) for the unfaulty class is the same as the FAR defined above.

3.7 Performance Measures

Cross-validation for getting the best-tuned parameters can typically be performed based on a performance measure called overall accuracy (followed by Ebrahimifakhar et al., 2020). The overall accuracy is the ratio of the total number of correct predictions from the AFDD classifier to the total number of actual cases in the testing set.

So, the best combination of tuning parameters can be obtained that yields the highest accuracy among all 80 tuning combinations discussed in Section 3.5. Therefore, the subsequent SVM classifier corresponding to the best-tuning parameters is the best-tuned SVM classifier (Final Classifier shown in Figure 9).

3.8   Fault Impact Limits

As mentioned in Yuill et al. (2014), for an RTU, each case (row) of simulated data corresponds to a specific operating condition (set of T_{RA}, WB_{RA}, T_{amb}, and specific fault intensity). Applying fault with a particular intensity impacts the RTU performance typically indicated by the values of capacity (Q) and Coefficient of Performance (COP). So, two Fault Impact Ratios (FIR) exist for capacity and COP, denoted by FIR_Q and FIR_{COP}, respectively. As per the definitions by Yuill and Mehrabi (2017),

$$FIR_Q = \frac{Q_{faulted}}{Q_{unfaulted}} \quad \ldots \quad (3.7)$$

$$FIR_{COP} = \frac{COP_{faulted}}{COP_{unfaulted}} \quad \ldots \quad (3.8)$$

Ideally, for an unfaulty operating condition, FIR = 1 for both of the above equations. Typically, the lower the FIR values, the more degraded the performance of the RTU for a specific faulty condition. In this study, the subtle faulty conditions (dictated by reasonable higher FIR_{COP} limits) will be treated as unfaulty conditions in the simulated data. To do that, reasonable lower and upper limits of FIR_{COP} need to be considered as unfaulty conditions. This will also mitigate the minority unfaulty (NoF) cases in the

simulated data, as evident from Table 8. The main idea is to consider some insignificant faults to be considered as unfaulty while training the machine-learning classifier. A wider limit (e.g., 97 to 103%) will increase the unfaulty cases and a tighter limit (e.g., 99 to 101%) will potentially neglect the insignificant faults to be considered as unfaulty. Thus, a tradeoff between the above two conditions must be met. So, the reasonable lower and upper limit of FIR_{COP} have been selected as 98% and 102%, respectively. So, for any case (i.e., row) of the simulated data, if the value of FIR_{COP} falls between 98% to 102%, inclusive, the associated fault class was converted from faulty to unfaulty case. However, for the lab data, although FIR_{COP} is known for all cases, the less significant faulty cases will not be considered unfaulty since the impact of fault is unknown.

To know how the ML classifiers perform at different fault impacts, five fault impact bins were considered (Yuill, 2014) based on different ranges of FIR_{COP} values: '>105%', '95-105%', '85-95%', '75-85%', and '<75%'.

3.9   Feature Selection

Feature selection can be defined as finding the reduced set of features for which there is a reasonable trade-off between the number of features and the prediction performance of the ML classifier. Reducing the number of features is also necessary because the cost associated with deploying reliable sensors for every feature might not be practical. In this study, simulated data for S3 has been considered to find the reduced set of features because, as evident from Table 8, it has more unique values for T_{RA}. Thus, simulated S3 data has more operating conditions than simulated S1 and S2.

As mentioned in the Introduction section, two feature selection approaches were
performed in this study. The filter approach followed in this study is based on the
correlation among the 15 features. In contrast, the wrapper approach followed in this
study is of backward elimination type, using radial SVM as its machine learning
algorithm with a 10-fold CV. The information regarding these two approaches using R
programming language (R Core Team, 2022) is shown in Table 13. The R codes and
outputs related to these two approaches are shown in Appendix F.

Table 13: Filter and Wrapper approaches using R software

Feature selection	R package	Functions used	Major arguments
Filter		findCorrelation()	cutoff = 0.90
Wrapper (backward elimination)	caret (Kuhn, 2008)	rfeControl()	function = caretFuncs, method = "cv", number = 10
		rfe()	method = "svmRadial"

The filter approach followed in this study provided a set of 8 relatively uncorrelated
features based on a correlation cut-off value of 90% (0.90 in Table 13). A lower cut-off
value eliminates more features, and a higher cut-off value retains more features; hence a
reasonable value of 90% was selected to have a better trade-off. To choose the
uncorrelated features, absolute values of pairwise correlations were considered. The first
pair is selected as $T_{sat,c}$ and P_{LL} because this pair has the highest correlation of 99.10%.
To eliminate one, the mean absolute correlation of each of these two with other variables
is considered. Since $T_{sat,c}$ has a higher mean absolute correlation with the rest of the
variables than that of the P_{LL}, $T_{sat,c}$ was eliminated. Next, the pair of P_{LL} and P_{dischg} is
selected since it has the next highest correlation of 98.5%. By using the same analysis,

P_{LL} was eliminated. Following this order, the next 5 eliminated features are P_{dischg}, $T_{air,ce}$, P_{suc}, $T_{sat,e}$, and WB_{SA}. In this way, 7 correlated (or, redundant) features were eliminated, and 8 uncorrelated features were obtained.

Backward elimination has been selected as a wrapper feature selection because previous studies (Chen et al. (2022, 2023)) have shown potential with this approach in getting statistically justified set of features. Backward elimination starts with full features and recursively eliminates less important features. It does that based on a machine-learning algorithm informing which features are adding less knowledge to the trained algorithm, marked by accuracy. To match with the same number of features, this study takes the top 8 features from backward elimination and will compare it with that of the 8 uncorrelated features.

Among many factors, the thermodynamic variables of a rooftop unit largely depend on the type of refrigerant used in the system. Thus, to fit a generalizable classifier for packaged rooftop units, refrigerant type was taken as an additional feature. Thus, in total, there are 9 features to fit classifiers using a reduced set of features.

Three machine learning classifiers were studied to compare the fault-predicting performance using the simulated testing set. The three classifiers are listed in Table 14.

Table 14: Three classifiers for fault prediction

Classifier	Description
C1	SVM classifier using 9 features from wrapper approach (backward elimination)
C2	SVM classifier using 9 relatively uncorrelated features from filter approach
C3	SVM classifier using full 16 features

Classifiers C1 and C2 were considered to show which reduced featured classifier (C1 or C2) showed better predicting performance. Classifier C3 was considered to compare the performance of the reduced featured classifiers (C1 or C2) with the full-featured classifier (C3).

Lab and field validations were performed for the C2 classifier only, as we noticed that this classifier performed better than C1 when tested with the simulated data.

3.10   Data Analysis Steps

In this study, following Figure 9, the whole dataset consists of the simulated data from S1, S2, and S3, discussed in Section 3.2. This study aims to develop a generalizable machine learning algorithm with a reduced number of features to use in the field to detect and diagnose the common fault classes discussed in Section 1.2.3. This generalized ML algorithm is like one of those FDD protocols evaluated by Yuill et al. (2014). Furthermore, SVM with radial kernel will be used to fit different machine learning classifiers since this classifier showed the best performance among all classifiers in a previous study (Ebrahimifakhar et al., 2020).

By following the machine learning workflow shown in Figure 9, the data analysis steps were as follows,

1. The whole dataset consists of simulated data from S1, S2, and S3. Therefore, the data has numerous rows and columns. Rows that did not provide meaningful values were discarded.

2. To increase the 'NoF' cases, fault cases that did not significantly impact the performance (evident by FIR_{COP}) were converted to 'NoF'. A reasonable FIR_{COP} range has been selected from 98% to 102%, to convert any fault to no fault. This conversion increases the number of 'NoF' cases significantly.

3. 16 features (including refrigerant type) and one response variable (fault class) were taken, and the whole dataset was shuffled.

4. The shuffled simulated dataset was randomly split into training and testing sets following the percentages of 80% and 20%, respectively.

5. The actual $Power_{comp}$ and T_{dischg} were appropriately normalized based on the methods discussed in Section 3.3. Next, the features were appropriately standardized discussed in Section 3.4.

6. The SVM classification algorithm (C3) was fitted with the training set with appropriate fine-tuning and cross-validation.

7. The performance of the full featured classifier obtained in step 6 was evaluated using the simulated testing set.

8. Next, feature selection, using both filter and wrapper approaches, has been performed to find the most cost-effective reduced set of features.

9. With the reduced number of 9 features, steps 6 and 7 were repeated to get the performance of the two SVM classifiers (C1 and C2) with the reduced set of 9 features. The results with reduced features (from C1 and C2) were compared with the results with 16 featured (C3, obtained in step 7) ones. The final generalized classifier, C2, was obtained since it was noticed that this classifier provides the best performance among the three.

10. An existing lab dataset (consisting of same S1, S2, and S3) was appropriately

    normalized and standardized (refer to step 5) and used as a new testing set. The

    performance of the generalized ML classifier (C2) obtained in step 9 was

    evaluated using this lab testing set. In addition, a comparison has been performed

    with the results obtained from the simulated testing set.

11. Field measurements (discussed elaborately in the next section) were performed

    for the reduced set of features using filter approach from an industrial facility. The

    generalized ML classifier, C2, obtained in step 9 was validated using the field-

    measured data.

Chapter 4. Field Testing

In this section, everything related to the field testing of the proposed machine-learning classifier (C2) will be presented.

4.1   Available Program at the UNL

The field testing was performed under the guidance of two organizations at the University of Nebraska-Lincoln (UNL): Nebraska Industrial Assessment Center (NIAC) and Partners in Pollution Prevention (P3). The P3 and NIAC programs at the UNL go hand in hand since both have common goals to fulfill.

The NIAC program (*About NIAC*, n.d.) is funded from the U.S. Department of Energy (U.S. DOE). This organization performs one-day energy assessments to small and medium manufacturing facilities, aiming to improve energy efficiency in the Midwest area. The NIAC is managed by Industrial Assessment Center (IAC) program (*Industrial Assessment Centers*, n.d.) under the U.S. DOE.

The P3 program (*P3: Partners in Pollution Prevention*, n.d.) provides pollution prevention to Nebraska businesses by minimizing waste and resource conservation. Some selected publications from P3 program are Dvorak et al. (2010), Kekilova et al. (2014), and Kuppig et al. (2016).

4.2 Manufacturing Facility with Rooftop Units

A suitable manufacturing facility located in Omaha, Nebraska was selected for the field validation of the machine-learning classifier. The suitable day for the field validation was July 14, 2023. This facility manufactures building drywall components and has a square footage of 140,000 ft^2 and has 4 rooftop units for air-conditioning the offices and shop areas. The 4 RTUs at this facility are marked by RTU 1, RTU 2, RTU 3, and RTU 4. Each of the 4 RTUs is a York brand with a model number of Y12AN44A9AA0ABF. The nameplate of this model is mentioned in Appendix A. RTU 4 is the same as system S4 described in Section 3.2.

As can be related from Section 1.2.2, the RTUs incorporate a VCRC with compressor, condenser, thermostatic expansion valve, and evaporator. Since each RTU has 2 independent cooling circuits (two stages), each RTU has 2 sets of heat exchanger coils (condenser and evaporator), 2 compressors, and 2 TXVs. For each RTU, stage 1 meets the primary cooling demand, and stage 2 turns on when stage 1 can't maintain the cooling demand alone. Table 15 shows the main equipment inside each RTU package.

Table 15: List of main equipment inside each field RTU

Equipment	Total Set	Quantity per Stage	
		Stage 1	Stage 2
Compressor	2	1	1
Condenser coil	2	1	1
Condenser fan	4	3	1
TXV	2	1	1
Evaporator coil	2	1	1
Supply air fan/blower	1	NA	NA

4.3   Detecting and Diagnosing Soft Faults

Each of the 4 rooftop units is identical and of the same age (installed in 2004). Since

those RTUs are relatively old, it is expected to be diagnosed with at least one of the 7 soft

faults discussed in Section 1.2.3. During the field testing, circuit 1 of RTU4 was found to

be running steadily, thus this circuit was considered for checking any potential soft faults.

Three types of measuring equipment were used for detecting and diagnosing purposes,

- Refrigerant pressure sensors: Fieldpiece JL3PR Job Link

- Refrigerant temperature sensors: Fieldpiece JL3PC Job Link

- Air temperature sensors: Fieldpiece JL3RH Job Link

The above sensors are shown in Figure 12. In this figure, from left to right, the first two

pieces are the air temperature sensors, the next two are the refrigerant pressure sensors,

and the last two are the refrigerant temperature sensors.

Figure 12: One time measuring equipment used (*JL3KH6 Job Link*, n.d.)

Faults were detected and diagnosed based on expert intuition and simple rules as discussed below.

For CA and EA: This can be checked in two ways.

- Visual inspection of condenser coil, evaporator coil, filters: Condenser and evaporator coils were visually checked for any significant fouling. The facility has been cleaning their coils and replacing their filters regularly. So, no condenser and evaporator fouling were present.

- Monitoring supply air fans, condenser fans, outside and return air dampers: No abnormalities were detected.

For LL: This can be checked by monitoring the liquid-line temperatures across the filter-drier. To check that, two refrigerant temperature sensors (Fieldpiece JL3PC) were installed before and after the filter drier. For RTU4 circuit 1, after the compressor of this circuit started running, the collected temperatures on every 15 seconds are listed in Table 16.

Table 16: Temperature before and after the filter in the liquid-line

Time	Temperature Before Filter (°F)	Temperature After Filter (°F)
0:00 (start)	75	76.4
0:15	88	84.2
0:30	106.1	107.1
0:45	106.9	109.5
1:00	96.7	100.3
1:15	93.2	95.8
1:30	93.9	95
1:45	95.7	96.8
2:00	97	98.2
2:15	97.7	98.6
2:30	93.6	96
Average (°F)	94.89	96.17

From Table 16, since the average temperature difference before and after the filter was found to be (96.17-94.89)°F = 1.28°F < 3°F, no LL fault is present as per the simple rules (Table 7) described in Chen and Braun (2001).

For NC: NC fault for system equipped with TXV is typically marked by higher subcooling values (Chen and Braun, 2001). No NC faults were detected as evident from the lower subcooling values discussed while checking for UC fault later.

Another procedure for diagnosing NC fault was mentioned in Hu et al. (2021). The non-condensable gases typically accumulate in the condenser when the RTU is off. Thus, when the RTU is off and reaches a steady state, the condenser pressure can be measured. The saturated temperature corresponding to this measured pressure can be obtained from the corresponding refrigerant (R22) properties. This saturated temperature can be compared with the measured condensing temperature. If the difference between these two temperatures is more than 1.7°C (3°F), the NC fault is diagnosed. In this study, this procedure was not followed since condensing temperature was unfortunately not measured, although the saturated condensing temperature (73.15°F) corresponding to the measured condensing pressure (142.9 psia) was known in RTU off condition.

For OC and UC: As mentioned in Chen and Braun (2001), for system equipped with TXVs, OC and UC faults are marked by higher and lower subcooling values, respectively. In addition to that, for systems equipped with FXO, OC and UC faults are also marked by lower and higher suction superheating values, respectively.

To measure the suction superheating and subcooling for RTU4 circuit 1, the methods listed in Appendix B can be followed. Appendix B shows that the suction superheating and subcooling values for the above system were calculated as 33.5°F and 2.2°F, respectively. Since these values significantly vary compared to the rated unfaulty system (discussed in Appendix B), RTU4 circuit 1 is diagnosed with severe UC fault.

For VL: VL is typically marked by excessive compressor noise. As the compressor noise was found normal, no VL is assumed to be in this system.

4.4   Installing Data Loggers

To detect and diagnose the above UC fault using the proposed machine-learning

classifier, steady-state data needs to be fed into the proposed classifier and observe the

predicting performance. The proposed machine-learning classifier in this study has 8

measurable variables. Thus, the 8 measurements need to be logged at sufficient interval

from the above faulty RTU. To do so, data loggers need to be installed. Following data

loggers were used in this study (refer to Appendix C for pictures),

a) Hobo UX120 4-channel (quantity: 2): For measuring T_{RA}, T_{SA}, T_{amb}, T_{LL}, T_{suc}, and

T_{dischg}.

b) PCE-PA 8000 power meter (quantity: 1): For measuring single phase voltage, three

phase currents, to ultimately get $Power_{comp}$.

c) Extech RHT10 humidity logger (quantity: 1): For measuring T_{RA} and RH_{RA}, to

ultimately get WB_{RA}.

In addition, 6 temperature probes were used for the 6 measurements in (a) that connect

into the two Hobo data loggers. The temperature probes for T_{RA}, T_{SA}, and T_{amb}, were

positioned in the corresponding air lines, and temperature probes for T_{LL}, T_{suc}, and T_{dischg},

were appropriately insulated and attached with the corresponding refrigerant lines.

4.5   Data Preprocessing

Data preprocessing involves converting the raw data into features for the machine-

learning classifier.

Data from the above sensors of Hobo, PCA-PA 8000, and Extech RHT10 had sampling intervals of 1 s, 10 s, and 30 s, respectively. Thus, the data from Hobo and PCA-PA 8000 needs to be filtered every 30 seconds to match the timing with Extech RHT10.

To get the $Power_{comp}$ from single phase voltage and three phase currents (mentioned in (b) in Section 4.4), first, for each phase, voltage and current are multiplied for getting the phase apparent power (in unit VA). Next, the three phase apparent powers were summed to get the total apparent power for the compressor. Since compressor current is measured in single phase (instead of three phase), the recorded total power factor showed a very low value (around 0.15). Thus, the total apparent power is multiplied with a constant power factor of 0.95 to get the value for $Power_{comp}$ in Watt.

To get WB_{RA} from the logged T_{RA} and RH_{RA}, the thermodynamic properties of R-22 refrigerant were utilized. For each set of T_{RA} and RH_{RA}, WB_{RA} can be determined using the software called Engineering Equation Solver, EES (Klein, 2018) (refer to Appendix D).

After getting the logged data in its corresponding featured version, the steady state cases need to be identified. In this study, the steady state cases were identified by considering steady T_{suc} and T_{dischg} values.

The data for the above 7 thermodynamic measurements were logged for around 5 days (from 06/14/2023 to 06/19/2023). However, due to the malfunctioning power meter, $Power_{comp}$ was only calculated for about 9 hours on 06/14/2023. Thus, the steady state conditions for the 8 variables need to be considered from those 9 hours range. It was found that, from those 9 hours window, the time from 14:38 to 18:15 need to be

considered because all the 8 variables only exist in this time range. The logged data for the 8 variables in that time window are graphically shown in Appendix E.

During the time between 14:38 to 18:15 on 06/14/2023, based on the steady-state operating conditions (marked by steady T_{suc} and T_{dischg}) of circuit 1 of the faulty RTU, 24 steady-state cases have been considered by taking a 5-minute average value for each of the 8 measurements. The time ranges and the averaged values for these 24 cases are shown in Table 22 located in Appendix E.

After getting the steady state cases for the 8 variables, the data were appropriately normalized and standardized as per the methodologies mentioned in Section 3.3 and Section 3.4, respectively. After adding the categorical feature of refrigerant type with R22 value, the 24 converted steady-state cases (with 9 features) are now ready to be fed into the proposed machine-learning classifier and get the predictions.

Chapter 5. Results

The classifier performance results will be discussed for different testing sets. Testing sets can be either simulated, experimental, or field measurement data. In the simulated testing set, as discussed in Section 3.8, FIR_{COP} values of 98% to 102% are considered to convert the faulty cases to unfaulty cases. However, FIR_{COP} will not be considered in lab and field data because it is impossible to know the fault impact on the performance of the VCRC in real life while applying the FDD method. First, feature selection will be performed by considering simulated S3 data since this dataset contains more operating conditions than S1 and S2. Then, the performance of C1, C2, and C3 with the simulated testing set will be provided. Next, the performance of C2 will be analyzed separately for simulated and lab testing sets. In addition to this, to validate the features in C2, the top 8 features from backward elimination will be considered.

5.1   Feature Selection Results

Using the Filter approach, the 8 relatively uncorrelated features, while analyzing randomly selected 70% of the simulated S3 data, are T_{RA}, WB_{RA}, T_{SA}, T_{amb}, T_{LL}, T_{suc}, T_{dischg}, and $Power_{comp}$ (Appendix F for R code). As we can observe, the obtained features are consistent with the knowledge of correlations among the features of a VCRC. These uncorrelated features match what Chen et al. (2022, 2023) selected as their 10 manually selected features, except they eliminated WB_{RA} and considered $T_{air,ce}$, $T_{sat,e}$, and $T_{sat,c}$. Also, the sensors associated with the above 8 features can be easily deployable to an RTU in the field for performing the field validation of the machine learning classifier.

As justified in Section 3.8, To validate the 8 uncorrelated features, this study will also consider a separate set of 8 features obtained using a Wrapper approach called backward elimination (with radial kernel-based SVM). Using the same data (70% of simulated S3 data), the top 8 features obtained using this approach are T_{SA}, P_{LL}, T_{LL}, P_{suc}, T_{suc}, T_{dischg}, $T_{sat, e}$, and $Power_{comp}$ (Appendix F for R code). Compared with the 10 features from SBS in Chen et al. (2022, 2023), T_{LL}, P_{suc}, and $Power_{comp}$ were absent, and they found WB_{RA}, WB_{SA}, P_{dischg}, $T_{air,ce}$ and T_{amb} as additional features.

As mentioned previously, to make a more generalizable classifier, refrigerant type was taken as an additional feature in all associated classifiers. It should also be noted that, before fitting any classifier, the values in the columns $Power_{comp}$ and T_{dischg} were normalized with the unfaulty cases based on the method described in Section 3.4.

