A UNIFIED META BASED MACHINE LEARNING MODEL FOR SUSTAINABLE MANUFACTURING USING CHARACTERIZATION AND REGRESSION OF MACHINING DATA

SANGEETHA ELANGO

DOCTOR OF PHILOSOPHY (ENGINEERING)

LEE KONG CHIAN FACULTY OF ENGINEERING AND SCIENCE

UNIVERSITI TUNKU ABDUL RAHMAN

MARCH 2023

LEE KONG CHIAN FACULTY OF ENGINEERING & SCIENCE

UNIVERSITI TUNKU ABDUL RAHMAN

Date: 20/03/2023

SUBMISSION OF THESIS

It is hereby certified that SANGEETHA ELANGO (ID No:**19UED05556**) has completed this thesis entitled "**A UNIFIED META BASED MACHINE LEARNING MODEL FOR SUSTAINABLE MANUFACTURING USING CHARACTERIZATION AND REGRESSION OF MACHINING DATA**" under the supervision of Dr Ezra Morris Abraham Gnanamuthu (Supervisor) from the Department of Electrical and Electronic Engineering, Lee Kong Chian Faculty of Engineering and Science and Ir. Prof. Dr. Rajkumar a/l Durairaj (Co-Supervisor) from the Department of Mechanical and Material Engineering, Lee Kong Chian Faculty of Engineering and Science.

I understand that the University will upload softcopy of my thesis in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,

(Sangeetha Elango)

DECLARATION

I **SANGEETHA ELANGO** hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTAR or other institutions.

(SANGEETHA ELANGO)

Date: 20/03/2023

A UNIFIED META BASED MACHINE LEARNING MODEL FOR SUSTAINABLE MANUFACTURING USING CHARACTERIZATION AND REGRESSION OF MACHINING DATA

By

SANGEETHA ELANGO

A thesis submitted to the Department of Lee Kong Chian Faculty of
Engineering & Science,
Universiti Tunku Abdul Rahman,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy (Engineering) March 2023

ABSTRACT

A UNIFIED META BASED MACHINE LEARNING MODEL FOR SUSTAINABLE MANUFACTURING USING CHARACTERIZATION AND REGRESSION OF MACHINING DATA

Sangeetha Elango

Polyoxymethylene(POM), Polytetrafluoroethylene(PTFE), Polyether ether ketone(PEEK) and multiwall carbon nano tubes reinforced PEEK(PEEK/MWCNT) are significant polymeric materials used in industrial applications and house-hold items and hence they were considered for this research. Considering its potentiality, this research was attempted to investigate and develop a unified meta based machine learning model for turning of different polymeric materials. Developed meta based model has two main parts: classifier and regressor. Classifier is the one initially taking the experimental data and classify them into reasonable and more accurate groups. Regressor use the output from the classifier and predict the surface finish. For classification, XGBoost algorithm and Logistic regression algorithm were investigated. k-fold cross validation method was adopted to apply all data patterns in learning of the model. Grid searching method was used to tune the hyper parameters for each algorithm. It was found from these results that Logistic Regression model is the better to be used as classifier. Once classifier model was confirmed, the output of the classifier was added to the database as a new feature. Now, with four independent features (including output of classifier), Support vector regressor and XGBoost algorithm were used to complete meta based model for each material. Further, a unified model (a model for all four polymeric materials) was developed using the same procedure as discussed above. It used an additional input feature known as material number. Interestingly observed that XGB model is the best model working great in both classification and regression. It resulted almost 100% accuracy in training and 98.86% in testing. After confirming the best model, a group of predicted results was generated from the prediction model and validated experimentally. Finally, user interface (API) was developed for the unified meta based model which industry can use in its production line for achieving high productivity.

ACKNOWLEDGEMENT

My highest gratitude to my supervisor Dr Ezra Morris Abraham Gnanamuthu from the Department of Electrical and Electronic Engineering, Lee Kong Chian Faculty of Engineering and Science and my co-supervisor Ir. Prof. Dr. Rajkumar a/l Durairaj from the Department of Mechanical and Material Engineering, Lee Kong Chian Faculty of Engineering and Science for their unconditional guidance and teaching throughout my doctoral study. It is not possible to bring out such an excellent research output without their involvement, professionalism, and preciseness. I thank God almighty for being with me in each and every step and lead me to the greatness and success. I thank my family members; my spouse Elango Natarajan, daughter Nikitha Mira and son Guru Kumaraa for their patience and cooperation during my study and fully involved research activities. And lastly, I thank my cousin brother Karthikeyan Mohanraj who assisted me in trouble shooting the models.

APPROVAL SHEET

This thesis entitled "**A UNIFIED META BASED MACHINE LEARNING MODEL FOR SUSTAINABLE MANUFACTURING USING CHARACTERIZATION AND REGRESSION OF MACHINING DATA**" was prepared by SANGEETHA ELANGO and submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering at Universiti Tunku AbdulRahman.

Approved by:

____

(Dr Ezra Morris Abraham Gnanamuthu)
Date:20/03/2023
Supervisor
Department of Electrical and Electronic Engineering,
Lee Kong Chian Faculty of Engineering and Science,
Universiti Tunku Abdul Rahman

____

(Ir. Prof. Dr. Rajkumar a/l Durairaj)
Date:20/03/2023
Co-supervisor
Department of Mechanical and Material Engineering,
Lee Kong Chian Faculty of Engineering and Science,
Universiti Tunku Abdul Rahman

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

AA	Aluminium Alloy
AE	Acoustic Emission
AI	Artificial Intelligence
AISI	Austenitic Stainless Steel
ANOM	Analysis of Means
ANOVA	Analysis of Variance
ANN	Artificial Neural network
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANFIS	Adaptive Neuro-Fuzzy Inference System
API	Application Program Interface
BNN	Bayesian Neural Network
BpNN	Backpropagation Neural Network Algorithm
CAT	Cat Boost Regression
CNN	Convolutional Neural Network
CNC	Computer Numerical Controlled
CPI	Continuous Process Improvement
CPS	Cyber Physical System
DF	Desirability Function
DFA	Design for Assembly
DFM	Design for Manufacturing
DOE	Design of Experiments
DO	Dandelion Optimizer
DT	Decision Tree
DTR	Decision Tree Regression
EMOTLBO	Enhanced Multi-Objective Teaching Learning-Based Optimization

EM-PCA	Expectation-Maximization for Principal Component Analysis
ETR	Extra Tree Regressor
FEA	Finite Element Analysis
FVM	Finite Volume Method
FFT-DNN	Fourier Transform-Deep Neural Networks
FFT-LSTM	Fast Fourier Transform Long Short-Term Memory Network
GBR	Gradient Boosting Regression
GP	Genetic Programming
H-ABC	Harmonic Artificial Bee Colony Algorithm
HDPE	High-Density Polyethylene
HNT/AI/Ep	Aluminium Reinforced Epoxy Matrix Hybrid Composite
IOT	Internet of Things
IOS	Internet of Services
ISOMAP	Isometric Feature Mapping
IT	Information Technology
ITD	Intrinsic Timescale Decomposition
KELM	Kernel Extreme Learning Machine
LSTM	Long Short-Term Memory Network
ML	Machine Learning
MLP	Multi-Layer Perception
MSE	Mean Squared Error
NB	Naïve Bayes
NSMTLBO	Non-dominated Sorting Modified Teaching-Learning Based Optimization
NSGA	Non-Sorting Genetic Algorithm
NSGA-II	Non-dominated sorting genetic algorithm-II