5.2   Performance with Simulated Testing Set

This section will compare the predicted performance of C1, C2, and C3 with the same unseen simulated testing set.

First, before fitting the three classifiers, as described in the Methodology section, the subtle fault cases ($FIR_{COP} = [98\%, 102\%]$) in the simulated data were converted to 'NoF'. Among many other factors, the performance of the classifiers will also depend on the distribution of the fault classes. The distribution of fault classes in the training data is shown in Figure 13.
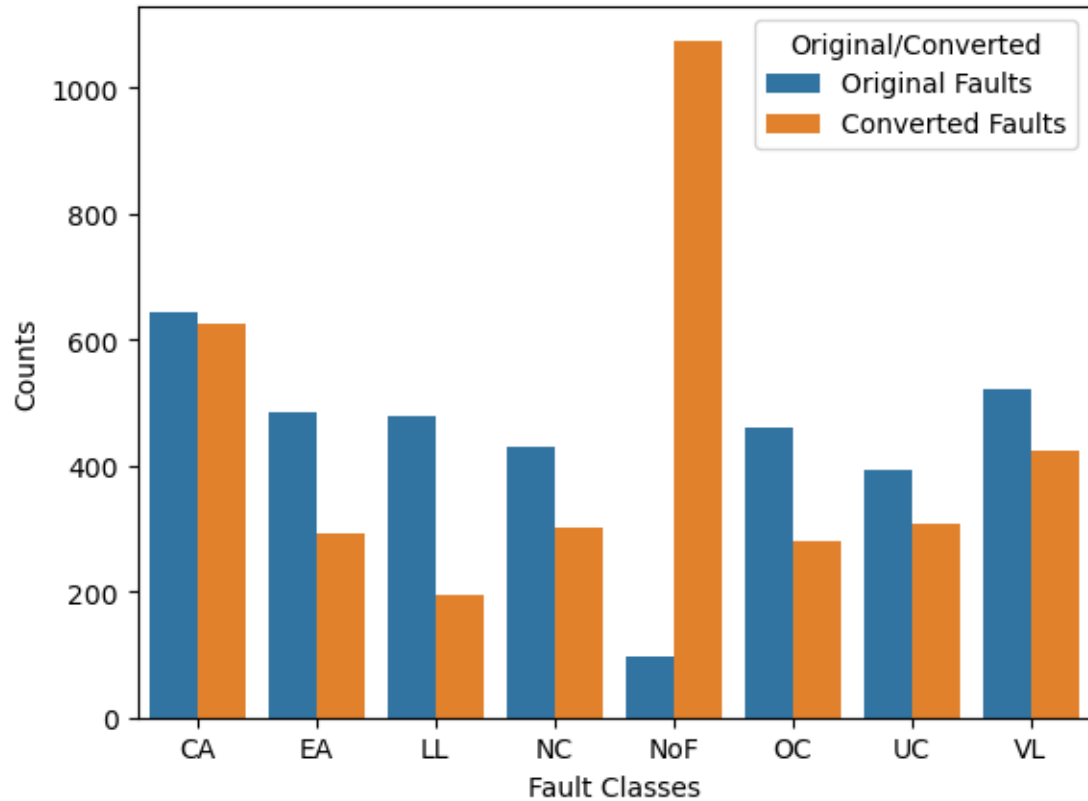
Figure 13: Distribution of fault classes

Figure 13 shows that most of the conversions occurred from LL, and the least occurred from CA classes, meaning, in the original simulated data, the LL class has the most subtle faults, and the CA class has the least subtle faults, among all the fault classes. Figure 13 also shows that the converted training set is imbalanced with higher NoF cases. This is intentionally done because false alarms are costlier in practical AFDD applications than missed detections and misdiagnoses. Higher NoF cases will yield lower FAR at the expense of higher MDR, which we will notice in the following section.

Next, using the 5-fold CV discussed previously, the best-tuned parameters (C and

gamma) for the three classifiers are shown in Table 17. Appendices J, K, and L can be

observed for related R codes and outputs.

Table 17: Best-tuned parameters for three classifiers

Classifier	C	gamma
C1	10^5	0.1
C2	10^3	0.1
C3	10^6	0.001

Using the same simulated testing set, the prediction performance of the three classifiers

can be expressed as three confusion matrices (CM) shown in Figure 14 (a, b, c). In these

CMs, the predicted classes are shown along the rows and the actual classes are shown

along the columns. Any CM can be quickly read by carefully observing the diagonal and

off-diagonal elements. Higher numbers in the diagonal elements (TP values) and lower

numbers in the off-diagonal elements indicate better performance from any CM. By

looking at the lower TP values (diagonal elements) of the CM from C1 (Figure 14 (a)), it

is evident that this classifier is the worst among all the three classifiers.

(a) CM from C1   (b) CM from C2

(c) CM from C3
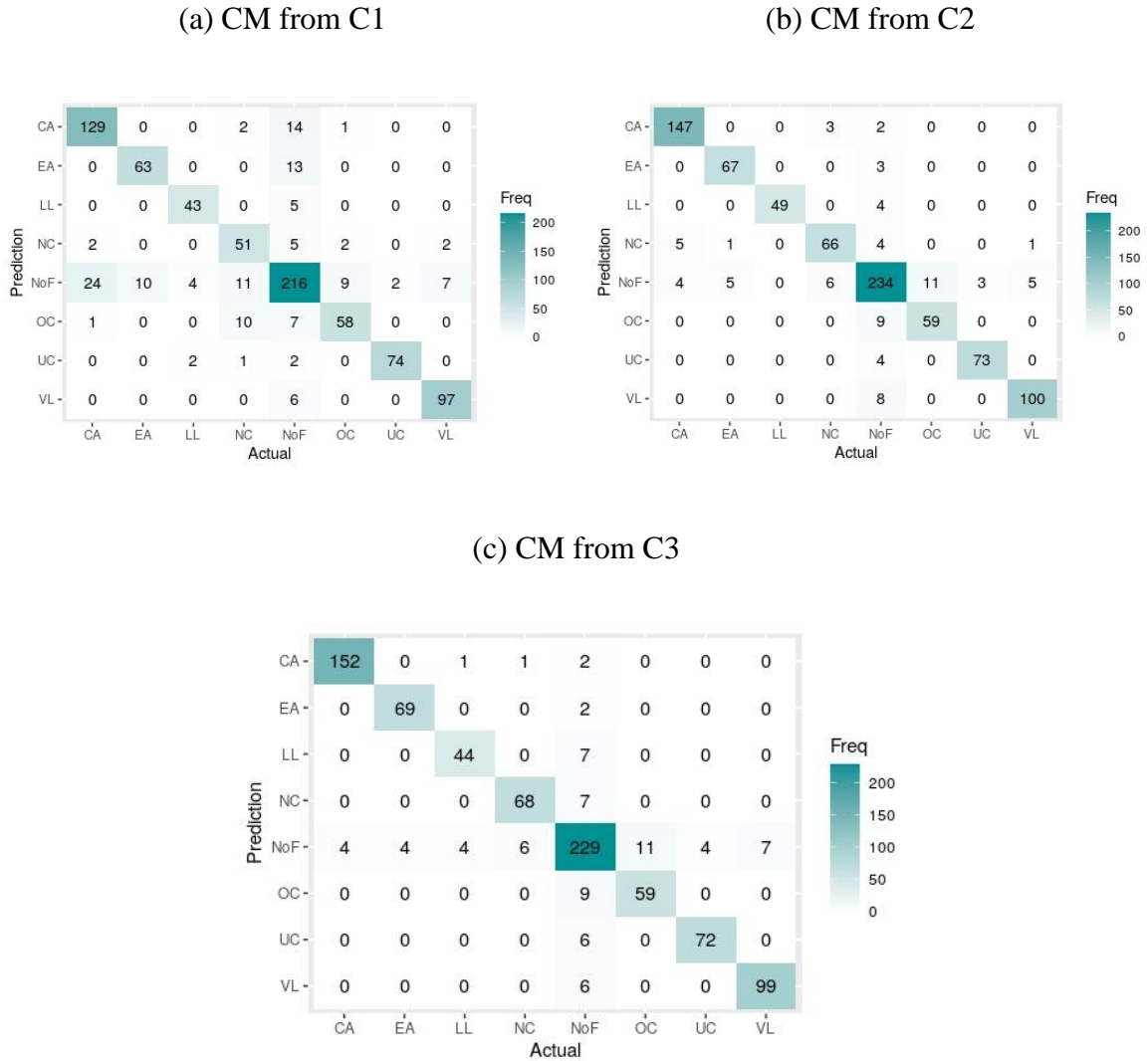
Figure 14: Three confusion matrices using same simulated testing set

In all the CMs, we could observe a higher number of missed detections (higher values in the NoF rows in the three CMs in Figure 14 (a,b,c)). This is expected since the classifiers were trained with a higher number of 'NoF' cases.

Using the simulated testing set, the performance metrics in terms of FAR, MR, MDR, and accuracy for the three classifiers are shown in Figure 15. This figure shows that the FAR from classifier C2 has the lowest (best) value of 12.7% among all three classifiers.

The performance in terms of MDR is also best (5.6%) with C2. Clearly, the performance

of C1 is the worst (higher frequency of occurrence rates and lower accuracy) among the

three classifiers. Overall, the accuracy of C2 is about 91%, highest among all the three
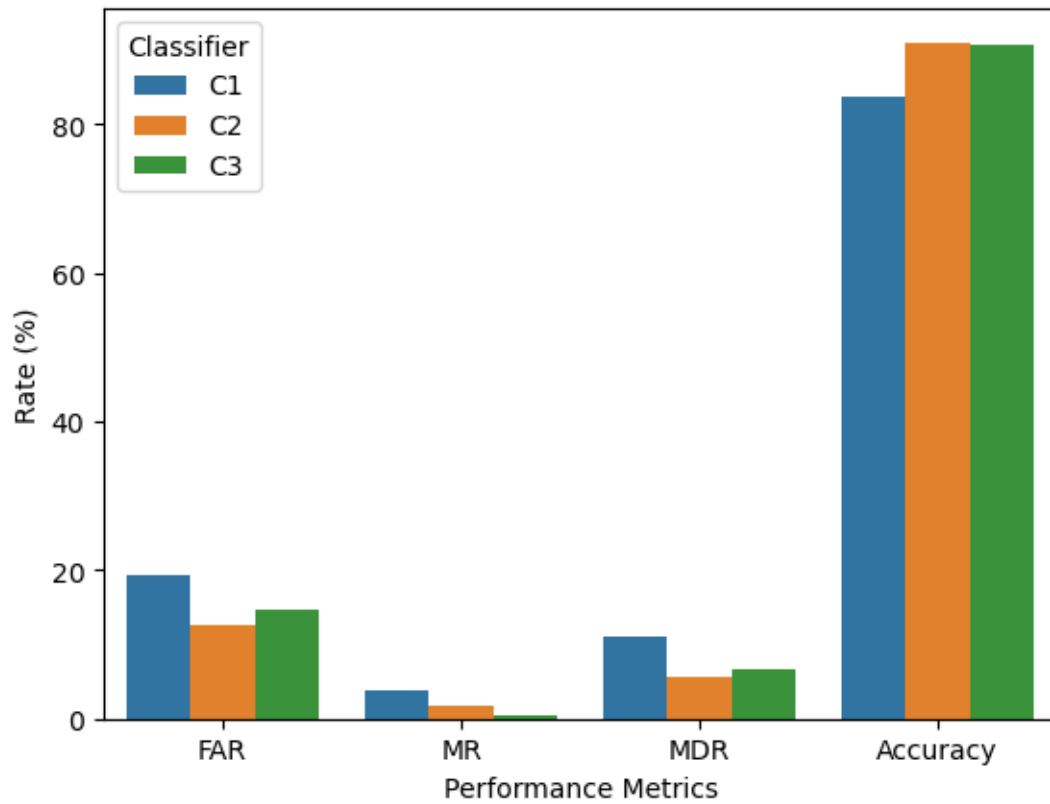
classifiers.

Figure 15: Performance metrics of the three classifiers with simulated testing set

Thus, the 9 relatively uncorrelated features are justified as the best-reduced set of features

for developing the ML algorithm for AFDD applications in packaged rooftop units. The

corresponding classifier (C2) for the confusion matrix shown in Figure 14 (b) is thus the

final ML classifier that needs to be validated with lab and field-measured data.

Next, the breakdown of each of the above performance metrics with respect to FIR_{COP}, except FAR, will be shown to see how the classifiers perform at different bins of fault impact. Five fault impact bins were chosen (Yuill et al., 2014) based on different ranges of FIR_{COP} values: '>105%', '95-105%', '85-95%', '75-85%', and '<75%'. The total number of cases under the five FIR_{COP} ranges in the testing set are presented in Figure 16. This figure shows that most of the cases fall under the FIR_{COP} of 95-105%. As expected, all the 'NoF' cases fall under this range. It also shows that the number of severe fault cases ($FIR_{COP}<85\%$) are really low as compared to the subtle fault cases.
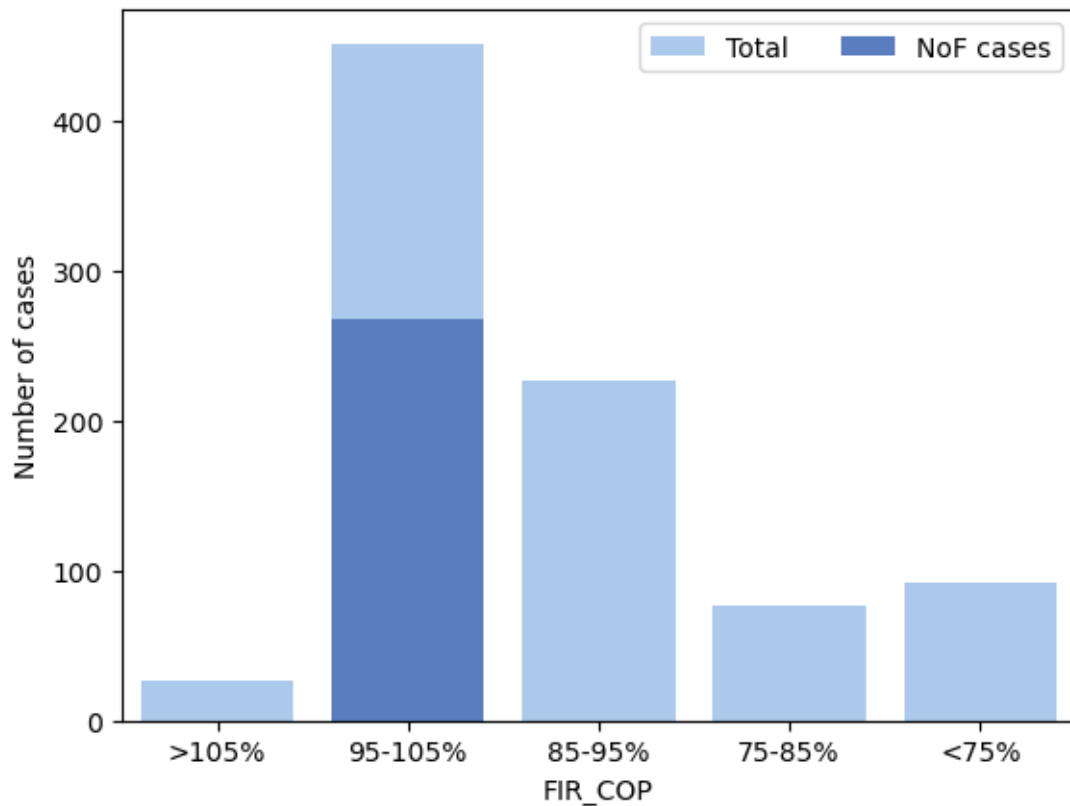
Figure 16: Breakdown of cases with respect to FIR_{COP} in simulated testing set

The breakdown of MR, MDR, and accuracy with respect to different FIR_{COP} bins is shown in Figure 17. The breakdown of FAR with respect to FIR_{COP} will not be shown

because the unfaulty cases only occur for the FIR_{COP} range of '95-105%', thus the false

alarms only make sense for that range only. Figure 17 (a) shows the MR Vs. FIR_{COP} for

the three classifiers. This shows that the MR's for all the three classifiers are low (<5%)

in all FIR_{COP} ranges. Classifier C1 performs worst in all the FIR_{COP} ranges (Figure 17

(a,b,c)). Compared to MR, MDR shows higher rates for $FIR_{COP} > 95\%$. This phenomenon

is expected as the classifiers are trained with higher number of 'NoF' cases, so the

classifiers are more prone to predicting 'NoF' for the faults that are subtle in nature (i.e.,

typically $FIR_{COP} > 95\%$). In all the performance metrics, the classifier performance

improves as we move from subtle fault cases to severe fault cases, meaning that the

classifiers are more accurate in predicting severe faults than subtle faults.

(a) MR

(b) MDR

(c) Accuracy
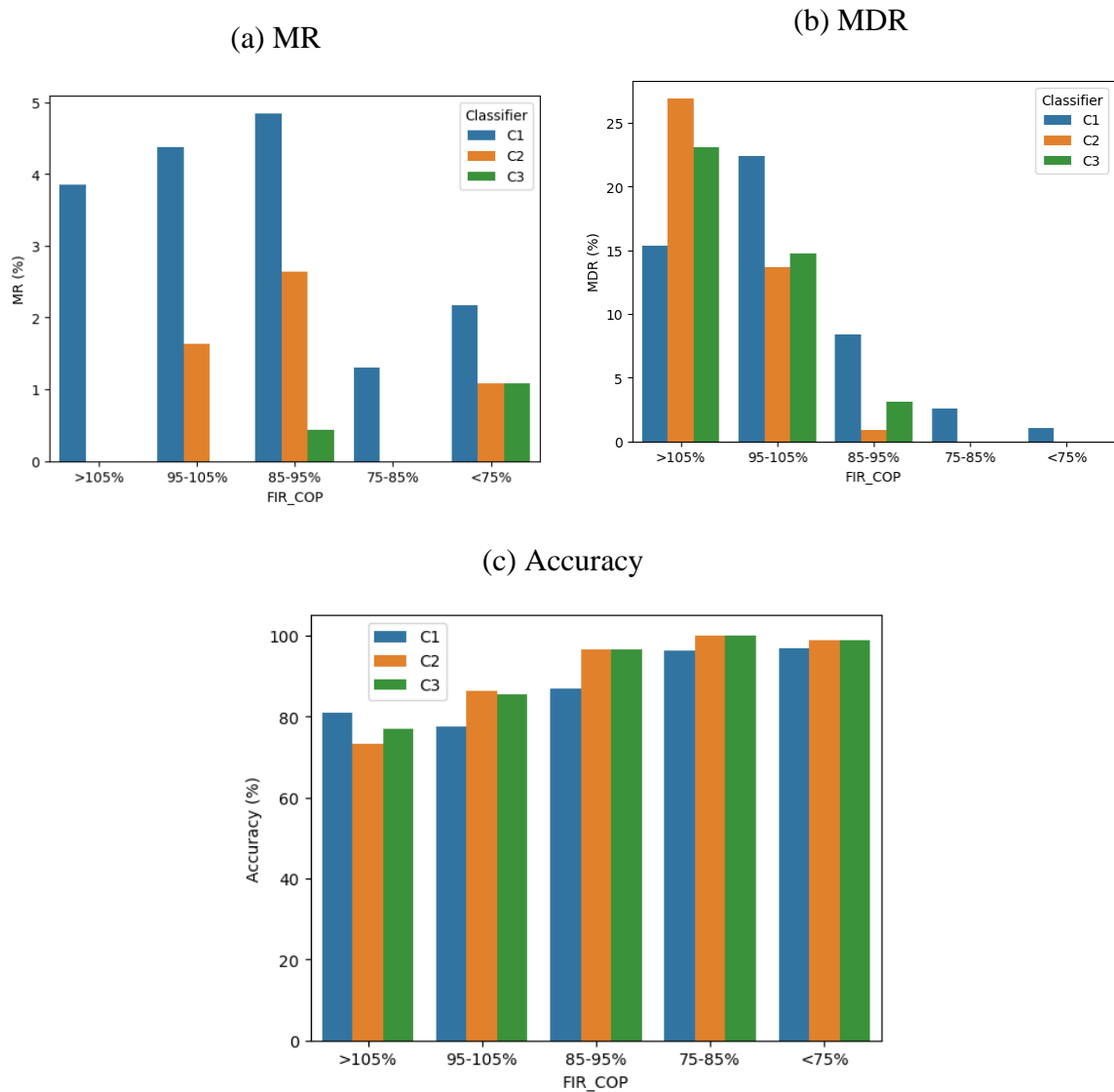
Figure 17: Breakdown of performance metrics in three classifiers

From the above results, it appears that, for systems like S1, S2, and S3, the classifier C2 can be treated as the final machine-learning classifier for AFDD in packaged rooftop units. This classifier (C2) needs to be validated using lab and field measurement data.

5.3   Performance with Lab Testing Set

In this section, using an existing lab testing dataset (consisting of the same three systems used for training the classifiers), the predicting performance of the classifier C2 was evaluated as this classifier, when tested with the unseen simulated data, shows significantly better performance than C1, and almost equal performance as C3. Appendix M can be seen for related R codes and outputs.