PDM	Product Data Management
PE	Polyethylene
PEEK	Polyether Ether Ketone
PEEK/MWCNT	Reinforced PEEK
PLM	Product Life Cycle Management
POM	Polyoxymethlene
PP	Polypropylene
PR	Proper Rotation
PTFE	Polytetrafluoroethylene
PU	Polyurethane
QPSO	Quantum Particle Swarm Optimization Algorithm
RF	Random Forest
RA	Regression Analysis
RBF	Radial Basis Function
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
RPM	Revolutions Per Minute
RSM	Response Surface Methodology / Response Surface Model
RUL	Remaining Useful Life
SDG	Sustainable Development Goals
SG	Stochastic Gradient
SME	Small and Medical Scale Industries
SVM	Support Vector Machine
SVM-RFE	Support Vector Machine Recursive Feature Elimination
SVR	Support Vector Regressor
TLBO	Teaching-Learning Based Optimization

UI	User Interface
WAAM	Wire Arc Additive Manufacturing
XGB	Extreme Gradient Boosting Regression
XGBoost	Extreme Boosting

TABLE OF CONTENTS

CHAPTER 1.0

INTRODUCTION

This research is intended to develop a unified meta based machine learning (ML) model which can predict surface roughness of the part (outcome response) during turning of polymeric materials. It takes speed, feed, depth of cut and name of the polymeric material as input features and predict the surface roughness corresponding to the input parameters.

1.1 Background

Production can be broadly classified into intermittent production and continuous production. Mass production is a subset of continuous production system where a large quantity of parts against a particular design is produced. Automated production or numerical controlled production can be easily implemented in mass production system, which can use the programmed instructions to manufacture parts. On the other hand, job shop production is a subset of intermittent production system where parts are produced to the requirement of customers and hence only one or some quantities are produced against a particular design. In mass production, the same machine setting can be utilized continuously to produce many quantities of items. But in job shop production, we change the machine setting once one or some items are produced.

Product development involves many tasks including market survey, materials selection, engineering design, engineering analysis, tools selection, process planning, production, assembly etc. The engineering industries apply many methods, strategies, and tools for accomplishing each of these tasks effectively. The effectiveness of the process is measured in each stage in order to ensure the productivity and profit. For instance, design of experiments (DOE), finite element analysis (FEA), finite volume method (FVM), design for manufacturing (DFM), design for assembly (DFA) are some of the tools used in conducting engineering design and manufacturing. The invention of high-end computers, technologies and networks have made a lead through pathway for the industries to apply information technology (IT) in each of its operation. The technologies like product data management (PDM) and product life cycle management (PLM) have been adopted by many industries since last two decades. PDM and PLM are frontiers of Industry 4.0 where product information is managed and controlled through networks.

Industry 4.0 is a holistic approach of utilizing internet of things (IOT) and internet of services (IOS) where resources are connected together in cyber physical system (CPS). Industry 4.0 has been adopted in developed countries and some developing countries in the recent years. Having the components such as Cyber-Physical Systems (IoT and IoS), Augmented reality, Virtual reality, Autonomous robots, Cloud computing, the Industry 4.0 perfectly conducts digital manufacturing for the satisfaction of the customers or clients.

The practice of Artificial Intelligence (AI) and Machine Learning (ML) ranges from big giant applications to the tiniest deployment of the technologies. Both are commonly practiced in the social media networks, business operations, and marketing. Their contribution in manufacturing sectors is however limited for various reasons. They may be constrained by budget, support from stockholders including the Government, limited resources, and facilities etc., in the worst-to-worst case, because of the unawareness of their usability, credibility and benefits.

Developed nations have started implementing Industry 4.0 in their production as attributed by intelligence and smartness in manufacturing. The automation and digitisation are two key elements in any modern industry. Automation of manufacturing process includes autonomous material handling, job loading/unloading and numerical controlled machining etc. Industry 4.0 technology based on cyber physical system (CPS) is clustered as below.

(a) Acquisition of data through sensors and processing them for decision making.
(b) Machine-to-machine (M2M) communication
(c) Human-machine interface (HMI) and communication

These tasks are achieved by connecting all resources in the factory through different modules and sensors, which are controlled and managed by group of programs and AI technologies. Tobias et al (2017) detailed how industry 4.0 impacts lean manufacturing system. Tommaso Harald Bauer et al (2018) presented challenges in implementing industry 4.0 and how new technological philosophy can be implemented in lean manufacturing. Sri Kolla et al (2019) presented challenges in small and medical scale industries (SMEs) and discussed how hybrid model consisting of lean manufacturing and industry 4.0 technologies will help SMEs, while readiness and maturity model was presented by Andreas Schumacher et al (2016). Ortt et al (2020) reviewed research articles published in this area in a last decade and discussed the implementation method. Gallo et al (2021) conducted a systematic review on tools for implementing industry 4.0 in lean manufacturing system.

The machining is considered to be one of the key elements in producing parts to the near-net shape. Almost all the industrial products take the machining operations in some form, irrespective of the manufacturing process it had taken. For instance, insulating socket of some plastic can be prepared from moulding, but it may further need some hole to use cables or screws. This considers to be the task of manufacturing the industrial product to the near net shape.

Drilling, milling, turning, grinding, and tapping are some machining processes that companies employ to produce their products or parts. The conventional method of machining has many limitations in terms of machine selection, cutter selection, cutting parameters setting as they are performed prescriptively. Apply artificial intelligence (AI) in machining of parts would help to identify the process outcome beforehand. The implicit benefits of AI implementation in machining processes are reduced wastage and increased profit.

1.2 Problem Statement

One of the challenges the industries face in the production line is that rejected parts due to defying with the design or standards. Production cost is increased when there are many rejected parts. The manufacturing process can be time-consuming and expensive for companies that do not adopt strategies or tools for the production. The solution to aforementioned problem is that the use of prediction model in the production. The application of appropriate ML model in the prediction of quality of the products helps in avoiding wastages. The less wastage in production leads the company to achieve high profit and sustainable manufacturing.

1.3 Research Objectives

In order to overcome the wastages in machining, this research is attempted to develop a unified meta based ML model. The following are four research objectives that serve the solution to the above problem.

- To prepare design of experiments (DoE) for different materials (PTFE, PEEK, DELRIN and PEEK/MWCNT) and conduct machining operation (Turning) according to DoE.
- To analyse the experimental data and develop a unified meta based machine learning (ML) model
- To investigate the effectiveness of the machine learning model suitable for industrial application.
- To deploy and demonstrate the developed model using application program interface (API).

1.4 Research Contribution

This research used four significant polymeric materials (PTFE, PEEK, DELRIN and PEEK/MWCNT) that are predominantly used in bio medical applications, industries and household items. A new model development approach was introduced in this research which is called as unified meta based model. This model is able to predict the surface roughness of the part before machining is taken place. Application Program Interface (API) was developed which is an interactive user interface where operator can input speed, feed, depth of cut and name of the polymeric material. It would result the surface roughness of the part if this input setting had been used.