In the practical application of FDD, the impact of fault on the COP (FIR_{COP}) is unknown. So, unlike the simulated testing set, $FIR_{COP} = [98\%, 102\%]$ will not be considered in the lab (and field) testing set to convert any fault class from faulty to unfaulty. Thus, it is expected that the classifier prediction for the lab testing set will have a higher MDR than that of the simulated testing set due to possible subtle fault cases (FIR_{COP} closer to 100%) in the lab testing set to be predicted as 'NoF'.

First, the confusion matrix while validating the final classifier (C2) with the lab testing set is shown in Figure 18.
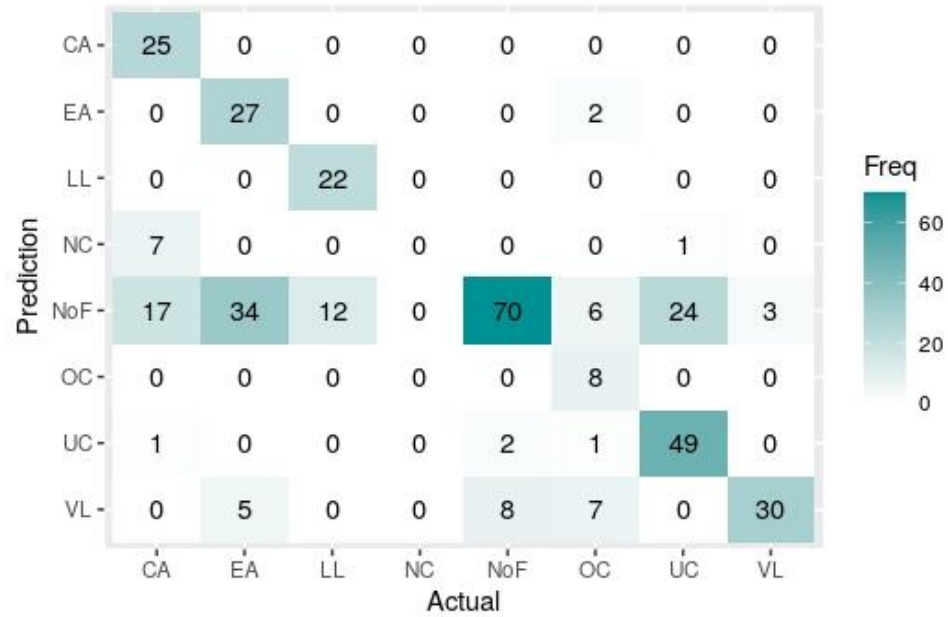
Figure 18: Confusion matrix for classifier C2 with lab testing set

Figure 18 shows, as expected since FIR$_{COP}$ was not considered, there exist many missed

detections (along the NoF row) when the final classifier (C2) is tested with the lab

dataset. Most of the missed detections occur for EA cases (#34), followed by UC (#24),

CA (#17), and so on. Figure 18 also shows that most of the false alarms (#8, along the

NoF column) occur with VL. Also, most of the misdiagnoses occur for CA and OC cases

(#7 CA cases misdiagnosed as NC, #7 OC cases misdiagnosed as VL). The CA and NC

faults typically show similar fault signatures by increased condensing pressure and

compressor power values. Thus, it's possible that the 7 CA cases were misdiagnosed as

NC due to these similar fault signatures. Similarly, OC and VL faults typically show

similar fault signature by decreased superheating values, thus it's possible that the 7 OC

cases were misdiagnosed as VL due to this similar fault behavior.

The performance metrics in terms of FAR, MR, MDR, and accuracy, for the simulated

and lab testing sets, are shown in Figure 19. As observed, when going from simulated to

lab testing set with classifier C2, FAR almost stays the same (12.7% Vs. 12.5%).

However, the MR increases from 1.7% to 8.5%, and MDR increases from 5.6% to

34.2%. As a result, the accuracy dropped from about 91% to 64%.
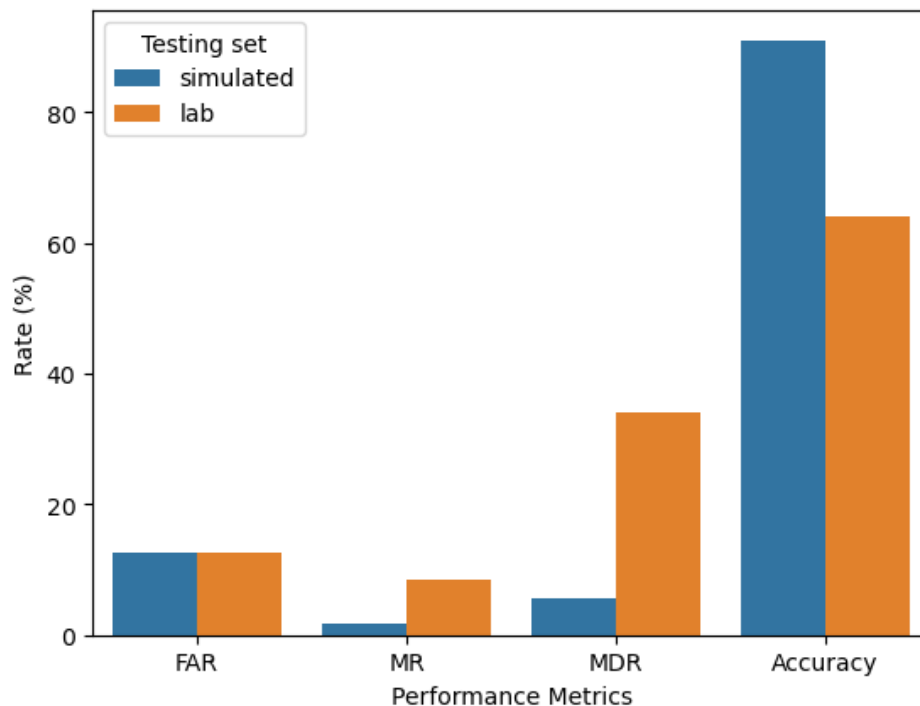
Figure 19: Performance metrics using simulated and lab testing sets

The results from classifier C2 with lab data can be compared with the FDD protocols

evaluated in Yuill et al. (2014b). This comparison will not necessarily be fair because the

FDD protocols use different features than C2; they are expected to lose significant

accuracy if their generalizability is tested with any testing set. Still, this comparison is

shown here since those FDD protocols were evaluated using the same lab testing set studied in this dissertation.

From Yuill et al. (2014b), if the FIR_{COP} threshold of 98% is considered, the best FDD protocol in terms of FAR was FDD E (FAR ~ 15%). The range of FAR across the 6 FDD protocols were from 15% to 95%. Thus, comparing with those 6 FDD protocols, classifier C2 in this study shows slightly better performance in terms of FAR (with FAR = 12.5%). When comparing C2 with the 6 FDD protocols in terms of MR, C2 is again showing significantly better performance (8.5% with C2 Vs. ~25% with FDD F). Although FDD D was showing 0% MR, it was a detection only protocol, thus cannot be directly compared with classifier C2 which is for both fault detection and diagnosis. Lastly, when comparing C2's performance in terms of MDR, the best comparable protocol was FDD B (34.2% with C2 Vs. ~10% with FDD B). However, although FDD B had lower MDR (~10%), it had a very high FAR of about 70% (at FIR_{COP} ~ 98%). Similarly, although FDD E had a comparable FAR of about 15%, it had a poor MDR value of about 65%. This indicates that the FDD protocols studied in Yuill et al. (2014b) did a poor job in finding a better trade-off between FAR and MDR. Although classifier C2 in this study has higher MDR, it has shown a significantly better trade-off between FAR and MDR, compared to the FDD protocols evaluated in Yuill et al. (2014b).

Figures 18 and 19 show that higher MDR while tested with the lab data can be largely attributed to the 34 missed detections with EA class. Out of these 34 MD's, 23 of them come from lab S3. As we can see, most of the MD's came from lab S3. From Table 8, since lab S3 has 26 cases in lab data, that means, most of the lab S3 cases were miss-detected. The reason most of the lab S3 cases were miss-detected can be explained in

terms of FIR_{COP}. In our simulated training data, mean FIR_{COP} (without percentage form) for the EA cases in the training S1, S2, and S3 data are 0.94, 0.95, and 0.95, respectively. Whereas, mean FIR_{COP} (without percentage form) for the EA faults for lab S1, S2, and S3 data are 0.95, 1.03, and 0.91, respectively. As we can see, since mean FIR_{COP} in lab S3 data indicate slightly severe EA fault cases than the mean FIR_{COP} of the training S3 data, it is possible that, those 23 severe lab S3 faulty cases were all predicted as 'NoF' because the classifier might be struggling to detect severe EA fault cases as it was trained with subtle EA fault cases. As a result, the classifier might be following its inherent bias of predicting those severe EA fault cases as 'NoF'. If the classifier is trained with severe EA fault cases, then those severe lab S3 fault cases could be correctly predicted.

Figure 20 shows how C2 performed at different fault intensities of the lab data. For better comparison, the bar plots in Figure 20 also show the previous performance of C2 with the simulated testing set.

(a) MR

(b) MDR

(c) Accuracy
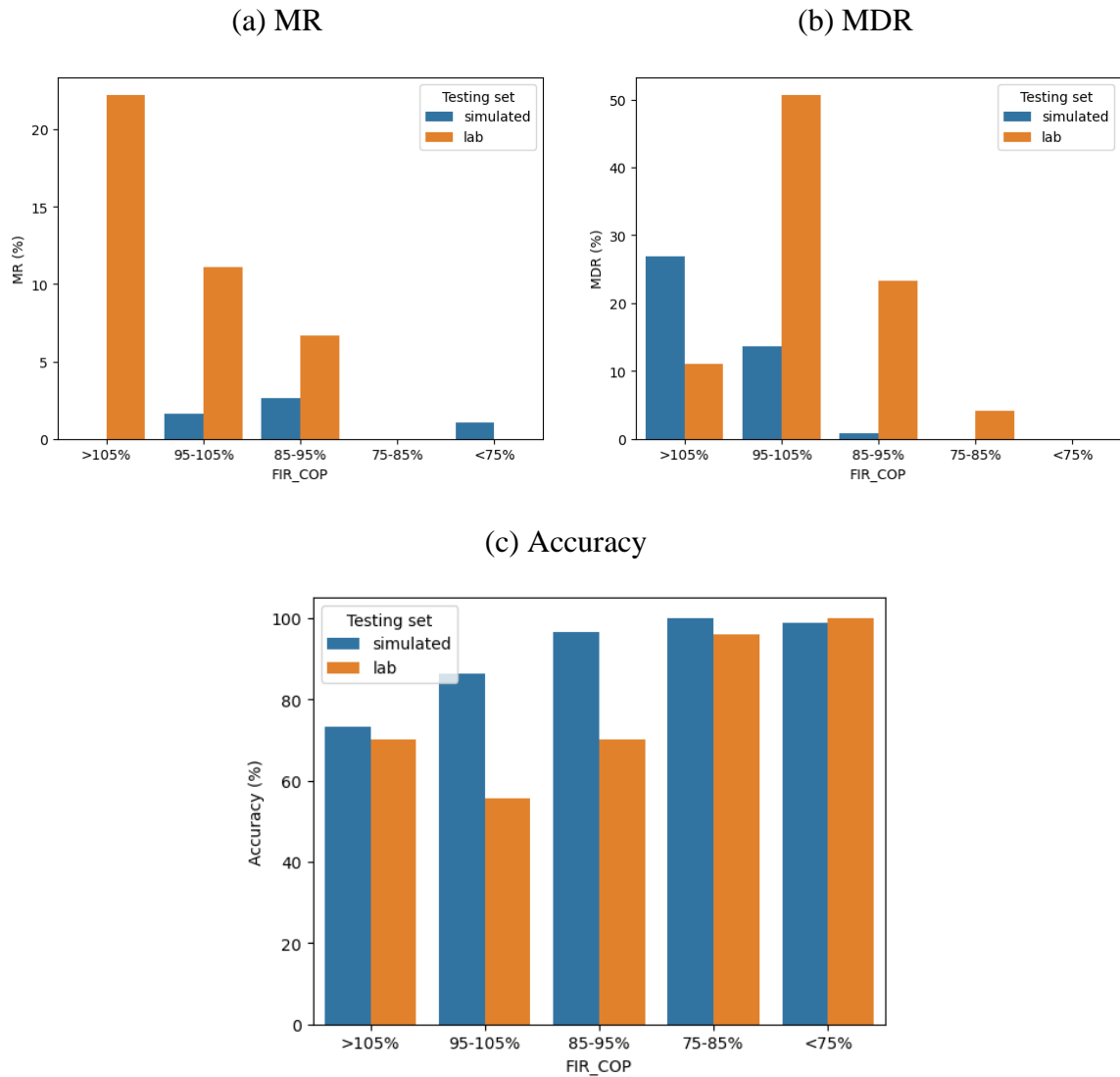
Figure 20: Breakdown of performance metrics using simulated and lab testing sets

Figure 20 shows that the performance of C2 in all FIR$_{COP}$ ranges deteriorates with the lab data than the simulated testing set. However, the bar plots for the lab data show almost similar patterns as with simulated data, i.e., the performance of C2 generally improves with increasing fault intensity.

The 6 FDD protocols studied in Yuill (2014) had varying predicting performance across different FIR_{COP} bins. The protocols did not show any predicting patterns unlike what C2 showed having increasing accuracy with increasing fault severity.

Overall, the drop in accuracy from the simulated to lab testing set is not as significant as what Chen *et al.* (2023) found when they tested their classifiers with systems different than the system their classifiers used for training.

5.4   Analyzing the Inferior Performance of C2 with Lab Testing Set

To know how generalizable the C2 is, it is necessary to analyze how it performs with each lab system (S1/S2/S3). In this section, for each of the three lab systems, the performance of C2 to different frequency of occurrence rates (FAR, MR, MDR, and accuracy) and individual class performance rates (recall, precision, and F1-score) will be analyzed to understand why the performance of C2 with lab data is much inferior to that of the simulated testing set.

Figure 21 shows the classifier C2's performance for the three lab systems regarding FAR, MR, MDR, and accuracy. The performance of C2 in terms of FAR and MR is best with lab S3, demonstrated by the lowest FAR and MR values among all lab systems. However, the performance of C2 in terms of MDR is worst with lab S3 (highest MDR value). It appears that the better performance of FAR with lab S3 is happening in expense of higher MDR value. Based on accuracy, if the performance of C2 is ordered from best to worst among the three lab systems, it is with S1, followed by S2 and S3.
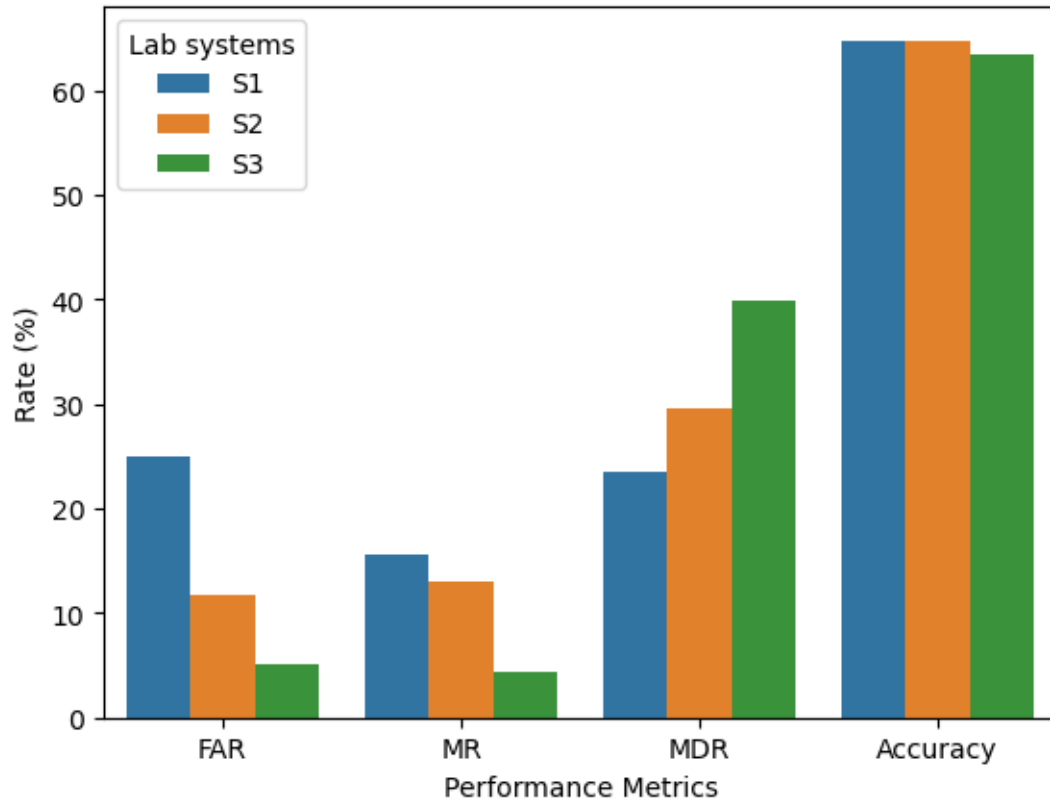
Figure 21: Classifier performance with the three lab systems

Table 18 shows the performance of C2 regarding individual class performance rates (recall, precision, and F1-score) for each lab system. In this Table, the cases that do not exist (e.g., absence of LL in lab S1) for a system show NA for all recalls. Also, if there are no predictions (no FP and TP) for a particular class, the precision value for that class is shown as NA. As a result, if any particular class has NA value in either of precision or recall, the F-1 score is shown as NA.

Table 18: Class performance rates of C2 with lab systems

Fault	Precision (%)			Recall (%)			F-1 score (%)		
	S1	S2	S3	S1	S2	S3	S1	S2	S3
CA	100	100	100	83	25	50	91	40	67
EA	85	100	100	52	68	12	65	81	21
LL	NA	NA	100	NA	NA	65	NA	NA	79
NC	0	NA	0	NA	NA	NA	NA	NA	NA
NoF	55	48	36	75	88	95	63	62	52
OC	100	100	NA	25	42	NA	40	59	NA
UC	87	100	95	80	73	53	83	85	68
VL	0	0	97	NA	NA	91	NA	NA	94

The results from Table 18 can be explained by introducing boxplots for the FIR_{COP} values for all classes in each set of the training and lab systems. The box plot of FIR_{COP} values is important because it shows a particular class's distribution of fault intensities. Thus, looking at the boxplots, if a classifier is trained with subtle fault cases for a specific class, the classifier could struggle to identify that specific class if the testing set appears with severe fault cases for that particular class. The boxplots of FIR_{COP} (without percentage form) for the 6 classes in the training and three lab systems are shown in Figure 22.

(a) CA, EA, LL

(b) OC, UC, VL
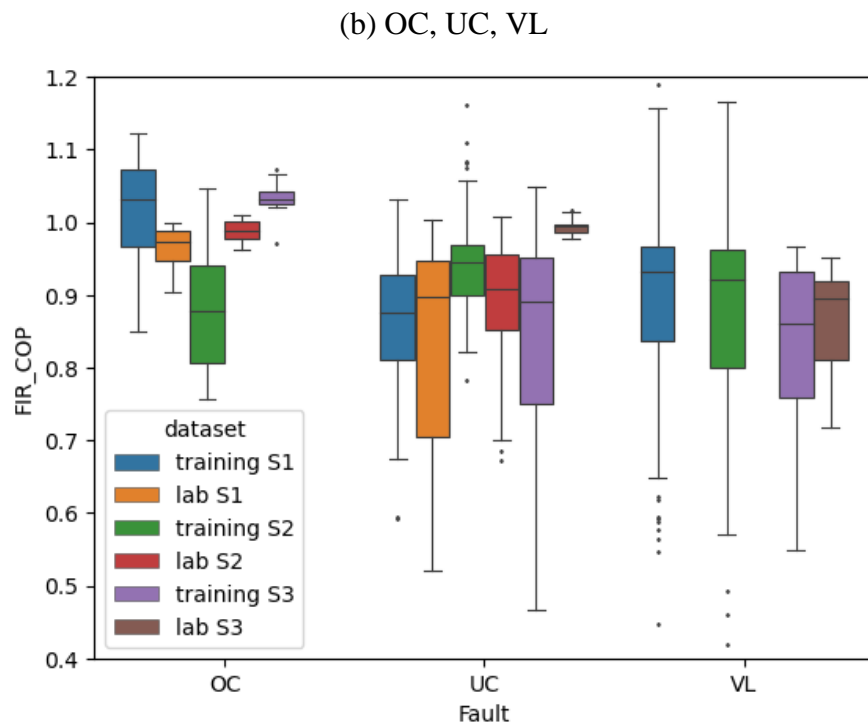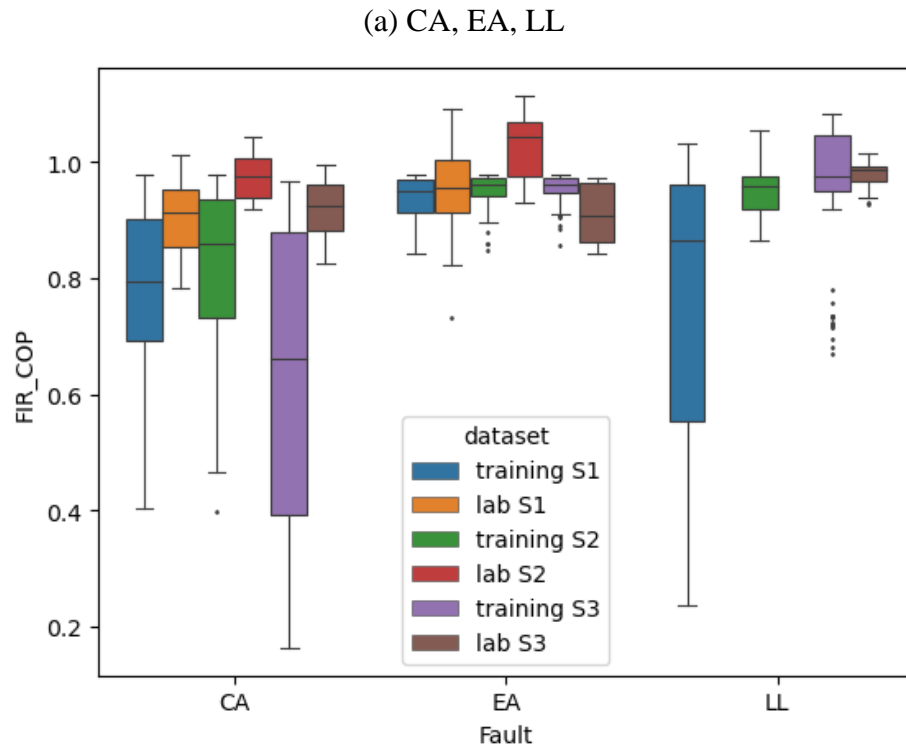
Figure 22: The distribution of FIR$_{COP}$ in training and lab testing sets

Table 19 shows the following results and discussions for the 8 fault classes by analyzing

Table 18 and Figure 22.

Table 19: Individual class performance with possible classifier improvement scopes

Fault	Performance	Possible Classifier Improvement Scope
CA	From Table 18, C2 shows best performance with S1 (F-1 score: 91%) and least performance with S2 (F-1 score: 40%).	Comparing FIR_{COP} values from Figure 22 (a), training S1, S2, and S3 have mostly severe fault cases, whereas the lab S1, S2, and S3 have mostly subtle fault cases, having lab S2 with more subtle CA cases. This could possibly lead C2 having the least performance with lab S2. Adding subtle CA fault cases in the training data might improve the predicting performance of CA faults.
EA	From Table 18, C2 is worst with S3 (F1-score = 21%) among all the three systems. Conversely, it performs best with S2 (F1-score = 81%).	Comparing FIR_{COP} values from Figure 22 (a), the EA cases in the training set have mostly subtle fault levels, whereas EA cases in lab S3 have more severe fault levels than the other systems. Thus, the classifier might follow its inherent bias toward predicting NoF for the majority of the S3 data. This could be why 23 out of 26 EA cases in S3 were miss-detected included in the 34 MD's in Figure 18. Adding severe EA fault cases in the training data might improve the predicting performance of EA faults.
LL	The LL faults only exist for lab S3 (34 cases). Table 18 shows an F1-score of 79%.	Need more lab testing data of LL faults with absent systems to check the predicting performance of C2 with those systems.
NC	There are no cases for NC; thus, the performance with lab data with this fault is unknown.	Need more lab testing data of NC faults to check the predicting performance of NC faults.
NoF	Table 18 shows that C2 is almost equally best with S1 and S3, based on the F1-score (63% and 62%). It shows comparatively bad performance with S3 (F-1 score: 52%).	The bad precision value (36%) using S3 mainly occurred due to the 23 missed-detections of EA faults. As mentioned earlier, adding severe EA fault cases in the training data might improve the predicting performance of EA faults, thus improving the precision of NoF class.

Fault	Performance	Possible Classifier Improvement Scope
OC	From Table 18, the C2 shows the worst performance with lab S1 (F-1 score: 40%).	In Figure 22 (b), lab S1 and S2 mostly show subtle OC cases. However, training S1 and S2 have mostly severe fault cases with S1's FIR_{COP} on the higher side, and S2's FIR_{COP} on the lower side. Adding more subtle OC cases could improve the predicting performance for OC. Need more lab testing data of OC faults with systems like lab S3 to check the predicting performance with such kind systems.
UC	Comparing C2's performance with lab S3, Table 18 shows better F-1 score with S1 and S2.	In Figure 22 (b), the UC cases in lab S3 are all subtle fault cases, thus 16 out of 34 UC cases are missed-detected which can be acceptable as those are subtle UC fault cases. To lower the above missed detections, more subtle UC cases can be added to the training data.
VL	The VL faults only exist for lab S3 (33 cases). It has an F1-score of 94% (Table 18).	The FIR_{COP} in the training and testing sets closely matches with each other (Figure 22 (b)), thus C2 is showing an acceptable predicting performance. Need more lab testing data of VL faults with lab systems like S1 and S2 to check the predicting performance with such systems.

From the above analysis, it appears that, for any fault class, the classifier usually shows better predicting performance if the classifier is trained with severe fault cases than the subtle fault cases for that particular class.

In addition to the above reasons, the difference in the distribution of the fault cases under different ranges of FIR_{COP} between the simulated and lab testing sets must have some impact on the degraded performance of C2 with the lab data. The distribution of fault cases under different FIR_{COP} ranges is shown in Table 20.

Table 20: Distribution of cases in testing data sets based on FIR_{COP}

Data		FIR_{COP}					Total
		>105%	95-105%	85-95%	75-85%	<75%	
Simulated	# Cases	26	451	227	77	92	873
	%	3%	52%	26%	9%	10%	100%
Lab	# Cases	10	223	90	24	14	361
	%	3%	62%	25%	6%	4%	100%

In Table 20, the distribution of data based on FIR_{COP} is not similar across the two testing sets. Both testing sets have most of the cases from FIR_{COP} of 95-105%. However, the percentage of cases of $FIR_{COP} = [95, 105]$ is higher for the lab testing set (62%) than that of the simulated testing set (52%). Also, the percentage of cases of FIR_{COP} less than 85% in the simulated testing set is 19%, higher than that of the lab testing set (10%). From the classifier (C2) performance results, the developed ML classifier performs better in predicting severe faults than subtle ones. Thus, in the lab testing set, the increased percentage of subtle fault cases and the decreased percentage of severe fault cases than that of the simulated testing set can lead to some degraded performance while predicting the lab testing set.

It is worth noting that the generalizability of the final classifier with an RTU different than S1, S2, and S3 is not tested using lab data but will be analyzed in the following section for a field-tested RTU different than S1, S2, and S3.

5.5   Performance with Field Testing Set

In this section, the final classifier's generalizability will be field tested using S4 (different

from S1, S2, and S3 in cooling capacity) with refrigerant undercharge faults. Appendix N

can be seen for related R codes and outputs. Unlike the lab testing set, the FIR_{COP} values

for those UC cases are unknown. However, the obtained UC faults are known to be

severe fault cases (FIR_{COP} values much lower than 1) marked by very high superheating

and low subcooling values.

As mentioned in the methodology section, 24 steady-state UC cases were identified for a

field-tested rooftop unit (S4). Figure 23 shows the prediction results in a confusion

matrix. Since no other fault cases exist for this system, only the associated fault classes

are shown in this figure.
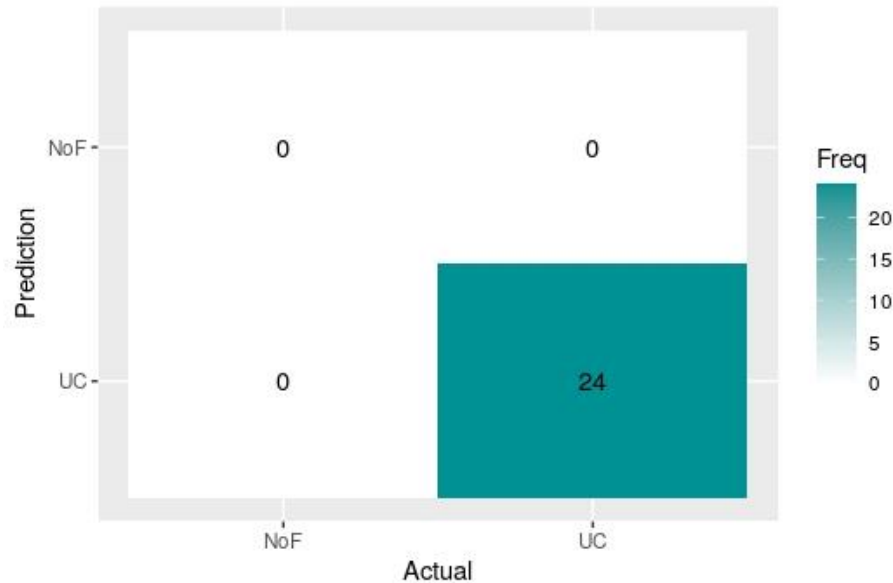
Figure 23: Confusion matrix with field testing set

As we can see, all of the UC cases were accurately predicted by classifier C2. The accuracy of the C2 from this field validation is thus found to be 100% with UC fault cases. As we can see, although our classifier was developed for RTUs with FXO, it can still detect UC faults with a system equipped with TXV, possibly because the RTU has a severe UC fault, leading its TXV to remain fully open to meet the cooling load, acting it like an FXO. This result could indicate that the classifier C2 could be treated as a generalizable classifier for predicting at least severe soft fault from any packaged rooftop unit with any of the three refrigerant types (R410a/R407c/R22).

Chapter 6. Conclusions and Future Study

In this section, the summary, potential applications, and the future steps related to this research will be presented.

6.1   Summary

To detect and diagnose the common soft faults in packaged rooftop units, this study proposes a machine-learning classifier (SVM classifier) with a reduced set of 9 relatively uncorrelated features (8 quantitative features from the filter approach, 1 qualitative refrigerant type feature) using the simulated data library from Cheung and Braun (2013a, 2013b). While analyzing the performance of this classifier with simulated, lab, and field-testing sets, the classifier shows variable but acceptable level of performance. In addition, the proposed classifier could be a generalizable classifier for predicting typical soft faults from any RTU equipped with fixed orifice valves. Following conclusions are drawn from the whole analysis,

(1) Feature selection using the filter approach (based on the correlation among 15 features) provides the most practical reduced set of 8 measurable features than the wrapper approach (backward elimination, SVM-radial classifier).

(2) While checking with the unseen simulated testing set, since the proposed 9 uncorrelated featured classifier (C2) shows little or no performance degradation compared with the full-featured and backward eliminated classifiers, these 9 features can be considered for fitting the final classifier for packaged rooftop units.

(3) The final classifier is validated using an existing lab testing set consisting of the same training systems and shows an FAR, MR, MDR, and accuracy of 12.5%, 8.5%, 34.2%, and 64%, respectively. The deteriorated performance with the lab testing set can be primarily attributed to varying fault intensities for each training and lab testing set class. The higher MDR with the lab testing set can be attributed to having the trained data largely imbalanced with unfaulty cases. The unconverted subtle faulty cases in the lab data also contributed to higher MDR.

(4) Higher missed detections from the final classifier can be acceptable since many of them were subtle fault cases.

(5) The final classifier's performance improves with the increasing fault intensity. This is true while tested with both simulated and lab testing sets.

(6) For a particular fault class and refrigerant type, a classifier trained with severe fault cases performs better predicting that class than if the classifier is trained with subtle fault cases.

(7) The developed ML classifier can be generalizable for the three refrigerant types (R410a, R407c, R22), accurately detecting most of the severe soft faults from packaged rooftop units.

(8) The proposed classifier could be generalizable for diagnosing severe refrigerant undercharge fault for RTU equipped with TXVs.

6.2   Potential Applications with Corrective Actions

The proposed classifier (C2) can be used in an IAC or P3 assessments to detect and

diagnose the common soft faults from an RTU at a facility. It can also be used to know if

the RTU is running efficiently. To predict the fault using C2 when the RTU is expected

to be running cooling mode, it would require the following data logger types:

- 3 dry-bulb temperatures

- 1 wet-bulb temperature

- 3 refrigerant temperatures, and

- 1 power meter

In a rooftop unit, the above data loggers need be installed in the following locations listed

in Table 21. It should be noted that, for personal safety, the unit should be turned off

before installing the data loggers.

Table 21: Data loggers with their installing locations for Classifier C2

Measuring Parameters	Sensor Type	Sensor Location
T_{RA}	Dry-bulb temperature	Return air duct
WB_{RA}	Wet-bulb temperature	Return air duct
T_{SA}	Dry-bulb temperature	Supply air duct
T_{amb}	Dry-bulb temperature	Outside air
T_{LL}	Refrigerant temperature	Refrigerant line just after the condenser
T_{suc}	Refrigerant temperature	Refrigerant line just before the compressor
T_{dischg}	Refrigerant temperature	Refrigerant line just after the compressor
$Power_{comp}$	Electric power	Compressor power cable

After installing the above loggers, the rooftop unit can be turned on, and the data loggers will start collecting the 8 readings at some specific intervals. The data should be collected for an extended period to get all possible operating conditions. After collecting the data, the RTU should be turned off to detach the data loggers. The logged data then needs to be transferred to a computer, and the data needs to be pre-processed based on the methodologies described in Section 3.3, 3.4, and 4.5. The pre-processing also includes adding the specific refrigerant type used in the RTU. The pre-processed steady-state cases then need to be fed into the Classifier C2 in an R software package. Finally, the Classifier C2 will predict the cases as a specific fault type or unfaulty type. In other words, there will be a predicted fault type for each steady-state case (operating condition).

The facility could do some preventive maintenance works to keep their RTU fault free. Certain routine maintenance tasks, such as cleaning the evaporator and condenser coils regularly, could prevent developing CA and EA faults.

Faults should be detected and diagnosed as early as possible, using any automated tool like the proposed classifier discussed in this study. After diagnosing a specific fault type, the facility could do some corrective actions to eliminate that fault. For example, UC fault could be eliminated by adding appropriate refrigerant amount to the system.

6.3   Future Steps

To improve the predicting performance of the proposed machine-learning classifier, the following future steps can be taken,

(1) Adding subtle CA fault cases in the training data could improve the predicting performance with CA faults.

(2) Adding severe EA fault cases in the training data could improve the predicting performance of EA faults.

(3) The lab data has random and bias errors; however, the simulated data only has bias errors (Yuill et al., 2014). Thus, training data consisting of both simulated and lab data could introduce both error characteristics in the trained classifier, improving its overall predicting performance with a new testing set.

(4) Testing the classifier with the remaining field faulty classes.

(5) Field testing the classifier with more RTU operating conditions.

In addition to the above, following studies can be performed to enrich the research knowledge in this area,

(1) Check the generalizability of a new classifier by taking EER instead of refrigerant type and normalizing compressor power with the rated cooling capacity. The proposed classifier in this dissertation has refrigerant type as a categorical feature, and the classifier can only be used for a system that has one of the three refrigerants: R410a, R407c, or R22. Thus, by taking EER as a continuous variable, the new classifier can be used for systems with any refrigerant type.

Also, in this dissertation, the compressor power was normalized with the unfaulty compressor power (obtained using compressor performance data) based on a specific mean condensing and evaporating temperatures obtained from the unfaulty operating conditions of the training data. However, the training data did not contain all the practical operating conditions, and thus the calculated unfaulty compressor power might not be accurate. To get a more accurate normalization, the actual compressor power can be normalized with the rated cooling capacity, and this normalized compressor power can be used to fit the above new classifier.

(2) Developing a classifier that can predict multiple simultaneous faults present in the system.

(3) Developing a classifier that can predict fault type from a TXV-equipped RTU.

(4) Quantifying the energy and cost savings if a specific fault type is eliminated from a system. For instance, if a system has refrigerant undercharge fault, to meet the cooling demand, the system could have an increased operating time, although the compressor power will be reduced. Since the energy consumption is mainly the multiplication of the compressor power and operating time, there could be energy saving if the above multiplication provides higher quantity than that of the unfaulty condition. This estimation could be possible by using an energy modeling software (e.g., EnergyPlus) that could simulate the undercharge fault and unfaulty cases. Similar studies could be possible for other faulty cases.

References

About NIAC | College of Engineering | University of Nebraska–Lincoln. (n.d.). Retrieved November 21, 2023, from https://engineering.unl.edu/iac/about-niac/

Albayati, M. G., Faraj, J., Thompson, A., Patil, P., Gorthala, R., & Rajasekaran, S. (2022). Semi-Supervised Machine Learning for Fault Detection and Diagnosis of a Rooftop Unit. *Big Data Mining and Analytics*, *5*. https://doi.org/10.26599/BDMA.2022.9020015

Behfar, A., Yuill, D. P., & Yu, Y. (2019). Automated fault detection and diagnosis for supermarkets–method selection, replication, and applicability. *Energy and Buildings*, *198*, 520–527. https://doi.org/10.1016/j.enbuild.2019.06.011

Brereton, R. G., & Lloyd, G. R. (2010). Support Vector Machines for classification and regression. *Analyst*, *135*(2), 230–267. https://doi.org/10.1039/b918972f

Breuker, M. S., & Braun, J. E. (1998). Common faults and their impacts for rooftop air conditioners. *HVAC&R Research*, *4*(3), 303–318. https://doi.org/10.1080/10789669.1998.10391406

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers and Electrical Engineering*, *40*(1), 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024

Chen, B., & Braun, J. E. (2001). Simple Rule-Based Methods for Fault Detection and Diagnostics Applied to Packaged Air Conditioners. *ASHRAE Transactions*, *107*,

847.
https://www.proquest.com/docview/192523498/citation/92F3287DF2184106PQ/1?a
ccountid=14692

Chen, Y., Ebrahimifakhar, A., Hu, Y., & Yuill, D. P. (2023). Generalizability of machine
learning-based fault classification for residential air-conditioners. *Energy & Buildings*, *295*(June). https://doi.org/10.1016/j.enbuild.2023.113263

Chen, Y., Ebrahimifakhar, A., & Yuill, D. P. (2022). Using Machine Learning for
Feature Selection in Automated Fault Detection and Diagnosis of Split System Air
Conditioners. *Proceedings of 19th International Refrigeration and Air Conditioning
Conference at Purdue*, *2351*, 1–10.

Cheung, H., & Braun, J. E. (2013a). Simulation of fault impacts for vapor compression
systems by inverse modeling. Part I: Component modeling and validation. *HVAC&R
Research*, *19*(7), 892–906. https://doi.org/10.1080/10789669.2013.824800

Cheung, H., & Braun, J. E. (2013b). Simulation of fault impacts for vapor compression
systems by inverse modeling. Part II: System modeling and validation. *HVAC&R
Research*, *19*(7), 907–921. https://doi.org/10.1080/10789669.2013.819769

Du, Z., Domanski, P. A., & Payne, W. V. (2015). Effect of common faults on the
performance of different types of vapor compression systems. *Applied Thermal
Engineering*, *98 (2016)*, 61–72.
https://doi.org/10.1016/j.applthermaleng.2015.11.108

Dvorak, B. I., Stewart, B. A., Hosni, A. A., Hawkey, S. A., & Nelsen, V. (2010).
Intensive Environmental Sustainability Education: Long-Term Impacts on

Workplace Behavior. *Journal of Professional Issues in Engineering Education and Practice*, *137*(2), 113–120. https://doi.org/10.1061/(ASCE)EI.1943-5541.0000054

Ebrahimifakhar, A., Kabirikopaei, A., & Yuill, D. P. (2020). Data-driven fault detection and diagnosis for packaged rooftop units using statistical machine learning classification methods. *Energy and Buildings*, *225*, 110318. https://doi.org/10.1016/j.enbuild.2020.110318

Fan, C., Liu, X., Xue, P., & Wang, J. (2021). Statistical characterization of semi-supervised neural networks for fault detection and diagnosis of air handling units. *Energy and Buildings*, *234*, 110733. https://doi.org/10.1016/J.ENBUILD.2021.110733

Farzad, M., & O'Neal, D. L. (1991). System performance characteristics of an air conditioner over a range of charging conditions. *International Journal of Refrigeration*, *14*(6), 321–328. https://doi.org/10.1016/0140-7007(91)90027-E

Fernandez, N., Katipamula, S., Wang, W., Xie, Y., & Zhao, M. (2018). *Energy savings potential from improved building controls for the US commercial building sector*. https://doi.org/10.1007/s12053-017-9569-5

Frequently Asked Questions (FAQs) - U.S. Energy Information Administration (EIA). (2021). https://www.eia.gov/tools/faqs/faq.php?id=1174&t=1

Granderson, J., Lin, G., Harding, A., Im, P., & Chen, Y. (2020). Building fault detection data to aid diagnostic algorithm creation and performance testing. *Nature*, *7*(65). https://doi.org/10.1038/s41597-020-0398-6

Hsu, C.-W., & Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines [Article]. *IEEE Transactions on Neural Networks*, *13*(2), 415–425. https://doi.org/10.1109/72.991427

Hu, Y., & Yuill, D. P. (2022a). Impacts of common faults on an air conditioner with a microtube condenser and analysis of fault characteristic features. *Energy and Buildings*, *254*. https://doi.org/10.1016/j.enbuild.2021.111630

Hu, Y., & Yuill, D. P. (2022b). Non-condensable gas in the refrigerant of air-source heat pumps: interactions between detection features, charge level, and temperature. *International Journal of Refrigeration*, *153*, 378–384. https://doi.org/10.1016/j.ijrefrig.2022.10.006

Hu, Y., Yuill, D. P., & Chen, Y. (2021). Impacts and Detection of Non-Condensable Gas in a Residential Air Source Heat Pump. *INTERNATIONAL REFRIGERATION AND AIR CONDITIONING CONFERENCE*. https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=3243&context=iracc