Research contributions are listed below.

1. Systematic experimental design was conducted, and experimental data were collected from turning operations.
2. Meta based model was developed for each material and appropriate ML algorithm suitable for the machining data was evaluated.
3. Unified meta based model was developed and validated which can predict the surface roughness of the part for any unseen data.
4. API was developed for ease interaction and service to the operator who can know the machining outcome before he starts the machining.

Though only four materials were used in developing the model, many more materials can also be added into this in the future. So that our model can be a complete unified model for manufacturing company.

1.4 Thesis Outline

In Chapter 2, various research articles published with meta-heuristic algorithm in the current problem is presented, followed by detailed discussion of literature on machine learning. It also comprehends the various method, ML algorithm, materials used in the past. Lastly, the research gap is presented to emphasize the need of the current research. In Chapter 3, method of experimental design, experimentation and data collection are detailed. In Chapter 4, ML model development for individual material and the unified ML model are detailed and discussed. In Chapter 5, results are presented and discussed, while Chapter 6 presents the conclusion and further improvements.

CHAPTER 2

LITERATURE REVIEW

Comprehensive literature related to the current research is presented in this chapter. The significance of using optimization algorithms and machine learning algorithms in turning operation is detailed. The model development methods and performance of models are also discussed.

2.1 Introduction

With the introduction of Industry 4.0, which utilizes AI as one of the key components, companies can achieve the set production target with no wastage or negligible wastage. AI can be implemented in each part of production line through which production capabilities can be improved to the maximum. AI can be implemented in machining of parts as it is a significant and mostly done operation in manufacturing. Implementation of ML models, which otherwise known as prediction models is prevalent strategy, which can estimate the quality and dimensional accuracy of the part before start of the machining. In this way wastage in production can be avoided and high profit can be achieved. SDG Goal 12: Responsible Consumption and Production, which focusses on sustainable consumption and production patterns can be ensured through this kind of sustainable manufacturing.

Turning operation is a subtractive process through which materials are removed using a lathe machine. Materials are removed to manufacture a part to the required dimensional accuracy and surface characteristic. Parameters such as speed, depth of cut, feed, coolant, tool geometry are some parameters that operator considers during turning operation. Figure 2.1 illustrates the parameters used in turning operation. Cutting speed refers to the speed of a tool that cuts the workpiece and is measured in m/s. The feed rate is the distance covered by the tool in one rotation of the spindle. It is measured in rev/minute. Depth of cut is the distance the tool is pushed deeper into the workpiece. It is measured in mm, and it may generally vary from 0.1 to 1.0 mm.

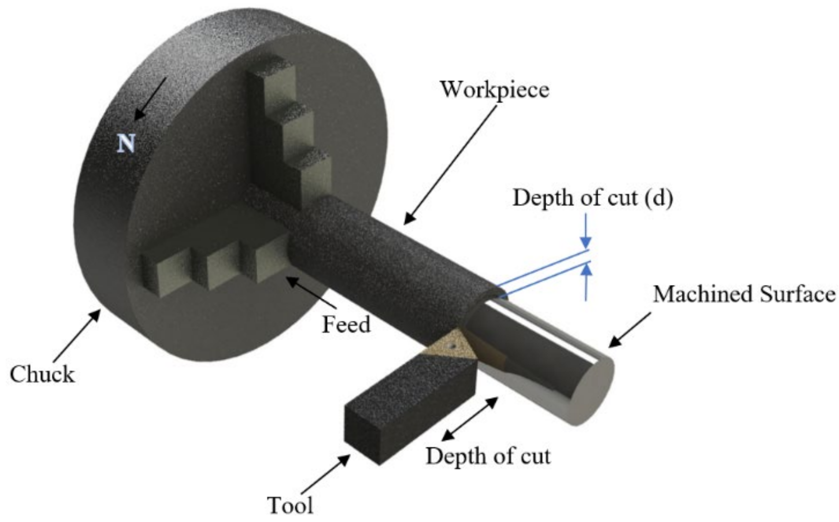

Figure 2.1: Illustration of cutting parameters and tool geometry

Some of the observed points are: (a) higher speed would result poor surface finish. (b) higher feed rate would increase the material removal rate but leave waviness on the surface (poor surface finish). When feed rate is increased, the cutting temperature and tool flank wear would also increase and hence the tool life would be affected. The possibility of recutting chips is higher when too lower feed rate is used. Meanwhile, too fast of a feed rate would cause a tool fracture. Feed rate and cutting speed are generally fixed based on the nature of the material to be cut. In addition to these two parameters, depth of cut, rigidity of the lathe, coolant and its condition must also be considered for obtaining higher surface finish. In general, industries do not follow any standards or templates in parameters setting during machining operations. These parameters are fixed based on the experience.

2.2 Research on Optimization of Cutting Parameters

The effect of cutting parameters have been examined, analysed by many researchers in last three decades. Their research falls on one or many below areas: materials to be cut, cutting tool material, coolant conditions (no coolant, minimal coolant, full coolant), bio coolant, cutting angles. Significant research was conducted on optimization of machining parameters in last two decades. It could be single objective or multi objective optimization problem. The focus of these research articles was on utilizing optimization algorithms to reduce surface roughness and/or material removal rate or tool wear.

Kaddeche et al (2012) used Response Surface Model (RSM) to achieve low surface roughness from turning of High-Density Polyethylene HDPE80 and HDPE100 polymers. Lazarevića et al (2012) used ANOM and ANOVA to analyse the cutting parameters that affect the surface roughness and material removal rate in turning of Polyethylene. Pang et al (2014) used Taguchi method to analyse end milling parameters for halloysite nanotube made of Aluminium reinforced epoxy matrix (HNT/Al/Ep) hybrid composite. Panda et al (2016) used ANOVA to analyse the cutting parameters that affect the surface roughness and material removal rate in turning of Nylon 6/6. Abdul Shukor et al (2016) used Taguchi method for design of experiments and signal-to-noise ratio to analyse the parameters in milling of Polypropylene (PP). . Hamlaoui et al (2017) used RSM, ANOVA and Desirability function (DF) to investigate the machinability of High Density Polyethelene (HDPE) resin. Chabbi et al (2017) used RSM to study the effect of turning parameters on surface roughness and material removal rate of Polyoxymethlene(POM).

Natarajan et al (2017) conducted turning experiments on ACETAL homopolymer (Delrin) material and developed RSM model to optimize the turning parameters. They further developed Enhanced Multi-Objective Teaching Learning-Based Optimization (EMOTLBO) algorithm and used it with Fuzzy decision maker for optimization. Kaviarasan et al (2018) conducted drilling experiments on the same material (DELRIN) and used RSM and Artificial Neural network (ANN) to optimize drilling parameters. Natarajan et al (2019) in their another research used RSM and non-dominated sorting modified teaching-learning based optimization (NSMTLBO) algorithm for optimizing machining parameters such as cutting speed, feed rate, depth of cut in turning of Polytetrafluoroethylene (PTFE). Jia et al (2021) used Non-dominated sorting genetic algorithm-II (NSGA-II) to optimize the turning parameters to achieve low energy consumption in turning of Steel. Kuntoğlu et al (2021) applied Harmonic Artificial Bee Colony Algorithm (H-ABC) for optimization of parameters in turning of AISI 5140 steel.

Palanikumar et al (2021) used Fuzzy and NSGA algorithm to optimize the thrust force in drilling of AA6061 aluminium alloy. Palanikumar et al (2022a) in their another research used a TLBO variant to optimize the cutting force and surface roughness in turning of Titanium alloy. Palanikumar et al (2022b) analysed cutting temperature in turning of Ti-6Al-4 V titanium alloy and used RSM model for modelling. Elango et al (2022a) used RSM and Adaptive Neuro-Fuzzy Inference System (ANFIS) to analyse and optimize the drilling parameters in drilling of Polytetrafluoroethylene (PTFE). Recently, Elango et al (2022b) used Dandelion optimizer (DO) to optimize the delamination in drilling of S-glass/Carbon fiber reinforced polymer composite. Koon Meng Ang et al (2022) considered five case studies relating to multi response machining processes and discussed how TLBO algorithm can be modified to achieve the best efficacy.

The observations from the above literature are:

(a) Substantial research on finding optimal machining parameters for turning, drilling or milling of different materials were published in last two to three decades. Different optimization strategies were adopted for different machining operations.

(b) Most of research was focused on metals, a very minimal literature are available on plastic materials.

(c) Only minimal research has been conducted on applying ML algorithms and development of prediction model for such operations. Search of optimal solution always results only the best parameters to be used in the production, while prediction model can help in getting the best parameters for any unseen response variable.

2.3 Research on AI Implementation in Manufacturing

Many manufacturing industries throughout the world focus currently on continuous process improvement (CPI) to maintain and increase their competitive level. The objective of CPI is to improve the quality of goods and services at all levels of the manufacturing operation. Smart manufacturing is the application of advanced technologies that enable the stable manufacturing of new products in less time (Baicum et al, 2021).

Ruholla Jafari-Marandi et al (2019) used thermal data of laser based additive manufacturing captured by IR camera, and microstructure of the post-production data from X-ray CT to develop multilayer perception (MLP) model. This is a deep learning model that can be applied in online quality control decision making. Gülçür and Whiteside (2021) developed a multilinear regression model to evaluate the quality of fingerprints in the micro injection moulding. Kai Kai Guo et al (2021) ML developed ML models trained from large material datasets that relate structure, properties and function of materials that can assist in material design. Ahmed Yaseer and Heping Chen (2021) developed a MLP model Wire to predict the layer surface roughness in Arc Additive Manufacturing (WAAM). Though the development of prediction models was started in last decade, plentiful research articles are available for the reference. Table 2.2 shows some of the AI research on machining being conducted and published in recent years. Reviews pinning ML algorithms pertaining to machining are available in (Dong-Hyeon, 2018; Kim et al, 2018; Nasir&Sassani, 2021; Aggogeri et al, 2021). Dong-Hyeon Kim (2018) presented use cases in milling, turning and drilling operations. They addressed that algorithm such as support vector machine (SVM), artificial neural network (ANN) and decision trees, probabilistic neural network, backpropagation neural network algorithm (BpNN) and random forest (RF) were commonly used. Particularly for the classification concerning tool wear monitoring, neural network algorithms were used. Among the algorithms, SVM was popularly acknowledged through the evaluation.

For the machine learning model development, most of the research was attempted on metals, particularly on titanium and aluminium. The reason is that titanium is one of hard metals to machine. Hence, most of the researchers attempted to investigate the effect of machining (particularly milling operation) on tool wear. As tool wear is directly connected to the surface quality and production down, online tool monitoring was examined using prediction models. Only a little amount of research was carried out on other parameters. In terms of ML algorithms, supervised algorithms were used. For online tool monitoring, mostly neural network or classification algorithms were used. For prediction of other parameter like surface roughness, regressor was used. Comprehensive literature review on ML models in machining is presented in the following section.

2.4 Research on ML Algorithms in Machining

Lela et al (2008) examined the influence of cutting speed, feed, and depth of cut on surface roughness and reported in their article that regression analysis (RA) is better than Support Vector Machines (SVM), and Bayesian Neural Network (BNN) for face milling of steel. Amit Kumar Gupta (2010) conducted turning operations on A356/20/SiCp-T6 metal matrix composite and developed a prediction model for surface roughness, tool wear and power required in turning operation. They used response surface methodology (RSM) for modelling, artificial neural network (ANN) and support vector regressor (SVR) algorithms for prediction model development.

Çaydaş et al (2012) used three different SVMs models such as least square SVM (LS-SVM), Spider SVM and SVM-KM and ANN models to predict the surface roughness on AISI304 Austenitic Stainless Steel in turning operation. They concluded that SVM models are better than ANN. Benkedjouh et al (2015) used non-linear reduction techniques; EM-PCA (Expectation-Maximization for Principal Component Analysis) and ISOMAP (isometric feature mapping) and SVR to predict tool wear assessment and remaining useful life (RUL) on milling of Steel material. Krishnakumar et al (2015) used Decision tree (J48) algorithm for feature selection and J48 and ANN for classification of cutting tool wear in highspeed turning of Titanium alloy (Ti-6Al-4 V). They concluded that ANN produces the better results compared to J48 algorithm. Gupta et al (2015) used SVM and ANN integrated with GA to determine the tool wear, power required and surface roughness in turning operation of Steel. Yiğit M. Arisoy & Tuğrul Özel (2015) investigated different tools and conditions to collect the characteristics of turned samples of Ti–6Al–4V alloy material. They further developed a Random Forest model to estimate the hardness and microstructure of the parts.

Lu et al (2016) used SVM to develop a prediction model for micro milling of Inconel 718 Steel. Sara Karam et al (2016) developed neural networks to monitor online tool life during turning of AISI 316 stainless steel. Žarko Ćojbašić et al (2016) used Extreme Booster algorithm to detect the surface roughness of Aluminium alloy AAASTM 6060 (EN: AW-6060; ISO: Al MgSi0.5) parts from abrasive water jet machining. They compared the results with genetic programming (GP) and artificial neural networks (ANNs) models and reported that extreme learning is better. Kilickap et al (2017) developed ANN model for cutting force, tool wear and surface roughness during milling operations of Ti-6242S alloy. Zurkovic et al (2018) used SVR, polynomial (quadratic) regression, and artificial neural network (ANN) for the prediction of independent cutting parameters in a high-speed turning of Steel. Duo et al (2018) considered process variables and statistical characteristics obtained from sensor signals for online drilling tool monitoring in drilling of 35CrMo4 steel. They used decision tree algorithms and evaluated the performance. Tran et al (2019) used continuous wavelet transform data as inputs to deep convolutional neural network (CNN) model to detect stable, transitive, and unstable cutting in milling of Al 6061-T6 alloy. Liu et al (2019) investigated Ring-shaped thin-walled metal discs and addressed that surface roughness on a circumference trajectory is not the same at all places, because of tool vibration. Hence, they considered cutting parameters and tool parameters in their machine learning model and reported that Gaussian-process-based Bayesian combined model is the best model for turning of ring-shaped thin-walled metal discs. Hui et al (2019) used milling tool vibration signals from milling of 45 heat-treatable steel as features for predicting tool wear. Support vector machine recursive feature elimination (SVM-RFE) algorithm was used to select the main features that are most relevant to tool wear states. SVM, decision tree (DT), naïve Bayes (NB) and Stochastic Gradient (SG) ensemble strategy were used and concluded that SG ensemble model has better recognition accuracy and stability than other models. Cherukuri et al (2019) used ANN to assess the vibration chatter in turning of Steel. Peng Wang et al