Hu, Y., Yuill, D. P., Ebrahimifakhar, A., & Rooholghodos, A. (2021). An experimental study of the behavior of a high efficiency residential heat pump in cooling mode with common installation faults imposed. *Applied Thermal Engineering*, *184*. https://doi.org/10.1016/j.applthermaleng.2020.116116

IAC: Search IAC Recommendations. (n.d.). Retrieved September 26, 2023, from https://iac.university/searchRecommendations

Industrial Assessment Centers. (n.d.). Retrieved November 21, 2023, from https://iac.university/

Isermann, R. (1984). Process fault detection based on modeling and estimation methods—A survey. *Automatica*, *20*(4), 387–404. https://doi.org/10.1016/0005-1098(84)90098-0

Iyengar, S., Lee, S., Irwin, D., Shenoy, P., & Weil, B. (2018). WattHome: A Data-driven Approach for Energy Efficiency Analytics at City-scale. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 396–405. https://doi.org/10.1145/3219819.3219825

JL3KH6 Job Link. (n.d.). Retrieved November 21, 2023, from https://www.fieldpiece.com/product/jl3kh6-job-link-probes-charging-and-air-kit/

Jung, D., & Sundstrom, C. (2019). A Combined Data-Driven and Model-Based Residual Selection Algorithm for Fault Detection and Isolation. *IEEE Transactions on Control Systems Technology*, *27*(2), 616–630. https://doi.org/10.1109/TCST.2017.2773514

Katipamula, S., & Brambley, M. R. (2005a). Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems-A Review, Part I. *HVAC&R Research*, *11*(1), 3–25. https://doi.org/10.1080/10789669.2005.10391123

Katipamula, S., & Brambley, M. R. (2005b). *Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems-A Review, Part II Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems-A Review, Part II*. *11*(2), 169–187. https://doi.org/10.1080/10789669.2005.10391133

Kekilova, A., Dvorak, B. I., & Williams, R. (2014). Motivations of Program Partners for Environmental Sustainability and Persistence of Benefits. *Journal of Professional*

Issues in Engineering Education and Practice, *141*(2), C4014003.
https://doi.org/10.1061/(ASCE)EI.1943-5541.0000222

Kim, M., Payne, W. V., Domanski, P. A., Yoon, S. H., & Hermes, C. J. L. (2009).
Performance of a residential heat pump operating in the cooling mode with single
faults imposed. *Applied Thermal Engineering*, *29*(4), 770–778.
https://doi.org/10.1016/j.applthermaleng.2008.04.009

Kim, W., & Braun, J. E. (2020). Development, implementation, and evaluation of a fault
detection and diagnostics system based on integrated virtual sensors and fault impact
models. *Energy and Buildings*, *226*. https://doi.org/10.1016/j.enbuild.2020.110368

Kim, W., & Katipamula, S. (2018). A review of fault detection and diagnostics methods
for building systems. *Science and Technology for the Built Environment*, *24*(1), 3–
21. https://doi.org/10.1080/23744731.2017.1318008

Kim, W., & Lee, J.-H. (2021). Fault detection and diagnostics analysis of air conditioners
using virtual sensors. *Applied Thermal Engineering*, *191*.
https://doi.org/10.1016/J.APPLTHERMALENG.2021.116848

Klein, S. A. (2018). *EES - Engineering Equation Solver* (V10.561, 2018-12-20).
https://fchartsoftware.com

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial
Intelligence*, *97*(1–2), 273–324. https://doi.org/https://doi.org/10.1016/S0004-
3702(97)00043-X

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of*

Statistical Software, *28*(5), 1–26.

https://doi.org/https://doi.org/10.18637/jss.v028.i05

Kuppig, V. D., Cook, Y. C., Carter, D. A., Larson, N. J., Williams, R. E., & Dvorak, B. I. (2016). Implementation of sustainability improvements at the facility level: Motivations and barriers. *Journal of Cleaner Production*, *139*, 1529–1538. https://doi.org/10.1016/J.JCLEPRO.2016.08.167

Langley, P. (1994). Selection of Relevant Features in Machine Learning. *Proceedings of the AAAI Fall Symposium on Relevance*, *184*, 245–271. www.aaai.org

Li, B., Cheng, F., Cai, H., Zhang, X., & Cai, W. (2021). A semi-supervised approach to fault detection and diagnosis for building HVAC systems based on the modified generative adversarial network. *Energy and Buildings*, *246*. https://doi.org/10.1016/J.ENBUILD.2021.111044

Li, H., & Braun, J. E. (2007a). A methodology for diagnosing multiple simultaneous faults in vapor-compression air conditioners. *HVAC&R Research*, *13*(2), 369–395. https://doi.org/10.1080/10789669.2007.10390959

Li, H., & Braun, J. E. (2007b). Economic Evaluation of Benefits Associated with Automated Fault Detection and Diagnosis in Rooftop Air Conditioners. *ASHRAE Transactions*, *113*(2), 200–210.

Li, Y., Liu, M., Lau, J., & Zhang, B. (2014). Experimental study on electrical signatures of common faults for packaged DX rooftop units. *Energy and Buildings*, *77*, 401–415. https://doi.org/10.1016/J.ENBUILD.2014.04.008

Llobet, E., Gualdrón, O., Vinaixa, M., El-Barbri, N., Brezmes, J., Vilanova, X., Bouchikhi, B., Gómez, R., Carrasco, J. A., & Correig, X. (2007). Efficient feature selection for mass spectrometry based electronic nose applications. *Chemometrics and Intelligent Laboratory Systems*, *85*(2), 253–261. https://doi.org/10.1016/j.chemolab.2006.07.002

Mehrabi, M., & Yuill, D. (2017a). Generalized effects of refrigerant charge on normalized performance variables of air conditioners and heat pumps. *International Journal of Refrigeration*, *76*, 367–384. https://doi.org/10.1016/j.ijrefrig.2017.02.014

Mehrabi, M., & Yuill, D. (2019). Fouling and Its Effects on Air-cooled Condensers in Split System Air Conditioners (RP-1705). *Science and Technology for the Built Environment*, *25*(6), 784–793. https://doi.org/10.1080/23744731.2019.1605197

Mehrabi, M., & Yuill, D. P. (2017b). Generalized effects of faults on normalized performance variables of air conditioners and heat pumps. *International Journal of Refrigeration*, *85*(2018), 409–430. https://doi.org/10.1016/j.ijrefrig.2017.10.017

Mirnaghi, M. S., & Haghighat, F. (2020). Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. *Energy and Buildings*, *229*. https://doi.org/10.1016/J.ENBUILD.2020.110492

Mujan, I., Anđelković, A. S., Munćan, V., Kljajić, M., & Ružić, D. (2019). Influence of indoor environmental quality on human health and productivity - A review. *Journal of Cleaner Production*, *217*, 646–657. https://doi.org/10.1016/j.jclepro.2019.01.307

Net Zero Coalition | United Nations. (n.d.). Retrieved December 28, 2022, from https://www.un.org/en/climatechange/net-zero-coalition

P3: Partners in Pollution Prevention | College of Engineering | University of Nebraska–Lincoln. (n.d.). Retrieved November 21, 2023, from

https://engineering.unl.edu/research/p3-partners-pollution-prevention/

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2022). https://www.r-project.org/

Rogers, A. P., Guo, F., & Rasmussen, B. P. (2019). A review of fault detection and diagnosis methods for residential air conditioning systems. *Building and Environment*, *161*. https://doi.org/10.1016/j.buildenv.2019.106236

Rossi, T. M. (1995). *Detection, diagnosis, and evaluation of faults in vapor compression equipment* [Purdue University].

https://www.proquest.com/docview/304211924?parentSessionId=sog%2FVZOpDU SkT7xvefzg7zTj58ol6KCrhKIDo6IEaLA%3D

Rossi, T. M. (2004). Unitary Air Conditioner Field Performance. *International Refrigeration and Air Conditioning Conference*. http://docs.lib.purdue.edu/iracc/666

Rossi, T. M., & Braun, J. E. (1995). Using Thermodynamic Impact for Detecting Refrigerant Leaks in Vapor Compression Equipment. *Proceedings of 1995 American Control Conference-ACC'95*, *6*, 4336–4340.

https://ieeexplore.ieee.org/document/532753

Rossi, T. M., & Braun, J. E. (1996). Minimizing Operating Costs of Vapor Compression Equipment with Optimal Service Scheduling. *HVAC&R Research*, *2*(1), 3–25.

https://doi.org/10.1080/10789669.1996.10391330

Rossi, T. M., & Braun, J. E. (1997). A Statistical, Rule-Based Fault Detection and Diagnostic Method for Vapor Compression Air Conditioners. *HVAC&R Research*, *3*(1), 19–37. https://doi.org/10.1080/10789669.1997.10391359

Roth, K. W., Llana, P., Feng, M., Westphalen, D., & Feng, M. Y. (2004). The Energy Impact of Faults in U.S. Commercial Buildings. *International Refrigeration and Air Conditioning Conference*. http://docs.lib.purdue.edu/iracc/665

Schantz, M. (2015). *What's on Your Roof? Rooftop Unit (RTU) Efficiency Advice and Guidance from the Advanced RTU Campaign | Department of Energy*. https://www.energy.gov/eere/buildings/articles/whats-your-roof-rooftop-unit-rtu-efficiency-advice-and-guidance-advanced

Seppänen, O. A., & Fisk, W. (2006). Some Quantitative Relations between Indoor Environmental Quality and Work Performance or Health. *HVAC&R Research*, *12*(4). https://doi.org/10.1080/10789669.2006.10391446

Wiggins, M., & Brodrick, J. (2012). HVAC Fault Detection. *ASHRAE Journal*, *54*(2), 78–80. http://tinyurl.com/TIAX2005.

Williams, R. E. (1993). *Investigation of the abrasive flow machining process and development of a monitoring strategy using acoustic emission*. University of Nebraska-Lincoln.

Yan, K., Ma, L., Dai, Y., Shen, W., Ji, Z., & Xie, D. (2018). Cost-sensitive and sequential feature selection for chiller fault detection and diagnosis. *International Journal of Refrigeration*, *86*, 401–409. https://doi.org/10.1016/j.ijrefrig.2017.11.003

Yang, H., Zhang, T., Li, H., Woradechjumroen, D., & Liu, X. (2014). HVAC Equipment, Unitary: Fault Detection and Diagnosis. *Encyclopedia of Energy Engineering and Technology, Second Edition*, *February 2015*, 854–864. https://doi.org/10.1081/e-eee2-120051345

Yang, L., Braun, J. E., & Groll, E. A. (2007). The impact of evaporator fouling and filtration on the performance of packaged air conditioners. *International Journal of Refrigeration*, *30*(3), 506–514. https://doi.org/10.1016/J.IJREFRIG.2006.08.010

Yuill, D. P. (2014). *Development of Methodologies for Evaluating Performance of Fault Detection and Diagnostics Protocols Applied to Unitary Air-Conditioning Equipment*. Purdue University.

Yuill, D. P., & Braun, J. E. (2012). Evaluating Fault Detection and Diagnostics Protocols Applied to Air-Cooled Vapor Compression Air-Conditioners. *International Refrigeration and Air Conditioning Conference*. http://docs.lib.purdue.edu/iracc/1307

Yuill, D. P., & Braun, J. E. (2013). Evaluating the performance of fault detection and diagnostics protocols applied to air-cooled unitary air-conditioning equipment. *HVAC&R Research*, *19*(7), 882–891. https://doi.org/10.1080/10789669.2013.808135

Yuill, D. P., & Braun, J. E. (2016). Effect of the distribution of faults and operating conditions on AFDD performance evaluations. *Applied Thermal Engineering*, *106*, 1329–1336. https://doi.org/10.1016/J.APPLTHERMALENG.2016.06.149

Yuill, D. P., & Braun, J. E. (2017). A figure of merit for overall performance and value of

AFDD tools. *International Journal of Refrigeration*, *74*, 651–661.
https://doi.org/10.1016/J.IJREFRIG.2016.11.015

Yuill, D. P., Cheung, H., & Braun, J. E. (2014a). Evaluating Fault Detection and
Diagnostics Tools with Simulations of Multiple Vapor Compression Systems.
International Refrigeration and Air Conditioning Conference, *July*.
http://docs.lib.purdue.edu/iracc/1543/

Yuill, D. P., Cheung, H., & Braun, J. E. (2014b). Validation of a Fault-Modeling
Equipped Vapor Compression System Model Using a Fault Detection and
Diagnostics Evaluation Tool. *International Refrigeration and Air Conditioning
Conference*. http://docs.lib.purdue.edu/iracc/1544

Yuill, D. P., & Mehrabi, M. (2017). *Simplified Models of Fault Effects on Unitary Air-
Conditioning Equipment for use in Building Simulation Tools*.
https://doi.org/10.26868/25222708.2017.378

Zhao, Y., Li, T., Zhang, X., & Zhang, C. (2019). *Artificial intelligence-based fault
detection and diagnosis methods for building energy systems: Advantages,
challenges and the future*. https://doi.org/10.1016/j.rser.2019.04.021

APPENDIX A: FIELD DATA- RTU NAMEPLATE

Figure 24: RTU nameplate info

APPENDIX B: FIELD DATA- UC FAULT DETECTION

1. Measuring Equipment Used

Two Fieldpiece pressure sensors (Fieldpiece JL3PR Job Link)

Two Fieldpiece temperature sensors (Fieldpiece JL3PC Job Link)

2. Installation of the Measuring Equipment
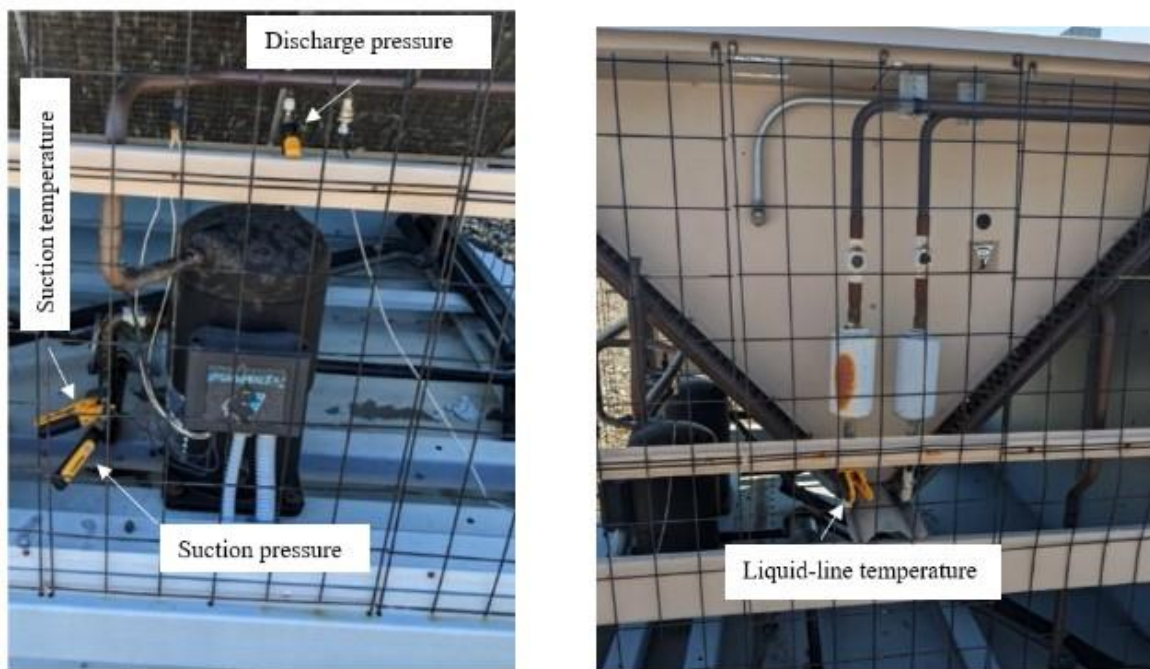
Figure 25: Installation of the one-time measuring equipment

APPENDIX B: FIELD DATA- UC FAULT DETECTION

3. Calculating Superheating (SH) and Subcooling (SC)

Measured parameters:

Suction (evaporating) pressure, P_{suc} = 54.3 psig

Discharge (condensing) pressure, P_{dischg} = 199.6 psig

Suction temperature, T_{suc} = 62.5°F

Liquid line temperature, T_{LL} = 98.8°F

From R-22 refrigerant properties:

Evaporating temperature (corresponding to P_{suc}), $T_{sat, e}$ = 29°F

Condensing temperature (corresponding to P_{dischg}), $T_{sat, c}$ = 101°F

Actual SH and SC calculation:

$$SH_{actual} = T_{suc} - T_{sat,e}$$

$$= 62.5°F - 29°F$$

$$= 33.5°F$$

$$SC_{actual} = T_{sat,c} - T_{LL}$$

$$= 101°F - 98.8°F = 2.2°F$$

APPENDIX B: FIELD DATA- UC FAULT DETECTION

Rated superheating and subcooling temperatures for Copeland Scroll ZR16M3E-TWD compressor are 20°F and 15°F, respectively.

The above calculations show that the actual SH is significantly higher than rated SH, and the actual SC is significantly lower than rated SC. Therefore, the field tested RTU is diagnosed with severe refrigerant undercharge (UC) fault.

APPENDIX C: FIELD DATA- LOGGING EQUIPMENT USED

Figure 26: Data loggers used

List of data loggers (shown in Figure 17),

a) Hobo UX120 4-channel

b) PCA-PA 8000 power meter

c) Extech RHT10 humidity logger

APPENDIX D: FIELD DATA- GETTING WB$_{RA}$ USING EES

EES software solves WB$_{RA}$ based on measured T$_{RA}$ and RH$_{RA}$. In the following figures,

WB$_{RA}$, T$_{RA}$, and RH$_{RA}$ are represented by wb, T, and R, respectively.
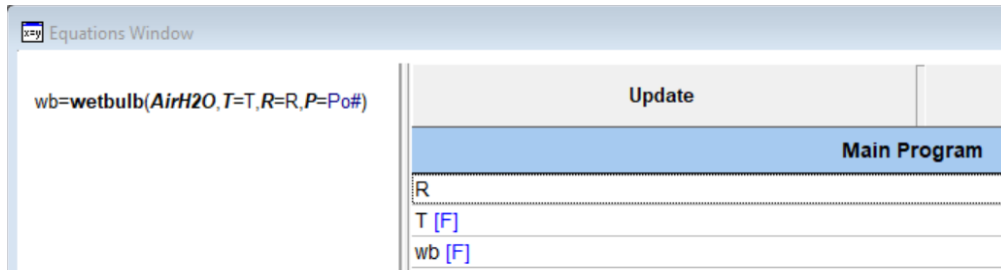
Figure 27: Screenshot of Equation Window in EES

	T [F]	R	wb [F]
Run 1	88.7	0.337	67.75
Run 2	89.4	0.3	66.75
Run 3	90.3	0.269	66.07
Run 4	91.2	0.251	65.89
Run 5	92.8	0.251	66.97
Run 6	94.6	0.237	67.51
Run 7	96.6	0.237	68.84
Run 8	98.6	0.227	69.65
Run 9	100.2	0.215	70.06
Run 10	101.8	0.212	70.93
Run 11	103.2	0.206	71.49
Run 12	104.5	0.199	71.91
Run 13	105.2	0.197	72.23
Run 14	106.1	0.196	72.74
Run 15	107.4	0.192	73.31

Figure 28: Screenshot of parametric table showing computed WB$_{RA}$

APPENDIX E: FIELD DATA- LOGGED MEASUREMENTS

Figure 29: Temperatures in return, supply, and ambient air
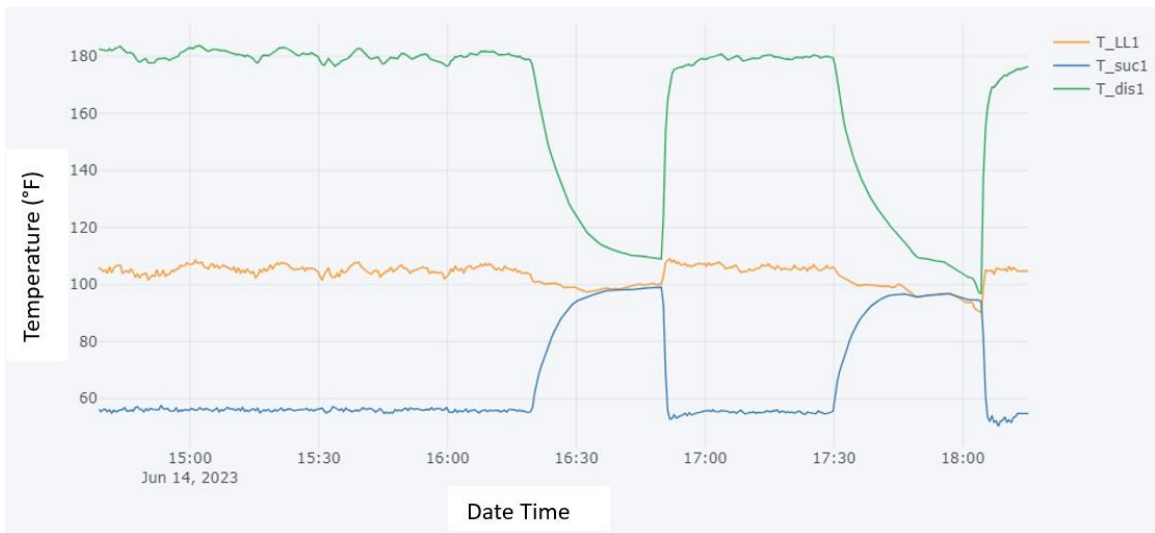
Figure 30: Temperatures in liquid line, suction, and discharge