(2019) combinedly used Convolutional Neural Networks (CNN) and Recurrent Neural Network (RNN) for milling tool monitoring. Surface and wear images of H13 steel and Inconel 718 milled samples were analysed using CNN to detect the surface roughness and wear impacts. The output of CNN was subsequently fed into a RNN to assess the relationship between degradation of tool and power. Xin-Cheng Cao et al (2019) used machine spindle vibration signals in the wavelet forms for the development of CNN model to monitor the tool wear in milling of S45C steel. Lu et al (2019) used wireless sensory tool holder and measured time domain, frequency domain, time-frequency domain features to predict surface roughness of milled Aluminum alloy 7109. Figure 2.2 depicts the way the features were acquired from sensors.

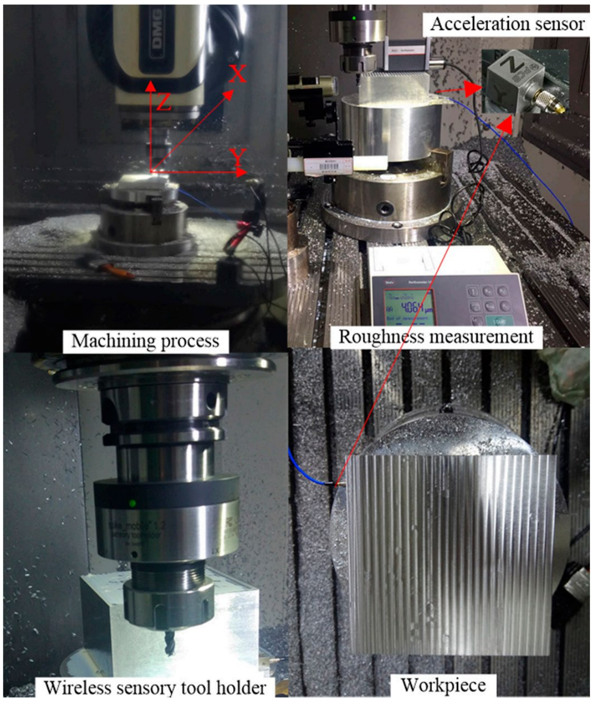

Figure 2.2: Illustration of measuring features from a milling operation (Lu et al, 2019)

Lin et al (2019) acquired features through vibration sensors and developed Fast Fourier Transform-Deep Neural Networks (FFT-DNN), Fast Fourier

Transform Long Short-Term Memory Network (FFT-LSTM), and one-dimensional convolutional neural network (1-D CNN) models for milling of S45C medium carbon steel. They concluded that FFT-LSTM or 1-D CNN is better for surface roughness prediction. Grzenda et al (2019) combined k-Nearest Neighbours algorithm and random forest technique to build the prediction model for predicting surface roughness of milled F114 steel.

Lei et al (2020) used an intrinsic timescale decomposition (ITD) technique to decompose sensor signals from different sources into several sets of proper rotation (PR) and combine them a kernel extreme learning machine (KELM) for developing a machine learning model for milling of 45 grade steel. Yuqing Zhou et al (2020) also developed a machine learning model for milling of 45 grade steel, but their approach is different from Lei et al (2020). They used many domains (time, frequency, wavelet) to compose feature parameters of tool condition and used BDE optimization algorithm to get optimal fewest feature parameters. They further used two-layer angle kernel extreme learning machine for the prediction. Laddada et al (2020) developed a ML model for Cast Iron and steel milling. The developed model monitored the tool condition for which it used complex continuous wavelet transform data as input features and extreme learning machine for prediction. Sebastian Schorr et al (2020) developed Extra Tree Regressor (ETR) model for in process quality control during drilling and reaming of valves. This model is to predict the hole concentricity and diameter based on torque input. Alajmi et al (2020) used Adaptive Neuro-Fuzzy Inference System (ANFIS) and The Quantum Particle Swarm Optimization Algorithm (QPSO) for modelling surface roughness in cryogenic turned AISI 304 stainless steel. Vuong et al (2020) attempted linear regression, quadratic regression, random forest, and Gradient Boosting algorithms to develop models for surface roughness and tool flank wear predictions in turned steel material.

Karthik et al (2021) investigated cryogenic milling of SS316 material using a hybrid bias SVR algorithm to predict the surface finish of parts. Van-Hai Nguyen et al (2021) used ANN, Cat Boost Regression (CAT), SVR, Gradient Boosting Regression (GBR), Decision Tree Regression (DTR) and Extreme Gradient Boosting Regression (XGB) to develop a prediction model for milling of Polycarbonate materials. Markus Brillinger et al (2021) used Decision tree algorithms such as Decision tree, Random Forest, and Boosted Random Forest algorithms for energy demand in CNC turning of Aluminuium alloy. Bustillo et al (2021) used Random Forest ensembles combined with Synthetic Minority Over-sampling Technique (SMOTE) balancing technique to develop a model for predicting flatness deviation in face milled AISI 1045 steel parts. Paweł Twardowski et al (2021) also used acoustic emission (AE) signals for milling tool monitoring in milling of Aluminum-SiC metal matrix composite. They k-nearest algorithm for model development. Tabaszewski et al (2022) developed a prediction model for turning of Gray cast-iron EN-GJL-250 material. They used Classification Tree CART, Induced Fuzzy Rules and Artificial Neural Network algorithms to monitor the tool performance from vibration acceleration signals. Shah et al (2022) used acoustic emission (AE) and vibration signals captured through sensors during face milling of Stainless Steel. They developed three long short-term memory network (LSTM) models: vanilla, stacked, and bidirectional and analysed the tool condition during face milling. This approach is deep learning approach. Sangeetha et al (2022a) investigated SVR and Polynomial regressors models suitable for turning of Polytetrafluoroethylene (PTFE) material and reported that SVR better model than polynomial model. Pan et al (2022) used Laser Doppler based single point laser vibrometer to acquire features in milling of Tungsten heavy alloy. They discretized surface roughness and fitted into classification problem. 6-layer convolution, 6-layer pooling, and 3-layer fully connected deep CNN models were developed. They used dropout method to prevent the overfitting in deep structure of CNN model and batch normalization method to re-parameterizing the CNN model.

Mirifar et al (2020) use feedforward neural network with Bayesian backpropagation to predict surface roughness in grinding of 100Cr6 steel. They used features extracted from acoustic emission signals. Sangeetha et al (2022b) recently showed the efficacy of Extreme boosting (XGBoost) model for drilling of Polyether ether ketone (PEEK) material.

2.5 Research Gap

Though AI and machine learning algorithms have been developed and implemented in recent decades, many different sectors have not started to utilize them for different applications. One of the key sectors where these technologies must be implemented is manufacturing industries. Because this sector is a backbone of the economy of the Nation. If AI is rightly applied in this sector, productivity can be greatly improved, and profit can be increased. Though academia has started doing research on implementation of AI methods in manufacturing, industrial involvement is still warranted. Particularly, an application that helps in finding the quality of the product beforehand is acknowledgeable for any industry. On reviewing the literature published in last decade, their focus was to implement online tool monitoring only. In most articles, they used sensors to collect the tool characteristics and use them in ML model training. Only a few articles have been published on surface quality.

Besides, polymeric materials are now seen as a vibrant alternative to metals because of its high weight-strength ratio. In last two decades, fiber reinforced materials have been adopted in transport, aerospace, space vehicles, biomedical and many more industrial applications. Particularly, glass fiber reinforced composites, carbon fiber reinforced composites have been widely used in these applications. It is understood from the literature review that the AI implementation in machining of polymeric materials is not found except in (Sangeetha et al, 2022a; Sangeetha et al 2022b; Van-Hai Nguyen et al, 2021). Considering its potentiality, this research is attempted to develop a unified meta based machine learning model for turning of different polymeric materials.

CHAPTER 3.0

METHODOLOGY

The first foremost step involved in this research is to conduct turning experiments on polymeric materials and collect datasets which can further be used in model development. In this chapter, method of research pinning to experimental design, experimentation and data measuring method are presented and discussed.

The below steps are summary of our meta based ML model development.

1. Identifying the process variables involved in turning of all these selected materials.
2. Preparing design of experiments (DoE) involving each material.
3. Conducting experiments according to DoE and measuring output response (surface roughness).
4. Analyzing the experimental data, understanding the significance of independent variables and data preparation.
5. Applying classification algorithm into the experimental data, training, k-fold cross validation.
6. Testing of classifier with performance metrics and tuning hyper parameters.
7. Repeating step 5 and 6 till the best classifier is achieved.

8. Training meta based model including with the output from classifier, k-fold cross validation.
9. Testing of meta based model with performance metrics and tuning hyper parameters.
10. Repeating step 8 and 9 till the best meta based model is achieved.
11. Developing holistic meta based model with the same procedure.
12. Generating predicted data (about 20 sets) from the developed model.
13. Conducting validation experiments and identifying deviation between predicted data and experimental data.
14. Continue train the model with new datasets.
15. Performing steps 11-14 iteratively until error is negligible or attaining optimal level of threshold value (GINI value).
16. Developing API for the developed model for end user application.

Turning of cylindrical specimen is generally done in computer numerical controlled (CNC) machine. In CNC machine, the tool or cutter is fixed in tool holder, while sample is fixed into chuck. The cutting tool removes the material from the outer diameter of a rotating workpiece, when the sample rotates at some speed. For turning operation, some of the parameters such as speed, feed, depth of cut is to be programmed. Once the sample is loaded into the machine, the operation is carried out according to the program. Figure 3.1 shows the cutting parameters used in the experimentation and modelling. *D* refers to diameter of the sample, *f* and *ap* refer to the feed and depth of cut during the turning operation.

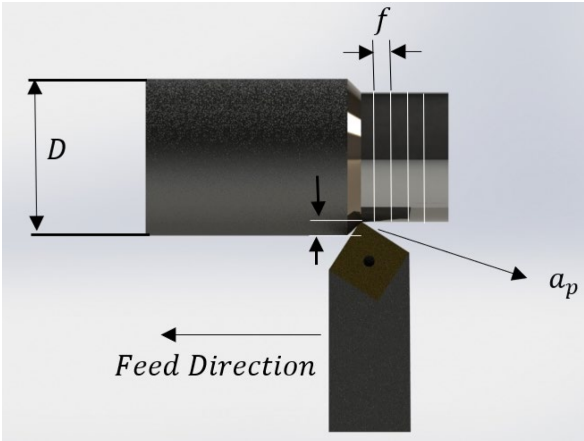

Figure 3.1: Illustration of parameters in turning operations

An overview of the cutting parameters used in turning operation is presented below.

Spindle speed (N): It refers to the speed at which the spindle and the workpiece rotates in revolutions per minute (rpm). It is equal to the ratio of the cutting speed and the circumference of the workpiece where the cut is being made.

Cutting speed or cutting velocity (V_c): It refers to tangential velocity of either the rotating workpiece or the rotating cutting tool. It is measured in m/minute.

Feed (f): It refers to the distance the cutting tool advances in one revolution of the spindle. It is measured in mm/revolution.

Depth of cut (ap): It refers to the depth of the tool advancement along the axis of the workpiece. It is measured in mm.

3.1 Materials used in Experimentation

The material chosen for the current research was Polyoxymethylene (POM), Polytetrafluoroethylene (PTFE), Polyether ether ketone (PEEK) and Multiwall carbon nano tubes reinforced PEEK (PEEK/MWCNT) composite. These materials are significant polymeric materials used in industrial applications and house-hold items. Moreover, once unified ML model is tested and validated, we may update the model with more other polymeric materials in the future. The following section gives an overview of each material.

Polyoxymethylene (POM) is a crystalline thermoplastic polymer material and is a version of Acetal homopolymer that offers outstanding mechanical properties, wear resistance and environmental properties. Delrin is a specific name used by DuPont (US) for its POM. Both copolymer Delrin and homopolymer Delrin resins are used in industries. Copolymer resins are used in the applications where mechanical properties, including stiffness, low water absorption, and dimensional stability are required. POM is widely used in electrical components of aircraft, automotive applications, wire insulation for particularly high temperature applications, wire couplings and fittings, electrical and electronic applications with higher service temperatures, monofilament for the production of woven products for filters, belting and meshes etc. Machining of POM is generally difficult because of its properties such as low stiffness, low rate of moisture absorption, high coefficient of thermal expansion, and high internal stresses.

Polytetrafluoroethylene (PTFE) is a fluoropolymer that has many excellent properties such as high hydrophobicity, high oleophobicity, high chemical resistance, high antifouling property, high sliding property, high thermal resistance, high weather resistance, low relative permittivity, and low dielectric loss tangent. PTFE is used in polymer bearing, gaskets, gears, valve seats, piston rings, seals, bushes, hose assemblies, high voltage switches, extension joints, cook wears, clinical, containers and pipework for reactive and corrosive chemicals. It is also used in biomedical applications such as oxygenator membrane, vascular graft and catheter coating. The specific grade of PTFE that has the greater dielectric strength is mostly used in wire and cable wrap, and to separate conductive surfaces in capacitors. As an instance, thick-walled extruded PTFE can be machined into standoff insulators and can be used in high voltage encapsulation devices with high dimensional accuracy and integrity.