APPENDIX E: FIELD DATA- LOGGED MEASUREMENTS

Figure 31: Logged single phase voltage
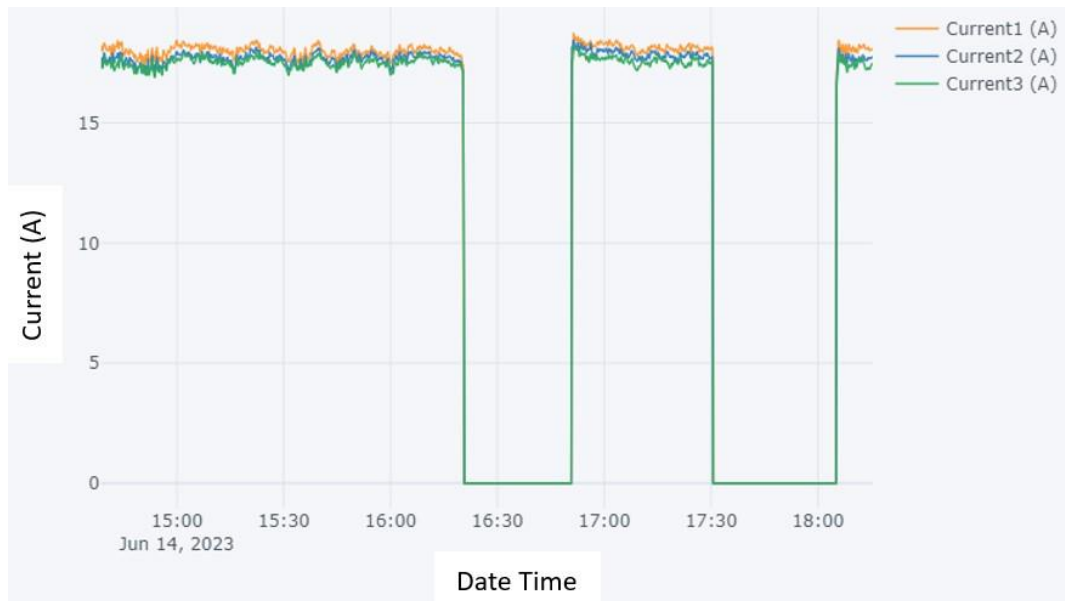
Figure 32: Logged three-phase current

APPENDIX E: FIELD DATA- LOGGED MEASUREMENTS

Figure 33: Logged compressor power (considering a power factor of 0.95)

APPENDIX E: FIELD DATA- AVERAGED MEASUREMENTS

Table 22: Averaged steady-state field data

Date Time Range	T_{RA} (°F)	WB_{RA} (°F)	T_{SA} (°F)	T_{amb} (°F)	T_{LL} (°F)	T_{suc} (°F)	T_{dischg} (°F)	$Power_{comp}$ (W)
2023-06-14 14:38:48 - 2023-06-14 14:43:48	75.65	56.94	54.23	91.78	105.55	55.81	182.48	8841.62
2023-06-14 14:44:18 - 2023-06-14 14:49:18	75.8	57.17	54.58	93.32	103.93	56.16	180.19	8754.71
2023-06-14 14:49:48 - 2023-06-14 14:54:48	76.99	57.29	54.84	96.14	103.4	56.46	178.57	8652.67
2023-06-14 14:55:18 - 2023-06-14 15:00:18	76.81	57.43	55.04	93.22	105.98	55.95	180.87	8763.62
2023-06-14 15:00:48 - 2023-06-14 15:05:48	76.2	57.48	55.21	92.11	106.95	56.16	182.71	8908.22
2023-06-14 15:06:18 - 2023-06-14 15:11:18	76.09	57.45	55.26	92.86	105.88	56.25	181.48	8915.81
2023-06-14 15:11:48 - 2023-06-14 15:16:48	75.57	57.37	55.18	92.58	104.24	55.96	179.47	8828.28
2023-06-14 15:17:18 - 2023-06-14 15:22:18	75.98	57.42	55.29	92.47	106.49	55.96	181.67	8926.92
2023-06-14 15:22:48 - 2023-06-14 15:27:48	76.14	57.46	55.43	92.82	105.64	56.12	181.08	8887.41
2023-06-14 15:28:18 - 2023-06-14 15:33:18	76.14	57.35	55.37	93.12	104.06	56.04	178.87	8792.28
2023-06-14 15:33:48 - 2023-06-14 15:38:48	77.12	57.4	55.46	95	105.12	56.03	179.25	8794.85
2023-06-14 15:39:18 - 2023-06-14 15:44:18	75.79	57.38	55.42	92.26	104.93	56.15	180.3	8868.85
2023-06-14 15:44:48 - 2023-06-14 15:49:48	76.07	57.3	55.39	93.49	105.14	55.53	179.19	8819.59
2023-06-14 15:50:18 - 2023-06-14 15:55:18	75.69	57.36	55.41	92.73	105.29	56.29	180.06	8867.53
2023-06-14 15:55:48 - 2023-06-14 16:00:48	75.92	57.28	55.32	92.95	103.92	56.2	177.95	8794.8
2023-06-14 16:01:18 - 2023-06-14 16:06:18	75.39	57.27	55.28	91.36	105.51	55.72	180.08	8912.12
2023-06-14 16:06:48 - 2023-06-14 16:11:48	75.59	57.17	55.32	90.7	106.43	55.75	181.32	8962.3
2023-06-14 16:12:18 - 2023-06-14 16:19:18	75.3	57.2	55.22	91	104.79	55.71	179.85	8883.42
2023-06-14 17:02:18 - 2023-06-14 17:07:18	75.74	57.91	56.15	89.2	106.63	55.46	179.77	8992.9
2023-06-14 17:07:48 - 2023-06-14 17:12:48	75.36	57.46	55.71	89.53	105.06	55.19	179.13	8873.21
2023-06-14 17:13:18 - 2023-06-14 17:18:18	75.28	57.19	55.49	89.17	106.17	55.54	179.95	8948.33
2023-06-14 17:18:48 - 2023-06-14 17:23:48	74.86	57.02	55.27	88.81	105.55	55.13	179.37	8909.43
2023-06-14 17:24:18 - 2023-06-14 17:29:18	75.25	57.02	55.27	89.75	105.77	55.02	179.86	8921.66
2023-06-14 18:13:18 - 2023-06-14 18:15:18	75.74	58.24	56.48	88.62	104.7	54.75	175.9	8894.98

APPENDIX F: FEATURE SELECTION

R code using filter approach:

#'RTU7' is a data frame containing all the simulated RTU7 data

Read the csv file containing RTU7 (simulated S3) data

RTU7 <- read.csv("RTU7.csv")

set.seed(502)

shuffled_RTU7 = RTU7 [sample(1:nrow(RTU7)),]

Considering 70% of the data as training data and find relevant features for this 70% data

library(caret)

set.seed(101)

rowIndex <- createDataPartition(shuffled_RTU7$Fault, p=0.7, list = F)

train <- shuffled_RTU7 [rowIndex,] #this is the data from which the filter approach will be applied.

Finding out the features that need to be removed based on correlation

ensure the results are repeatable

set.seed(7)

load the library

APPENDIX F: FEATURE SELECTION

library(mlbench)

library(caret)

calculate correlation matrix considering the 15 features

correlationMatrix3 <- cor(train[,1:15])

summarize the correlation matrix

print(correlationMatrix3)

find attributes that are highly corrected

highlyCorrelated8 <- findCorrelation(correlationMatrix3, cutoff=0.9)

highlyCorrelated9 <- findCorrelation(correlationMatrix3, cutoff=0.9, names = TRUE)

print indexes of highly correlated attributes with both col index and col name

print(highlyCorrelated8)

print(highlyCorrelated9)

APPENDIX F: FEATURE SELECTION

Output of the above R code:

```
              T_RA        WB_RA       T_SA       WB_SA      T_amb       P_LL        T_LL
T_RA       1.00000000  0.13056545  0.5063737  0.10123815 -0.02521454 -0.02707400 -0.03322909
WB_RA      0.13056545  1.00000000  0.4632515  0.94053986 -0.02145096  0.05333585  0.01598351
T_SA       0.50637373  0.46325153  1.0000000  0.66162772  0.22035033  0.19952676  0.22813260
WB_SA      0.10123815  0.94053986  0.6616277  1.00000000  0.10650775  0.15384094  0.14005901
T_amb     -0.02521454 -0.02145096  0.2203503  0.10650775  1.00000000  0.85916123  0.87819903
P_LL      -0.02707400  0.05333585  0.1995268  0.15384094  0.85916123  1.00000000  0.97521030
T_LL      -0.03322909  0.01598351  0.2281326  0.14005901  0.87819903  0.97521030  1.00000000
P_suc      0.03796421  0.20195875 -0.2274301  0.06900174  0.25813579  0.46932933  0.37227187
T_suc      0.50983383  0.45863676  0.8816886  0.60570968  0.25885284  0.22104671  0.27372502
P_dischg  -0.02980154  0.06501198  0.1409537  0.13919426  0.83163989  0.98493189  0.93531113
T_dischg   0.16413450  0.16997054  0.6881002  0.38512977  0.52864595  0.35633700  0.46826182
T_air_ce  -0.01600859  0.02174233  0.1777474  0.11997065  0.98932137  0.89296658  0.89685517
T_sat_e    0.02916347  0.18192295 -0.2839548  0.03124077  0.21619438  0.42739291  0.33209244
T_sat_c   -0.02647914  0.05578604  0.2131801  0.16205034  0.87608902  0.99039157  0.96945441
Power_comp 0.03874580  0.21185561 -0.2587562  0.06198682  0.28617214  0.45145355  0.37416574
              P_suc       T_suc      P_dischg    T_dischg    T_air_ce    T_sat_e     T_sat_c
T_RA       0.03796421  0.5098338 -0.02980154  0.1641345 -0.01600859  0.02916347 -0.02647914
WB_RA      0.20195875  0.4586368  0.06501198  0.1699705  0.02174233  0.18192295  0.05578604
T_SA      -0.22743007  0.8816886  0.14095374  0.6881002  0.17774741 -0.28395482  0.21318007
WB_SA      0.06900174  0.6057097  0.13919426  0.3851298  0.11997065  0.03124077  0.16205034
T_amb      0.25813579  0.2588528  0.83163989  0.5286459  0.98932137  0.21619438  0.87608902
P_LL       0.46932933  0.2210467  0.98493189  0.3563370  0.89296658  0.42739291  0.99039157
T_LL       0.37227187  0.2737250  0.93531113  0.4682618  0.89685517  0.33209244  0.96945441
P_suc      1.00000000 -0.1815186  0.55245913 -0.3901174  0.37940955  0.99333230  0.45064199
T_suc     -0.18151862  1.0000000  0.15679130  0.7643650  0.22937347 -0.21449498  0.22910410
P_dischg   0.55245913  0.1567913  1.00000000  0.2504691  0.87984859  0.51345157  0.98608429
T_dischg  -0.39011738  0.7643650  0.25046906  1.0000000  0.45057425 -0.43720846  0.37712077
T_air_ce   0.37940955  0.2293735  0.87984859  0.4505742  1.00000000  0.34182283  0.90863494
T_sat_e    0.99333230 -0.2144950  0.51345157 -0.4372085  0.34182283  1.00000000  0.40879068
T_sat_c    0.45064199  0.2291041  0.98608429  0.3771208  0.90863494  0.40879068  1.00000000
Power_comp 0.94406936 -0.1363444  0.53713663 -0.3330777  0.41126780  0.95579963  0.44679259
             Power_comp
T_RA        0.03874580
WB_RA       0.21185561
T_SA       -0.25875625
WB_SA       0.06198682
T_amb       0.28617214
P_LL        0.45145355
T_LL        0.37416574
P_suc       0.94406936
T_suc      -0.13634437

[1] 14  6 10 12  8 13  4
[1] "T_sat_c"  "P_LL"     "P_dischg" "T_air_ce" "P_suc"    "T_sat_e"  "WB_SA"
```

Figure 34: Outputs of filter feature selection

Therefore, using the filter approach, the features that need to be eliminated are,

"T_sat_c", "P_LL", "P_dischg", "T_air_ce", "P_suc", "T_sat_e", and "WB_SA"

APPENDIX F: FEATURE SELECTION

R code using wrapper approach:

```
set.seed(767)

subsets.svm2 <- c(1:15)

rfeCtrl.svm2 <-rfeControl (functions=caretFuncs,method="cv", number=10,

verbose=TRUE)

rfProfile.svm2 <- rfe(x=train[,-16], y=train$Fault, sizes = subsets.svm2,

rfeControl=rfeCtrl.svm2, method="svmRadial")

rfProfile.svm2

rfProfile.svm2$optVariables
```

APPENDIX F: FEATURE SELECTION

Output of the above R code:

Variables <S3: AsIs>	Accuracy <S3: AsIs>	Kappa <S3: AsIs>	AccuracySD <S3: AsIs>	KappaSD <S3: AsIs>
1	0.2458	0.1127	0.05861	0.06667
2	0.3156	0.1956	0.03535	0.04133
3	0.3776	0.2736	0.06977	0.08038
4	0.4378	0.3460	0.08303	0.09602
5	0.4869	0.4035	0.06063	0.06914
6	0.5181	0.4403	0.04661	0.05325
7	0.5659	0.4952	0.05416	0.06213
8	0.5594	0.4880	0.05528	0.06329
9	0.5563	0.4839	0.04080	0.04653
10	0.5594	0.4871	0.04465	0.05096
11	0.5605	0.4882	0.05299	0.06047
12	0.6245	0.5641	0.04319	0.04945
13	0.6308	0.5712	0.04442	0.05117
14	0.6352	0.5761	0.05230	0.06007
15	0.6394	0.5807	0.04795	0.05509

```
Recursive feature selection

Outer resampling method: Cross-Validated (10 fold)

Resampling performance over subset size:

The top 5 variables (out of 15):
   T_SA, T_suc, T_dischg, Power_comp, T_sat_e

 [1] "T_SA"       "T_suc"      "T_dischg"   "Power_comp" "T_sat_e"     "P_suc"       "T_LL"
"P_LL"       "WB_SA"      "T_sat_c"    "P_dischg"   "T_air_ce"
[13] "T_amb"       "WB_RA"      "T_RA"
```

Figure 35: Outputs of backward feature selection

Therefore, the top 8 features using backward elimination are,

"T_SA" , "T_suc", "T_dischg", "Power_comp", "T_sat_e", "P_suc", "T_LL", and "P_LL"

APPENDIX G: SAMPLE SIMULATED DATA

	T_RA	WB_RA	T_SA	WB_SA	T_amb	P_LL	T_LL	P_suc	T_suc	P_dischg	T_dischg	T_air_ce	T_
1	75	55	49.72963	43.77969	60	241.2884	66.26343	119.54017	53.05113	261.5700	137.4107	75.81244	
3	75	65	57.80031	55.11994	60	246.8553	66.11288	128.31972	65.80686	253.7083	139.0906	76.75127	
4	85	65	61.72680	54.54706	60	249.1156	66.21167	132.32284	71.04175	270.2204	147.5650	77.23656	
5	75	75	64.96741	66.85754	60	250.3059	66.03294	133.00732	70.16624	259.7509	141.9667	77.31284	
6	85	75	68.87322	66.32295	60	252.8542	66.10273	137.74104	77.91554	263.5028	147.2411	77.81829	
7	75	55	50.22203	44.64625	70	282.6233	77.74505	124.97325	44.44741	318.2577	146.1292	86.36205	
9	75	65	57.41538	55.00454	70	290.9632	77.52610	137.40090	62.17814	309.5861	150.3029	87.69081	
10	85	65	60.88751	54.55810	70	291.9635	77.59222	139.52373	66.29294	310.7547	153.1268	87.90288	
11	75	75	64.99787	66.67862	70	294.5625	77.33834	142.43641	70.04042	288.7721	146.4862	88.17280	
12	85	75	68.54858	66.12305	70	298.8184	77.43012	148.50395	75.91828	318.1743	158.3522	88.93456	
13	75	55	50.50377	44.12810	80	328.6558	89.02894	131.85183	44.05854	354.9706	152.8027	96.83306	
14	85	55	57.92961	42.48317	80	335.3678	88.92328	141.33282	56.50499	364.6364	160.8699	97.88004	
15	75	65	57.46525	55.21041	80	337.9806	88.79644	144.27972	59.08789	368.3102	162.3099	98.22882	
16	85	65	60.49960	54.83842	80	338.6147	88.85377	145.63184	61.68202	369.6413	164.3204	98.36423	
17	75	75	65.20931	66.64227	80	344.9779	88.52790	152.88682	69.60406	362.4157	164.3259	99.16176	
18	75	55	51.53008	44.62603	90	377.7496	100.20319	136.77781	41.99727	399.5463	161.1828	106.98987	
19	85	55	58.84798	42.94491	90	385.6137	100.04817	147.25845	55.73370	409.3063	169.9279	108.13237	
20	75	65	57.82405	55.62912	90	387.7278	99.93410	149.50819	57.82961	411.8537	171.0701	108.39366	
21	85	65	60.59129	55.31017	90	388.0215	99.97933	150.24290	59.35398	412.3738	172.2362	108.46434	

Showing 1 to 19 of 4,377 entries, 20 total columns

Figure 36: Sample simulated data

APPENDIX H: SAMPLE LAB DATA

	T_RA	WB_RA	T_SA	WB_SA	T_amb	P_LL	T_LL	P_suc	T_suc	P_dischg	T_dischg	T_air_ce	T
1	79.86000	62.76209	65.65944	57.78000	115.12000	479.5200	128.91000	122.61000	71.09000	488.3100	240.9200	128.77000	
2	79.94000	59.74145	60.66217	52.40000	114.61000	522.6500	127.25000	158.11000	52.03000	531.3200	198.5700	134.04000	
3	80.06000	59.87766	60.95004	52.62000	114.79000	519.0300	128.17000	157.54000	55.52000	527.8600	204.2200	134.02000	
4	80.14000	62.83141	62.42731	56.34000	115.16000	503.4600	132.20000	143.56000	69.95000	511.3100	222.6300	131.18000	
5	80.16000	62.89508	59.97840	55.44000	115.03000	517.7200	129.07000	153.99000	60.60000	527.1700	207.1600	132.55000	
6	80.21000	62.84102	59.87512	55.57000	115.05000	538.0500	123.45000	158.24000	51.51000	550.5800	170.8400	133.22000	
7	79.84000	66.50350	60.52279	57.45000	94.39000	434.8600	102.25000	158.20000	51.63000	449.0700	155.7300	116.23000	
8	79.85000	66.76952	69.97621	62.41000	94.43000	356.2500	105.82000	96.99000	71.71000	366.1400	229.3100	106.37000	
9	79.96000	66.84838	63.25841	58.95000	94.94000	398.4200	109.21000	141.81000	69.58000	409.5600	191.9000	113.11000	
10	80.09000	66.50255	60.36094	57.32000	94.76000	439.7100	101.46000	158.92000	51.57000	454.6500	148.1700	116.06000	
11	80.13000	66.81195	61.81439	58.19000	94.69000	409.0200	106.39000	151.37000	68.40000	421.6900	185.9600	114.42000	
12	80.20000	66.95005	67.99722	61.81000	94.94000	369.1900	109.00000	108.84000	71.75000	379.7900	217.3800	108.60000	
13	80.26000	66.80598	60.78864	57.62000	94.69000	426.8300	103.39000	157.79000	51.91000	440.9700	167.6000	115.94000	
14	80.38000	66.95008	62.24750	58.49000	94.73000	404.8900	107.06000	147.87000	69.16000	417.6400	186.9200	113.75000	
15	80.40000	66.89136	61.36555	57.77000	94.79000	414.4400	105.40000	154.98000	66.15000	427.9800	182.4200	114.96000	
16	80.43000	66.67265	64.07279	59.20000	94.45000	391.4200	110.44000	135.69000	69.66000	403.4200	193.8600	111.94000	
17	80.50000	67.00990	61.49295	58.11000	94.53000	415.9200	105.02000	156.14000	65.74000	429.9100	181.0500	115.12000	
18	80.61000	67.05286	60.89787	57.69000	94.74000	422.3600	104.29000	158.19000	56.82000	435.8500	172.9000	115.37000	
19	79.64000	60.18613	53.87646	50.53000	82.00000	362.0500	89.77000	142.37000	45.02000	376.3500	133.8900	101.69000	

Showing 1 to 19 of 361 entries, 20 total columns

Figure 37: Sample lab data

APPENDIX I: COMPRESSOR PERFORMANCE DATA

Rating Conditions	AIR CONDITIONING	ZP32K3E-PFV
20 °F Superheat		HFC-410A
15 F Subcooling	Blue Area Restrictions: 3-Phase Compressors Only	COPELAND SCROLL®
95 °F Ambient Air Over		PFV 208/230-1-60
60 Hz Operation		

Condensing Temperature °F (Sat. Dew Pt. Pressure, psig) — **Evaporating Temperature °F (Sat. Dew Pt. Pressure, psig)**