Polyether ether ketone (PEEK) is a biomaterial that has superior mechanical properties and high temperature durability. The ultimate tensile strength of this thermoplastic material is in the range of 90 to 100 MPa, its modulus of elasticity is about 3.6 GPa and the glass transition temperature is about 143°C to 250°C. It is preferred in many industrial applications including valves, bearings, pistons, seals manufacturing and bio-medical application. The implants or bone plates made of PEEK are viable alternative to Stainless steel and Titanium alloys. The orthopedic implants, bone plates and medical instruments are manufactured by casting, forging, sintering, machining, and recently additive manufacturing. These parts require machining like turning, drilling, grinding etc. The geometry of the joint implants, surgical instruments, molds or forging dies are different in shape and complex as well. Though the dimensional accuracy is easy to achieve, the surface finish is challenged to ensure.

The above said three materials are polymeric materials, while PEEK/MWCNT material is a reinforced polymer composite. The reason for including a polymer composite in this study is that characteristic of polymer composite is different from virgin polymer or plastic. Polymer composite is generally preferred for high strength applications. The reason for choosing turning operation in this study is that it is the most used operation in bringing the final product to near net shape.

3.2 Experimental Design

Design of Experiments (DoE) is generally done to identify the minimum number of experiments to be conducted in the selected levels. The reason to identify the minimum experiments is to reduce experimental costs and time. For preparing experimental design, firstly, the appropriate range and levels of each control parameter must be fixed. The range of values for each parameter is generally fixed based on the characteristics of the material.

In our experimentation, Speed (V_c), feed(f), depth of cut (ap) are control parameters (independent variables), while Surface finish (R_a) is the response variable (dependent variable). Table 3.1 shows the levels chosen for each material. The levels were fixed based on preliminary experiments conducted. L_{27} DoE matrix consisting of different combination of values was prepared for each material using Minitab software.

Table 3.1: Depicting features and levels used in DoE

Material	Machining Parameter	Level		
		I	II	III
POM	Speed (V_c) (m/minute)	90	135	180
	Feed (f) (mm/rev)	0.1	0.3	0.5
	Depth of cut (ap) (mm)	0.5	1.0	1.5
PTFE	Speed (V_c) (m/minute)	80	120	160
	Feed (f) (mm/rev)	0.1	0.3	0.5
	Depth of cut (ap) (mm)	0.5	0.75	1.0
PEEK	Speed (V_c) (m/minute)	95	125	155
	Feed (f) (mm/rev)	0.2	0.4	0.6
	Depth of cut (ap) (mm)	0.25	0.5	0.75
PEEK/MWCNT composite	Speed (V_c) (m/minute)	750	1500	2250
	Feed (f) (mm/rev)	0.15	0.45	0.75
	Depth of cut (ap) (mm)	0.1	1	1.8

3.3 Experimentation and Data Collection

For the experimentation, cylindrical rods in the size of 10 mm in diameter and 500 mm in length were purchased from the local supplier in Malaysia. Turning operations were performed as per respective DoE using CNC turning centre (Model: Sprint 16TC Fanuc 0i T Mate Model C). The experiments were carried out with servo super cut coolant 32t as per advice from the supplier. High carbon cemented carbide tool (Grade CNMG 120408 QM) was used in all experiments. The specification of the tool is Rhombie shape, insert angle=80°, tolerance=±0.13, insert size=12mm, insert thickness=4.76mm, insert clearance=0º. Figure 3.2 shows photographic illustration of turning of a sample.

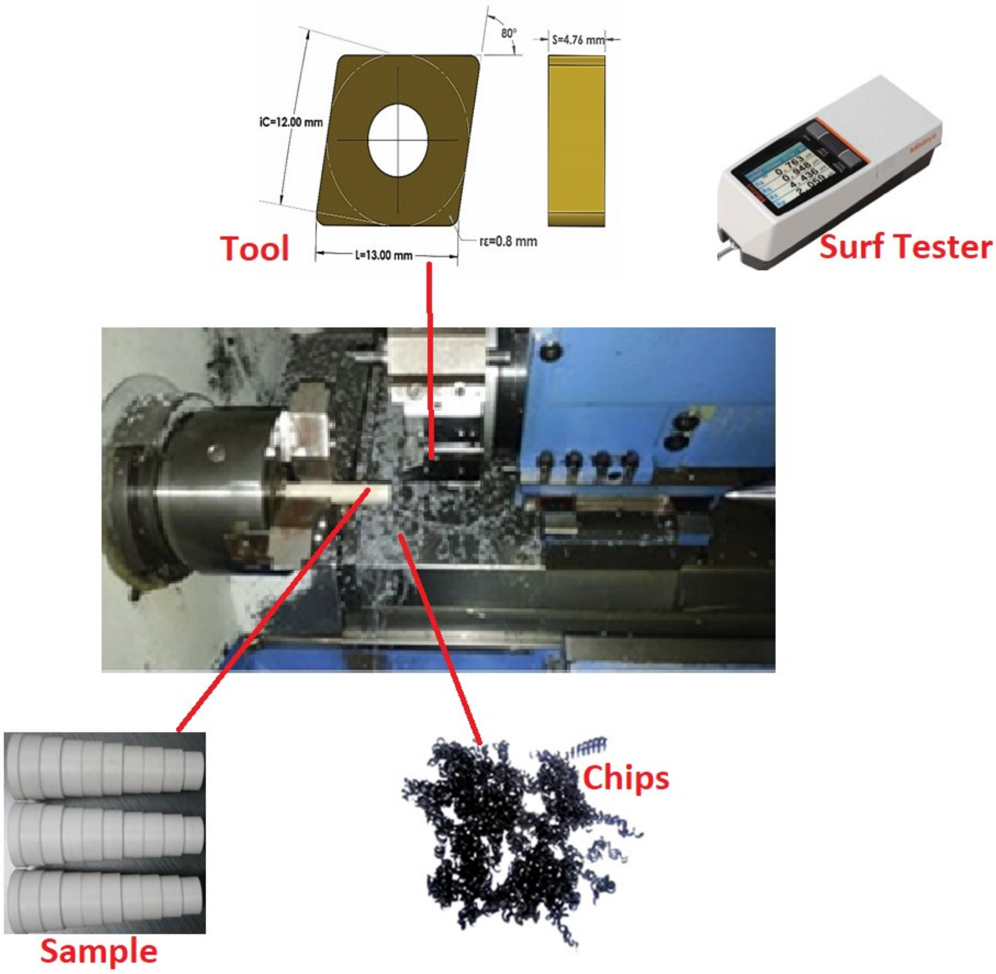

Figure 3.2: Photographic illustration of turning of a sample

Surface roughness of the machine sample was instantaneously measured with Mitutoyo make surf tester. Three trials were done for each experiment and the average of the measurements was recorded as shown in Table 3.2 – Table 3.5.