Condensing Temp		-10.0(36)	0.0(48)	10.0(62)	20.0(78)	30.0(97)	40.0(118)	45.0(130)	50.0(143)	55.0(156)
150.0 (613)	C						22,800	25,500	28,400	31,600
	P						4,180	4,150	4,110	4,080
	A						19.2	19.0	18.9	18.7
	M						398	442	489	539
	E						5.5	6.2	6.9	7.7
	%						55.4	58.2	60.8	63.2
145.0 (576)	C						24,200	27,000	30,000	33,300
	P						3,890	3,860	3,830	3,800
	A						17.8	17.7	17.6	17.4
	M						404	447	493	543
	E						6.2	7.0	7.8	8.8
	%						57.8	60.5	63.0	65.1
140.0 (541)	C					20,500	25,600	28,500	31,600	34,900
	P					3,680	3,630	3,600	3,570	3,540
	A					16.9	16.7	16.6	16.4	16.3
	M					331	409	452	498	547
	E					5.6	7.1	7.9	8.9	9.9
	%					54.5	60.2	62.7	65.0	66.9
130.0 (477)	C				18,100	22,700	28,200	31,300	34,600	38,100
	P				3,250	3,210	3,160	3,130	3,100	3,080
	A				15.0	14.8	14.6	14.5	14.4	14.3
	M				275	341	418	461	506	555
	E				5.6	7.1	8.9	10.0	11.2	12.4
	%				53.5	59.4	64.4	66.6	68.3	69.7
120.0 (418)	C			15,700	19,900	24,900	30,700	34,000	37,500	41,200
	P			2,860	2,840	2,800	2,750	2,730	2,710	2,690
	A			13.3	13.2	13.0	12.8	12.7	12.7	12.6
	M			227	284	350	426	468	514	562
	E			5.5	7.0	8.9	11.2	12.5	13.8	15.4
	%			52.1	58.1	63.4	67.7	69.3	70.4	71.0
110.0 (365)	C		13,400	17,200	21,700	27,000	33,100	36,500	40,200	44,200
	P		2,540	2,520	2,490	2,450	2,410	2,390	2,370	2,350
	A		11.9	11.8	11.7	11.5	11.3	11.3	11.2	11.1
	M		185	234	291	357	433	475	520	569
	E		5.3	6.8	8.7	11.0	13.7	15.3	17.0	18.8
	%		50.3	56.3	61.8	66.5	69.7	70.5	70.7	70.3
100.0 (318)	C	11,400	14,600	18,600	23,300	28,900	35,400	39,000	42,900	47,000
	P	2,240	2,240	2,220	2,190	2,150	2,110	2,100	2,080	2,070
	A	10.6	10.6	10.5	10.3	10.2	10.0	10.0	9.9	9.9
	M	150	191	239	296	362	438	481	526	575
	E	5.1	6.5	8.4	10.7	13.4	16.7	18.6	20.6	22.7
	%	48.3	54.1	59.6	64.5	68.1	69.8	69.7	68.7	66.9
90.0 (274)	C	12,300	15,700	19,900	24,900	30,700	37,500	41,300	45,400	49,800
	P	1,980	1,970	1,950	1,920	1,890	1,860	1,840	1,830	1,820
	A	9.5	9.4	9.3	9.2	9.0	8.9	8.9	8.8	8.8
	M	155	195	243	300	366	443	486	531	579
	E	6.2	8.0	10.2	12.9	16.3	20.2	22.5	24.8	27.3
	%	51.8	57.1	62.1	66.0	68.2	67.8	66.4	63.9	60.4
80.0 (236)	C	13,200	16,800	21,100	26,300	32,500	39,600	43,600	47,900	52,500
	P	1,740	1,730	1,710	1,680	1,650	1,620	1,610	1,610	1,600
	A	8.5	8.4	8.3	8.1	8.0	7.9	7.8	7.8	7.8
	M	158	198	246	303	370	447	489	535	584
	E	7.6	9.7	12.3	15.6	19.7	24.4	27.0	29.8	32.7
	%	54.6	59.4	63.5	66.1	66.4	63.3	60.2	55.8	50.1

C: Capacity (Btu/hr), P: Power (W), A: Current (Amps), M: Mass Flow (lb/hr), E: EER (Btu/Wh), %: Isentropic Efficiency (%)

Nominal Performance Values (±5%) based on 72 hours run-in. Subject to change without notice. Current @ 230 V

Copeland brand products

© 2016 Emerson Climate Technologies, Inc.
Autogenerated Compressor Performance

Printed 1/7/2016
97-515

EMERSON Climate Technologies

Figure 38: Compressor data for system S1[1]

---

[1] Search by compressor model, ZP32K3E-PFV: https://webapps.emerson.com/online-product-information/

APPENDIX I: COMPRESSOR PERFORMANCE DATA

Rating Conditions

20 °F Superheat
15 F Subcooling
95 °F Ambient Air Over

60 Hz Operation

AIR CONDITIONING

ZRD61KCE-PFV

HFC-407C - Dew Pt.
COPELAND SCROLL®
PFV 208/230-1-60

Condensing Temperature °F (Sat. Dew Pt. Pressure, psig)

Evaporating Temperature °F (Sat. Dew Pt. Pressure, psig)

Cond. Temp		-10.0 (12)	0.0 (19)	10.0 (28)	20.0 (38)	30.0 (50)	40.0 (63)	45.0 (71)	50.0 (79)	55.0 (88)
150.0 (402)	C						44,400	49,800	56,000	63,500
	P						7,150	7,250	7,350	7,500
	A						33.0	33.5	34.0	34.6
	M						743	830	927	1,036
	E						6.2	6.9	7.6	8.5
	%						57.2	59.2	61.2	63.2
140.0 (352)	C					36,100	46,800	53,500	60,500	69,000
	P					6,150	6,350	6,450	6,550	6,650
	A					27.4	28.5	29.2	29.9	30.8
	M					574	734	828	932	1,049
	E					5.9	7.4	8.3	9.3	10.4
	%					53.7	58.8	61.2	63.3	65.2
130.0 (307)	C				28,200	38,100	50,500	57,500	66,000	75,000
	P				5,250	5,400	5,600	5,700	5,850	5,950
	A				22.4	23.8	25.1	25.8	26.8	28.0
	M				428	571	741	839	949	1,070
	E				5.4	7.0	9.0	10.1	11.3	12.6
	%				48.9	55.7	61.1	63.4	65.3	66.7
120.0 (266)	C			20,600	30,100	41,100	54,500	62,500	71,000	81,000
	P			4,460	4,630	4,790	4,980	5,100	5,200	5,300
	A			18.0	19.6	20.8	22.2	23.1	24.3	25.8
	M			300	432	580	755	856	967	1,091
	E			4.6	6.5	8.6	11.0	12.3	13.7	15.2
	%			41.9	51.4	58.4	63.3	65.1	66.3	66.8
110.0 (229)	C		13,000	22,500	32,700	44,500	58,500	67,000	76,000	86,000
	P		3,790	3,940	4,090	4,240	4,430	4,530	4,650	4,790
	A		14.0	16.0	17.1	18.1	19.5	20.6	21.9	23.7
	M		184	312	445	593	768	869	980	1,103
	E		3.4	5.7	8.0	10.5	13.2	14.7	16.3	18.0
	%		31.6	45.5	54.8	60.9	64.4	65.2	65.3	64.6
100.0 (196)	C	5,520	15,400	25,100	35,500	47,500	62,000	70,000	79,500	89,500
	P	3,220	3,360	3,480	3,610	3,760	3,940	4,050	4,180	4,320
	A	10.5	12.7	13.9	14.5	15.3	16.7	17.9	19.4	21.4
	M	77	206	330	459	603	773	871	979	1,099
	E	1.7	4.6	7.2	9.8	12.6	15.7	17.3	19.0	20.7
	%	16.2	36.9	49.8	57.8	62.1	63.3	62.8	61.4	59.2
90.0 (167)	C	8,550	18,200	27,600	37,800	49,500	63,500	72,000	81,000	91,000
	P	2,860	2,960	3,060	3,180	3,320	3,520	3,630	3,770	3,920
	A	9.8	11.0	11.3	11.4	11.9	13.3	14.6	16.3	18.5
	M	110	232	347	467	601	761	853	956	1,070
	E	3.0	6.1	9.0	11.9	14.9	18.1	19.8	21.5	23.2
	%	24.1	42.9	53.7	59.2	60.9	59.1	57.0	54.0	50.2
80.0 (141)	C	11,800	20,700	29,400	38,900	49,800	63,000	71,000	80,000	89,500
	P	2,530	2,600	2,680	2,790	2,940	3,140	3,270	3,410	3,580
	A	8.3	8.4	7.9	7.5	7.7	9.0	10.3	12.2	14.6
	M	144	252	354	460	579	723	808	903	1,009
	E	4.7	8.0	11.0	13.9	17.0	20.1	21.8	23.4	25.0
	%	32.6	47.9	55.5	57.8	56.1	51.1	47.4	42.9	37.6

C: Capacity (Btu/hr), P: Power (W), A: Current (Amps), M: Mass Flow (lb/hr), E: EER (Btu/Wh), %: Isentropic Efficiency (%)

Nominal Performance Values (±5%) based on 72 hours run-in. Subject to change without notice. Current @ 230 V

Copeland
brand products

© 2016   Emerson Climate Technologies, Inc.
Autogenerated Compressor Performance

Printed   1/6/2016
07-367

EMERSON
Climate Technologies

Figure 39: Compressor data for system S2[2]

---

[2] Search by compressor model, ZRD61KCE-PFV: https://webapps.emerson.com/online-product-information/

APPENDIX I: COMPRESSOR PERFORMANCE DATA

Rating Conditions	AIR CONDITIONING	CRH3-0275-TFD
20 °F Superheat		COPELAWELD® HCFC-22
15 F Subcooling		COMPRESSOR
95 °F Ambient Air Over		TFD 460-3-60
60 Hz Operation		

Condensing Temperature °F (Sat. Dew Pt. Pressure, psig) — Evaporating Temperature °F (Sat. Dew Pt. Pressure, psig)

Cond.		-10.0(17)	0.0(24)	10.0(33)	20.0(43)	30.0(55)	40.0(69)	45.0(76)	50.0(84)	55.0(93)
150.0 (382)	C					16,200	22,400	26,100	30,200	34,700
	P					3,050	3,500	3,720	3,930	4,130
	A					4.5	5.1	5.3	5.6	5.8
	M					271	368	426	489	558
	E					5.3	6.4	7.0	7.7	8.4
	%					52.8	55.6	57.0	58.2	59.2
140.0 (337)	C				13,800	19,200	26,200	30,300	34,900	39,600
	P				2,560	2,990	3,380	3,570	3,750	3,920
	A				4.0	4.5	4.9	5.1	5.3	5.5
	M				221	304	406	469	535	607
	E				5.4	6.4	7.7	8.5	9.3	10.2
	%				53.3	55.9	56.7	59.9	60.8	61.5
130.0 (297)	C			11,300	16,000	22,200	30,000	34,500	39,400	44,800
	P			2,140	2,540	2,910	3,240	3,400	3,550	3,690
	A			3.5	3.9	4.4	4.7	4.9	5.1	5.3
	M			175	244	333	444	507	575	649
	E			5.3	6.3	7.6	9.2	10.1	11.1	12.1
	%			52.8	55.1	58.1	60.7	61.6	62.2	62.3
120.0 (260)	C			12,900	18,300	25,200	33,700	38,500	43,800	49,600
	P			2,140	2,490	2,800	3,080	3,210	3,340	3,460
	A			3.5	3.9	4.2	4.5	4.7	4.8	5.0
	M			191	266	361	476	541	611	687
	E			6.0	7.3	9.0	10.9	12.0	13.1	14.3
	%			53.7	56.3	59.5	61.6	62.2	62.3	61.9
110.0 (225)	C		10,100	14,700	20,600	28,200	37,300	42,600	48,200	54,500
	P		1,780	2,110	2,410	2,670	2,900	3,010	3,110	3,210
	A		3.1	3.4	3.8	4.1	4.3	4.4	4.6	4.7
	M		145	207	287	386	505	572	644	721
	E		5.7	6.9	8.6	10.6	12.9	14.1	15.5	16.9
	%		50.6	53.7	57.2	60.1	61.6	61.6	61.1	59.9
100.0 (196)	C		11,500	16,500	23,100	31,200	41,000	46,500	52,500	59,000
	P		1,780	2,050	2,300	2,520	2,700	2,790	2,870	2,950
	A		3.1	3.4	3.7	3.9	4.1	4.2	4.3	4.3
	M		158	224	308	411	532	601	674	751
	E		6.5	8.0	10.0	12.4	15.2	16.7	18.3	20.0
	%		50.7	54.3	57.8	60.0	60.4	59.7	58.4	56.3
90.0 (168)	C	8,920	13,000	18,600	25,700	34,300	44,700	50,500	56,500	63,500
	P	1,460	1,730	1,970	2,170	2,340	2,490	2,560	2,620	2,680
	A	2.7	3.0	3.3	3.5	3.7	3.8	3.9	4.0	4.0
	M	119	172	241	329	434	558	627	701	779
	E	6.1	7.5	9.4	11.8	14.7	17.9	19.7	21.7	23.7
	%	48.5	51.5	55.1	58.0	59.1	58.0	56.3	53.9	50.6
80.0 (144)	C	10,300	14,800	20,800	28,300	37,500	48,300	54,500	61,000	68,000
	P	1,430	1,660	1,850	2,010	2,150	2,260	2,310	2,360	2,400
	A	2.7	2.9	3.2	3.3	3.5	3.6	3.6	3.6	3.7
	M	132	188	260	350	457	583	652	726	805
	E	7.2	8.9	11.2	14.1	17.4	21.4	23.5	25.8	28.3
	%	50.1	52.9	55.9	57.6	57.2	53.9	51.0	47.1	42.2

C: Capacity (Btu/hr), P: Power (W), A: Current (Amps), M: Mass Flow (lb/hr), E: EER (Btu/Wh), %: Isentropic Efficiency (%)

Nominal Performance Values (±5%) based on 72 hours run-in. Subject to change without notice. Current @ 460 V

Copeland brand products

© 2016 Emerson Climate Technologies, Inc.
Autogenerated Compressor Performance

Printed 1/4/2016
P81-84R1

EMERSON Climate Technologies

Figure 40: Compressor data for system S3[3]

---

[3] Search by compressor model, CRH3-0275-TFD: https://webapps.emerson.com/online-product-information/

APPENDIX I: COMPRESSOR PERFORMANCE DATA

Rating Conditions	AIR CONDITIONING	ZR16M3E-TWD
20 °F Superheat		HCFC-22
15 F Subcooling		COPELAND SCROLL®
95 °F Ambient Air Over		TWD 460-3-60
60 Hz Operation		

Condensing Temperature °F (Sat. Dew Pt. Pressure, psig)

Evaporating Temperature °F (Sat. Dew Pt. Pressure, psig)

Cond.		-10.0 (17)	0.0 (24)	10.0 (33)	20.0 (43)	30.0 (55)	40.0 (69)	45.0 (76)	50.0 (84)	55.0 (93)
150.0 (382)	C						120,000	133,000	147,000	162,000
	P						16,900	16,900	16,900	16,900
	A						23.6	23.6	23.6	23.5
	M						1,968	2,172	2,386	2,611
	E						7.1	7.9	8.7	9.6
	%						61.7	64.0	66.1	67.8
140.0 (337)	C					105,000	130,000	144,000	159,000	175,000
	P					15,100	15,100	15,100	15,100	15,100
	A					21.3	21.2	21.2	21.2	21.2
	M					1,653	2,027	2,227	2,436	2,662
	E					6.9	8.6	9.5	10.5	11.6
	%					60.4	65.2	67.2	68.8	70.1
130.0 (297)	C				90,000	113,000	140,000	154,000	170,000	186,000
	P				13,400	13,500	13,600	13,500	13,500	13,500
	A				19.2	19.2	19.2	19.1	19.1	19.2
	M				1,370	1,704	2,071	2,269	2,479	2,703
	E				6.7	8.4	10.3	11.4	12.5	13.8
	%				58.4	63.8	67.8	69.2	70.3	71.0
120.0 (260)	C			75,500	97,500	121,000	149,000	164,000	180,000	198,000
	P			12,000	12,100	12,200	12,200	12,200	12,200	12,200
	A			17.5	17.5	17.5	17.5	17.5	17.5	17.5
	M			1,114	1,414	1,740	2,102	2,301	2,512	2,738
	E			6.3	8.0	10.0	12.2	13.5	14.8	16.3
	%			55.5	61.6	66.0	69.0	69.8	70.2	70.1
110.0 (226)	C		61,500	81,500	104,000	129,000	157,000	173,000	190,000	208,000
	P		10,700	10,800	10,900	10,900	10,900	11,000	11,000	11,000
	A		16.0	16.1	16.1	16.1	16.0	16.0	16.1	16.1
	M		884	1,153	1,442	1,763	2,125	2,324	2,536	2,767
	E		5.8	7.5	9.5	11.8	14.3	15.8	17.3	18.9
	%		51.6	58.5	63.5	66.9	68.7	68.7	68.2	67.0
100.0 (195)	C	48,600	67,000	87,000	109,000	135,000	165,000	181,000	199,000	219,000
	P	9,450	9,650	9,750	9,850	9,850	9,900	9,950	10,000	10,100
	A	14.6	14.8	14.9	14.9	14.8	14.8	14.9	14.9	15.0
	M	677	918	1,174	1,457	1,776	2,140	2,342	2,559	2,792
	E	5.1	6.9	8.9	11.1	13.7	16.6	18.2	20.0	21.7
	%	46.4	54.3	60.1	64.0	66.2	66.2	65.3	63.7	61.3
90.0 (168)	C	53,000	71,000	91,000	114,000	141,000	172,000	190,000	209,000	229,000
	P	8,600	8,750	8,800	8,900	8,950	9,050	9,100	9,200	9,350
	A	13.7	13.8	13.9	13.8	13.8	13.8	13.9	13.9	14.1
	M	708	935	1,183	1,462	1,781	2,151	2,356	2,578	2,816
	E	6.2	8.1	10.3	12.8	15.8	19.0	20.8	22.7	24.6
	%	49.0	55.6	60.2	63.0	63.6	61.6	59.4	56.4	52.5
80.0 (144)	C	56,000	74,000	94,500	118,000	146,000	179,000	198,000	218,000	239,000
	P	7,800	7,900	7,950	8,050	8,150	8,350	8,450	8,600	8,800
	A	12.8	12.9	12.9	12.9	12.9	12.9	13.0	13.1	13.3
	M	722	938	1,180	1,459	1,782	2,157	2,368	2,594	2,839
	E	7.2	9.4	11.8	14.7	17.9	21.5	23.4	25.3	27.2
	%	50.2	55.5	58.9	60.1	58.8	54.2	50.6	46.1	40.6

C: Capacity (Btu/hr), P: Power (W), A: Current (Amps), M: Mass Flow (lb/hr), E: EER (Btu/Wh), %: Isentropic Efficiency (%)

Nominal Performance Values (±5%) based on 72 hours run-in. Subject to change without notice. Current @ 460 V

Copeland brand products

© 2015   Emerson Climate Technologies, Inc.
Autogenerated Compressor Performance

Printed   12/31/2015
03-750

EMERSON Climate Technologies

Figure 41: Compressor data for system S4[4]

---

[4] Search by compressor model, ZR16M3E-TWD: https://webapps.emerson.com/online-product-information/

APPENDIX J: CLASSIFIER C1 (BACKWARD FEATURED)

R code:

```
#function that returns FAR, MR, MDR, and accuracy

perf.metrics1 <- function(predicted, actual) {

library(caret)

conf.mat1=confusionMatrix(predicted, actual)

#find FAR

recall_NoF=conf.mat1$byClass[45]

FAR=1-recall_NoF

#find MR

class.number=5

true.loc = (class.number-1)*8+class.number

TP.NoF = conf.mat1$table[true.loc]

total.actual=sum(conf.mat1$table)

presence=total.actual-sum(actual=='NoF')-sum(predicted=='NoF')+TP.NoF

presence1=total.actual-sum(actual=='NoF')

trueDiag.woNoF=sum(diag(conf.mat1$table))-TP.NoF
```

APPENDIX J: CLASSIFIER C1 (BACKWARD FEATURED)

Misd=presence-trueDiag.woNoF

MR=Misd/presence

#find MDR

MD=sum(predicted=='NoF')-TP.NoF

MDR=MD/presence1

#find accuracy

sum.TP=sum(diag(conf.mat1$table))

accuracy=sum.TP/total.actual

#find balanced accuracy

bal_acc=mean(conf.mat1$byClass[41:48], na.rm=TRUE)

perf=data.frame(FAR=FAR, MR=MR, MDR=MDR, accuracy=accuracy,

bal_acc=bal_acc)

return(perf)

}

function for fitting C1 and testing with simulated data

fun95 <- function(lower, upper, data){

  for (i in 1:nrow(data)){

APPENDIX J: CLASSIFIER C1 (BACKWARD FEATURED)

```
if (data[i, 20]>lower & data[i, 20]<upper){

data[i, 19]="NoF"

  }

 }
```

Load the required library

```
library(dplyr)
```

Define the specific numbers for each category

```
specific_numbers <- data.frame (Refrigerant = c("R410a", "R407c", "R22"), power_NoF
= c(2720.35, 5639.18, 3106.93), T_dischg_NoF = c(178.92, 193.26, 204.10))
```

Merge the specific_numbers data frame with my data based on the category column

```
data <- left_join (data, specific_numbers, by = "Refrigerant")  #after this line,
power_NoF and T_dischg_NoF columns will be created
```

Create the new column by dividing the power column by the corresponding divisor

```
data <- data %>%

  mutate(FIR_power = Power_comp / power_NoF) #after this line, FIR_power column
will be created. This is the normalized power.

data <- data %>%
```

APPENDIX J: CLASSIFIER C1 (BACKWARD FEATURED)

```
mutate(FIR_T_dischg = T_dischg / T_dischg_NoF) #after this line, FIR_T_dischg

column will be created

#convert Refrigerant column into factor   #seems unnecessary

data$Refrigerant <- factor (data$Refrigerant, levels = c("R410a", "R407c", "R22"))

library(caret)

set.seed(101)

data.shuffle= data[sample(1:nrow(data)), ]

set.seed(101)

rowIndex <- createDataPartition (data.shuffle$Fault, p=0.8, list = F)

train.sim <- data.shuffle[rowIndex, ]

test.sim <- data.shuffle[-rowIndex, ]

#take 16 features and 1 response variable (Fault)

train.sim.15=train.sim[,c(1:10,24,12,13,14,23,18,19)]

test.sim.15=test.sim[,c(1:10,24,12,13,14,23,18,19)]

#take 9 backward features (including Refrigerant type as feature) 1 response variable

(Fault)

train.sim.8=train.sim.15[,c(3,6,7,8,9,11,13,15,16,17)]

test.sim.8=test.sim.15[,c(3,6,7,8,9,11,13,15,16,17)]
```

APPENDIX J: CLASSIFIER C1 (BACKWARD FEATURED)

#preprocess now

library(caret)

preProc10 <- preprocess (train.sim.8, method = c("center", "scale")) #preprocessing will be done based on mean & sd from train.sim.8

#Apply the processing to the train and test data

train.sim.8.scaled <- predict(preProc10, train.sim.8)

test.sim.8.scaled <- predict(preProc10, test.sim.8)

#tuning & fitting model

library(caret)

grid <- expand.grid(C=c(1, 10, 10^2,10^3,10^4,10^5,10^6,10^7,10^8,10^9), sigma=c(1e-07,1e-06,1e-05,1e-04,1e-03,1e-02,1e-01,1)) #grid search

set.seed(1234)

svmrad.caret <- train(Fault ~ ., data = train.sim.8.scaled, method = "svmRadial", tuneGrid = grid, trControl = trainControl(method = "cv", number = 5, verboseIter = TRUE))

best.para=svmrad.caret$bestTune

#fitting model using e1071 package

library(e1071)

APPENDIX J: CLASSIFIER C1 (BACKWARD FEATURED)

```
model = svm(Fault~.,data=train.sim.8.scaled, gamma=svmrad.caret$bestTune[1],

cost=svmrad.caret$bestTune[2])

#predict using train data

pred1 = predict(model, newdata = test.sim.8.scaled[,-10])

#confusion matrix

conf.matx1 = table(pred1, test.sim.8.scaled[,10]) #whole lab data

acc1 = 1-mean(pred1 !=test.sim.8.scaled[,10])  #whole lab data

metrics=perf.metrics1(pred1, test.sim.8.scaled[,10])

newList <- list("train.sim" = train.sim, "test.sim" = test.sim, "data.shuffle" = data.shuffle,

"train.sim.15" = train.sim.15, "test.sim.15" = test.sim.15, "preProc"=preProc10,

"train.sim.8.scaled" = train.sim.8.scaled, "test.sim.8.scaled" = test.sim.8.scaled,

"best.para"=best.para, "model"= model,"pred" = pred1,

"con_mat_whole"=conf.matx1,"acc_whole"=acc1, "metrics"=metrics)

return(newList)

}

#Inputting FIR_COP thresholds and raw simulated data ('sim.data.347.1')

mru116_0.98_1.02_wo_smote= fun95(0.98, 1.02, sim.data.347.1)

mru116_0.98_1.02_wo_smote$con_mat_whole
```

APPENDIX J: CLASSIFIER C1 (BACKWARD FEATURED)

mru116_0.98_1.02_wo_smote$best.para

mru116_0.98_1.02_wo_smote$metrics

Outputs from the above R code:

```
pred1  CA  EA  LL  NC  NoF  OC  UC  VL
  CA  129   0   0   2   14   1   0   0
  EA    0  63   0   0   13   0   0   0
  LL    0   0  43   0    5   0   0   0
  NC    2   0   0  51    5   2   0   2
 NoF   24  10   4  11  216   9   2   7
  OC    1   0   0  10    7  58   0   0
  UC    0   0   2   1    2   0  74   0
  VL    0   0   0   0    6   0   0  97
```

sigma	C
<dbl>	<dbl>
0.1	1e+05

FAR	MR	MDR	accuracy	bal_acc
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
0.1940299	0.04275093	0.1107438	0.8373425	0.846351

Figure 42: Outputs from classifier C1

APPENDIX K: CLASSIFIER C2 (UNCORRELATED FEATURED)

R code:

```
# function for fitting C2 and testing with simulated data

fun83 <- function (lower, upper, data){

  for (i in 1:nrow(data)){

    if (data[i, 20]>lower & data[i, 20]<upper){

data[i, 19]="NoF"

    }

  }
```

```
# Load the required library

library(dplyr)
```

```
# Define the specific numbers for each category

specific_numbers <- data.frame (Refrigerant = c("R410a", "R407c", "R22"), power_NoF

= c(2720.35, 5639.18, 3106.93), T_dischg_NoF = c(178.92, 193.26, 204.10))
```

```
# Merge the specific_numbers data frame with my data based on the category column

data <- left_join (data, specific_numbers, by = "Refrigerant")  #after this line,

power_NoF and T_dischg_NoF columns will be created
```

```
# Create the new column by dividing the power column by the corresponding divisor

data <- data %>%
```

APPENDIX K: CLASSIFIER C2 (UNCORRELATED FEATURED)

mutate(FIR_power = Power_comp / power_NoF) #after this line, FIR_power column will be created. This is the normalized power.

data <- data %>%

  mutate(FIR_T_dischg = T_dischg / T_dischg_NoF) #after this line, FIR_T_dischg column will be created

#convert Refrigerant column into factor   #seems unnecessary

data$Refrigerant <- factor (data$Refrigerant, levels = c("R410a", "R407c", "R22"))

library(caret)

set.seed(101)

data.shuffle= data[sample(1:nrow(data)),]

set.seed(101)

rowIndex <- createDataPartition (data.shuffle$Fault, p=0.8, list = F)

train.sim <- data.shuffle[rowIndex,]

test.sim <- data.shuffle[-rowIndex,]

#take 16 features and 1 response variable (Fault)

train.sim.15=train.sim[,c(1:10,24,12,13,14,23,18,19)]

test.sim.15=test.sim[,c(1:10,24,12,13,14,23,18,19)]

APPENDIX K: CLASSIFIER C2 (UNCORRELATED FEATURED)

#take 9 uncorrelated features (including Refrigerant type as feature) and 1 response variable (Fault)

train.sim.8=train.sim.15[,c(1,2,3,5,7,9,11,15,16,17)]

test.sim.8=test.sim.15[,c(1,2,3,5,7,9,11,15,16,17)]

#preprocess now

library(caret)

preProc10 <- preprocess (train.sim.8, method = c("center", "scale")) #preprocessing will be done based on mean & sd from train.sim.8

#Apply the processing to the train and test data

train.sim.8.scaled <- predict(preProc10, train.sim.8)

test.sim.8.scaled <- predict(preProc10, test.sim.8)

#tuning & fitting model

library(caret)

grid <- expand.grid(C=c(1, 10, 10^2,10^3,10^4,10^5,10^6,10^7,10^8,10^9),

sigma=c(1e-07,1e-06,1e-05,1e-04,1e-03,1e-02,1e-01,1)) #grid search

set.seed(1234)

APPENDIX K: CLASSIFIER C2 (UNCORRELATED FEATURED)

svmrad.caret <- train(Fault ~ ., data = train.sim.8.scaled, method = "svmRadial",

tuneGrid = grid, trControl = trainControl(method = "cv", number = 5, verboseIter =

TRUE))

best.para=svmrad.caret$bestTune

#fitting model using e1071 package

library(e1071)

model = svm(Fault~.,data=train.sim.8.scaled, gamma=svmrad.caret$bestTune[1],

cost=svmrad.caret$bestTune[2])

#predict using train data

pred1 = predict(model, newdata = test.sim.8.scaled[,-10])

#confusion matrix

conf.matx1 = table(pred1, test.sim.8.scaled[,10]) #whole lab data

acc1 = 1-mean(pred1 !=test.sim.8.scaled[,10]) #whole lab data

metrics=perf.metrics1(pred1, test.sim.8.scaled[,10])

newList <- list("train.sim" = train.sim, "test.sim" = test.sim, "data.shuffle" = data.shuffle,

"train.sim.15" = train.sim.15, "test.sim.15" = test.sim.15, "preProc"=preProc10,

"train.sim.8.scaled" = train.sim.8.scaled, "test.sim.8.scaled" = test.sim.8.scaled,

"best.para"=best.para, "model"= model,"pred" = pred1,

"con_mat_whole"=conf.matx1,"acc_whole"=acc1, "metrics"=metrics)

APPENDIX K: CLASSIFIER C2 (UNCORRELATED FEATURED)

return(newList)

}

#Inputting FIR_COP thresholds and raw simulated data ('sim.data.347.1')

mru104_0.98_1.02_wo_smote= fun83(0.98, 1.02, sim.data.347.1)

mru104_0.98_1.02_wo_smote$con_mat_whole

mru104_0.98_1.02_wo_smote$best.para

mru104_0.98_1.02_wo_smote$metrics

Outputs from the above R code:

```
pred1  CA  EA  LL  NC NoF  OC  UC  VL
  CA  147   0   0   3   2   0   0   0
  EA    0  67   0   0   3   0   0   0
  LL    0   0  49   0   4   0   0   0
  NC    5   1   0  66   4   0   0   1
 NoF    4   5   0   6 234  11   3   5
  OC    0   0   0   0   9  59   0   0
  UC    0   0   0   0   4   0  73   0
  VL    0   0   0   0   8   0   0 100
```

sigma	C
<dbl>	<dbl>
0.1	1000

FAR	MR	MDR	accuracy	bal_acc
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
0.1268657	0.01751313	0.05619835	0.9106529	0.9200037

Figure 43: Outputs from classifier C2

APPENDIX L: CLASSIFIER C3 (FULL FEATURED)

R code:

```
# function for fitting C3 and testing with simulated data

fun94 <- function (lower, upper, data){

  for (i in 1:nrow(data)){

    if (data[i, 20]>lower & data[i, 20]<upper){

data[i, 19]="NoF"

    }

  }
```

```
# Load the required library

library(dplyr)
```

```
# Define the specific numbers for each category

specific_numbers <- data.frame (Refrigerant = c("R410a", "R407c", "R22"), power_NoF
= c(2720.35, 5639.18, 3106.93), T_dischg_NoF = c(178.92, 193.26, 204.10))
```

```
# Merge the specific_numbers data frame with my data based on the category column

data <- left_join (data, specific_numbers, by = "Refrigerant")  #after this line,
power_NoF and T_dischg_NoF columns will be created
```

```
# Create the new column by dividing the power column by the corresponding divisor
```

APPENDIX L: CLASSIFIER C3 (FULL FEATURED)

data <- data %>%

mutate(FIR_power = Power_comp / power_NoF) #after this line, FIR_power column will

be created. This is the normalized power.

data <- data %>%

  mutate(FIR_T_dischg = T_dischg / T_dischg_NoF) #after this line, FIR_T_dischg

column will be created

#convert Refrigerant column into factor   #seems unnecessary

data$Refrigerant <- factor (data$Refrigerant, levels = c("R410a", "R407c", "R22"))

library(caret)

set.seed(101)

data.shuffle= data[sample(1:nrow(data)),]

set.seed(101)

rowIndex <- createDataPartition (data.shuffle$Fault, p=0.8, list = F)

train.sim <- data.shuffle[rowIndex,]

test.sim <- data.shuffle[-rowIndex,]

#take 16 features and 1 response variable (Fault)

train.sim.15=train.sim[,c(1:10,24,12,13,14,23,18,19)]

test.sim.15=test.sim[,c(1:10,24,12,13,14,23,18,19)]

APPENDIX L: CLASSIFIER C3 (FULL FEATURED)

#preprocess now

library(caret)

preProc10 <- preprocess (train.sim.15, method = c("center", "scale")) #preprocessing will be done based on mean & sd from train.sim.15

#Apply the processing to the train and test data

train.sim.15.scaled <- predict(preProc10, train.sim.15)

test.sim.15.scaled <- predict(preProc10, test.sim.15)

#tuning & fitting model

library(caret)

grid <- expand.grid(C=c(1, 10, 10^2,10^3,10^4,10^5,10^6,10^7,10^8,10^9), sigma=c(1e-07,1e-06,1e-05,1e-04,1e-03,1e-02,1e-01,1)) #grid search

set.seed(1234)

svmrad.caret <- train(Fault ~ ., data = train.sim.15.scaled, method = "svmRadial", tuneGrid = grid, trControl = trainControl(method = "cv", number = 5, verboseIter = TRUE))

best.para=svmrad.caret$bestTune

#fitting model using e1071 package

library(e1071)

APPENDIX L: CLASSIFIER C3 (FULL FEATURED)

```
model = svm(Fault~.,data=train.sim.15.scaled, gamma=svmrad.caret$bestTune[1],

cost=svmrad.caret$bestTune[2])

#predict using train data

pred1 = predict(model, newdata = test.sim.15.scaled[,-17])

#confusion matrix

conf.matx1 = table(pred1, test.sim.15.scaled[,17]) #whole lab data

acc1 = 1-mean(pred1 !=test.sim.15.scaled[,17])  #whole lab data

metrics=perf.metrics1(pred1, test.sim.15.scaled[,17])

newList <- list("train.sim" = train.sim, "test.sim" = test.sim, "data.shuffle" = data.shuffle,

"train.sim.15" = train.sim.15, "test.sim.15" = test.sim.15, "preProc"=preProc10,

"train.sim.15.scaled" = train.sim.15.scaled, "test.sim.15.scaled" = test.sim.15.scaled,

"best.para"=best.para, "model"= model,"pred" = pred1,

"con_mat_whole"=conf.matx1,"acc_whole"=acc1, "metrics"=metrics)

return(newList)

}

#Inputting FIR_COP thresholds and raw simulated data ('sim.data.347.1')

mru115_0.98_1.02_wo_smote= fun94(0.98, 1.02, sim.data.347.1)

mru115_0.98_1.02_wo_smote$con_mat_whole
```

APPENDIX L: CLASSIFIER C3 (FULL FEATURED)

mru115_0.98_1.02_wo_smote$best.para

mru115_0.98_1.02_wo_smote$metrics

Outputs from the above R code:

```
pred1  CA  EA  LL  NC NoF  OC  UC  VL
   CA 152   0   1   1   2   0   0   0
   EA   0  69   0   0   2   0   0   0
   LL   0   0  44   0   7   0   0   0
   NC   0   0   0  68   7   0   0   0
  NoF   4   4   4   6 229  11   4   7
   OC   0   0   0   0   9  59   0   0
   UC   0   0   0   0   6   0  72   0
   VL   0   0   0   0   6   0   0  99
```

sigma <dbl>	C <dbl>
0.001	1e+06

FAR <dbl>	MR <dbl>	MDR <dbl>	accuracy <dbl>	bal_acc <dbl>
0.1455224	0.003539823	0.0661157	0.9072165	0.912857

Figure 44: Outputs from classifier C3

APPENDIX M: TESTING C2 WITH LAB DATA

R code:

Create a function to preprocess 'data2' and predict from C2

fun84 <- function(data2, preproc, model){

specific_numbers <- data.frame(Refrigerant = c("R410a", "R407c", "R22"), power_NoF = c(2720.35, 5639.18, 3106.93), T_dischg_NoF = c(178.92, 193.26, 204.10))

library(dplyr)

Merge the specific_numbers dataframe with your_data

data2 <- left_join(data2, specific_numbers, by = "Refrigerant") #after this line, power_NoF and T_dischg_NoF columns will be created

Create the new column by dividing the power column by the corresponding divisor

data2 <- data2 %>%

mutate(FIR_power = Power_comp / power_NoF) #after this line, FIR_power column will be created

data2 <- data2 %>%

 mutate(FIR_T_dischg = T_dischg / T_dischg_NoF) #after this line, FIR_T_dischg column will be created

#convert Refrigerant column into factor #seems unnecessary

APPENDIX M: TESTING C2 WITH LAB DATA

```
data2$Refrigerant <- factor(data2$Refrigerant, levels = c("R410a", "R407c", "R22"))

library(caret)

data.lab.8 <- data2[,c(1,2,3,5,7,9,24,23,18,19)]

data.lab.8$Fault <- factor(data.lab.8$Fault, levels = c("CA", "EA",

"LL","NC","NoF","OC","UC","VL"))

data.lab.8.scaled <- predict(preproc, data.lab.8)

library(e1071)

#predict the lab data

pred1 = predict(model, newdata = data.lab.8.scaled[,-10])  #predict whole lab data

#confusion matrix

conf.matx1 = table(pred1, data.lab.8.scaled[,10]) #whole lab data

acc1 = 1-mean(pred1 !=data.lab.8.scaled[,10])  #whole lab data

metrics=perf.metrics1(pred1, data.lab.8.scaled[,10])

newList <- list("data2" = data2,"lab.8" = data.lab.8, "lab.8.scaled" = data.lab.8.scaled,

"model"= model,"pred" = pred1, "con_mat_whole"=conf.matx1,"acc_whole"=acc1,

"metrics"=metrics)

 return(newList)

}
```

APPENDIX M: TESTING C2 WITH LAB DATA

#Inputting raw lab data ('dat.lab.updated2'), preprocessing info, and model C2 to 'fun84'

mru105_0.98_1.02_wo_smote=fun84(dat.lab.updated2,

mru104_0.98_1.02_wo_smote$preProc, mru104_0.98_1.02_wo_smote$model)

mru105_0.98_1.02_wo_smote$con_mat_whole

mru105_0.98_1.02_wo_smote$metrics

Outputs from the above R code:

```
pred1 CA EA LL NC NoF OC UC VL
   CA 25  0  0  0   0  0  0  0
   EA  0 27  0  0   0  2  0  0
   LL  0  0 22  0   0  0  0  0
   NC  7  0  0  0   0  0  1  0
  NoF 17 34 12  0  70  6 24  3
   OC  0  0  0  0   0  8  0  0
   UC  1  0  0  0   2  1 49  0
   VL  0  5  0  0   8  7  0 30
```

FAR <dbl>	MR <dbl>	MDR <dbl>	accuracy <dbl>	bal_acc <dbl>
0.125	0.1297297	0.341637	0.6398892	0.6193909

Figure 45: Outputs of C2 while tested with lab data

APPENDIX N: TESTING C2 WITH FIELD DATA

R code:

```
# Create a function to preprocess 'data2' and predict from C2

fun85 <- function(data2, preproc, model){

data2$Refrigerant="R22" #total 11 columns

data2 = data2[,c(1:9,11,10)] #rearrange columns

library(dplyr)

# Create the new column by dividing the power column by the corresponding divisor

data2 <- data2 %>%

  mutate(FIR_power = Power_comp / 11642.65) #after this line, FIR_power (12th col)
column will be created

data2 <- data2 %>%

  mutate(FIR_T_dischg = T_dischg / 204.10) #after this line, FIR_T_dischg (13th col)
column will be created

data2 <- data2[,c(2:7,13,12,10,11)] #rearrange those 10 columns

data.lab.8 <- data2

#convert Refrigerant column into factor
```

APPENDIX N: TESTING C2 WITH FIELD DATA

```
data.lab.8$Refrigerant <- factor(data.lab.8$Refrigerant, levels = c("R410a", "R407c",

"R22"))

data.lab.8$Fault <- factor(data.lab.8$Fault, levels = c("CA", "EA",

"LL","NC","NoF","OC","UC","VL"))

data.lab.8.scaled <- predict(preproc, data.lab.8)

# predict using lab data

pred1 = predict(model, newdata = data.lab.8.scaled[,-10])

#confusion matrix

conf.matx1 = table(pred1, data.lab.8.scaled[,10]) #whole lab data

acc1 = 1-mean(pred1 !=data.lab.8.scaled[,10])  #whole lab data

metrics=perf.metrics1(pred1, data.lab.8.scaled[,10])

newList <- list("data2" = data2, "lab.8" = data.lab.8, "lab.8.scaled" = data.lab.8.scaled,

"model"= model, "pred"

=pred1,"con_mat_whole"=conf.matx1,"acc_whole"=acc1,"metrics"=metrics)

return(newList)

}

#Inputting field data ('data_field_UC_5min_SS'), preprocessing info, and model C2 to

'fun84'
```

APPENDIX N: TESTING C2 WITH FIELD DATA

mru106_0.98_1.02_wo_smote=fun85(data_field_UC_5min_SS,

mru104_0.98_1.02_wo_smote$preProc, mru104_0.98_1.02_wo_smote$model)

mru106_0.98_1.02_wo_smote$con_mat_whole

mru106_0.98_1.02_wo_smote$metrics

Outputs from the above R code:

```
pred1 CA EA LL NC NoF OC UC VL
   CA  0  0  0  0   0  0  0  0
   EA  0  0  0  0   0  0  0  0
   LL  0  0  0  0   0  0  0  0
   NC  0  0  0  0   0  0  0  0
  NoF  0  0  0  0   0  0  0  0
   OC  0  0  0  0   0  0  0  0
   UC  0  0  0  0   0  0 24  0
   VL  0  0  0  0   0  0  0  0
```

FAR <dbl>	MR <dbl>	MDR <dbl>	accuracy <dbl>	bal_acc <dbl>
NA	0	0	1	1

Figure 46: Outputs of C2 while tested with field data