Table 3.2: Cutting condition and surface roughness measured from turning of DELRIN material

Expt. No.	**Cutting speed, V_c (mm/minute)**	**Feed Rate, *f* (mm/revolution)**	**Depth of Cut, *ap* (mm)**	***Surface roughness R_a (μm)***
1	90	0.1	0.5	0.79
2	90	0.1	1.0	0.61
3	90	0.1	1.5	0.56
4	90	0.3	0.5	1.88
5	90	0.3	1.0	1.78
6	90	0.3	1.5	1.74
7	90	0.5	0.5	1.67
8	90	0.5	1.0	1.59
9	135	0.5	1.5	1.65
10	135	0.1	0.5	1.18
11	135	0.1	1.5	0.84
12	135	0.1	1.0	0.66
13	135	0.3	0.5	1.60
14	135	0.3	1.5	1.72
15	135	0.3	1.0	1.66
16	135	0.5	0.5	1.50
17	135	0.5	1.5	1.80
18	135	0.5	1.0	1.43
19	180	0.1	0.5	1.19
20	180	0.1	1.5	0.89
21	180	0.1	1.0	0.67
22	180	0.3	0.5	1.62
23	180	0.3	1.5	1.65
24	180	0.3	1.0	1.60
25	180	0.5	0.5	1.42
26	180	0.5	1.5	1.61
27	180	0.5	1.0	1.59

Table 3.3: Cutting condition and surface roughness measured from turning of PTFE material

Expt. No.	**Cutting speed, *V_c* (mm/minute)**	**Feed Rate, *f* (mm/revolution)**	**Depth of Cut, *ap* (mm)**	***Surface roughness R_a* (*μm*)**
1	120	0.3	1	2.26
2	120	0.3	0.75	2.67
3	80	0.5	0.75	3.19
4	160	0.3	0.75	2.69
5	160	0.3	0.75	2.85
6	160	0.3	0.75	2.46
7	120	0.5	0.75	2.19
8	80	0.3	0.5	1.76
9	80	0.3	0.75	2.14
10	160	0.5	0.75	2.32
11	160	0.1	0.75	3.22
12	80	0.3	0.75	2.0
13	120	0.3	0.5	1.89
14	120	0.1	0.75	1.9
15	120	0.1	0.5	2.48
16	120	0.3	0.75	2.69
17	120	0.5	1	2.96
18	120	0.5	0.5	2.23
19	160	0.3	1	2.29
20	80	0.3	1	2.54
21	120	0.3	0.75	2.15
22	120	0.1	1	1.77
23	120	0.5	0.75	2.84
24	120	0.3	1	2.1
25	120	0.3	0.5	2.43
26	120	0.1	0.75	2.18
27	80	0.1	0.75	1.54

Table 3.4: Cutting condition and surface roughness measured from turning of PEEK material

Expt. No.	**Cutting speed, V_c (mm/minute)**	**Feed Rate, f (mm/revolution)**	**Depth of Cut, ap (mm)**	***Surface roughness R_a (μm)***
1	95	0.2	0.25	1.156
2	95	0.2	0.5	1.193
3	95	0.2	0.75	1.446
4	95	0.4	0.25	4.57
5	95	0.4	0.5	5.29
6	95	0.4	0.75	5.37
7	95	0.6	0.25	6.09
8	95	0.6	0.5	8.14
9	95	0.6	0.75	8.47
10	125	0.2	0.25	1.203
11	125	0.2	0.5	1.15
12	125	0.2	0.75	1.5
13	125	0.4	0.25	4.51
14	125	0.4	0.5	4.83
15	125	0.4	0.75	6.293
16	125	0.6	0.25	7.16
17	125	0.6	0.5	8.18
18	125	0.6	0.75	8.74
19	155	0.2	0.25	1.02
20	155	0.2	0.5	1.126
21	155	0.2	0.75	1.24
22	155	0.4	0.25	4.6
23	155	0.4	0.5	4.41
24	155	0.4	0.75	5.21
25	155	0.6	0.25	6.38
26	155	0.6	0.5	8.32
27	155	0.6	0.75	8.72

Table 3.5: Cutting condition and surface roughness measured from turning of PEEK/MWCNT composite material

Expt. No.	**Cutting speed, V_c (mm/minute)**	**Feed Rate, *f* (mm/revolution)**	**Depth of Cut, *ap* (mm)**	***Surface roughness R_a (μm)***
Units	rpm	mm/rev	Mm	μm
1	750	0.15	0.2	0.84
2	750	0.15	1	0.92
3	750	0.15	1.8	1.02
4	750	0.45	0.2	1.24
5	750	0.45	1	0.99
6	750	0.45	1.8	1.42
7	750	0.75	0.2	1.93
8	750	0.75	1	2.22
9	750	0.75	1.8	2.32
10	1500	0.15	0.2	0.83
11	1500	0.15	1	0.88
12	1500	0.15	1.8	1.04
13	1500	0.45	0.2	1.84
14	1500	0.45	1	1.69
15	1500	0.45	1.8	2.06
16	1500	0.75	0.2	2.30
17	1500	0.75	1	2.43
18	1500	0.75	1.8	2.66
19	2250	0.15	0.2	0.85
20	2250	0.15	1	0.88
21	2250	0.15	1.8	1.00
22	2250	0.45	0.2	1.45
23	2250	0.45	1	1.51
24	2250	0.45	1.8	1.68
25	2250	0.75	0.2	2.15
26	2250	0.75	1	2.26
27	2250	0.75	1.8	2.47

From the experimentation, 27 datasets for each material were collected. Each dataset has three input features and one output response variable. These data were used for training and testing in machine learning. The details of the model development is presented and discussed in Chapter 4.0

CHAPTER 4.0

META BASED ML MODEL DEVELOPMENT

An overview of machine learning and ML algorithms is firstly presented in this chapter. Meta based ML model development is further presented and detailed.

4.1 Machine Learning (ML)

Machine learning is a practice of programming machine to understand and learn the pattern of datasets and predict the response for unseen data. ML algorithm is used to iteratively learn from data to improve, describe data, and predict outcomes. There are four subset of Artificial intelligence (AI) methods exist as illustrated in Figure 4.1

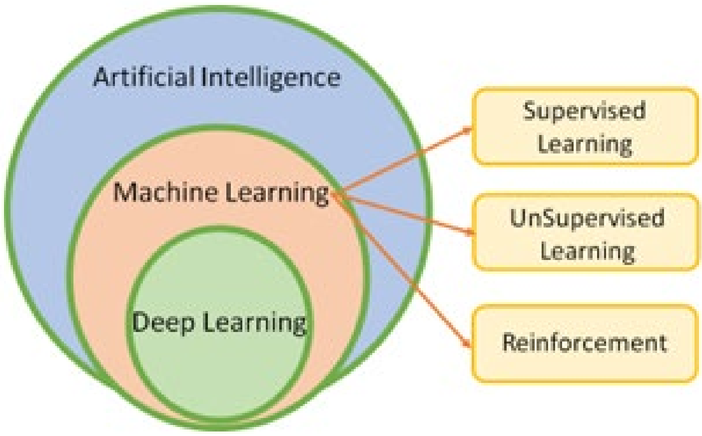

Figure 4.1: Types of Machine Learning Algorithms

In most of the use cases, supervised learning algorithms are used. In supervised learning, an established set of features are used to train the machine. Hence machine can certainly classify the data, understand the pattern of the data, and further apply in decision making. The features used in supervised learning are labeled features that define the meaning of data. For example, in the current research, speed, feed, depth of cut and surface roughness are features of the model. As the ML algorithm ingests training data, it is then able to produce more precise ML model based on the data. After training, the developed ML model can be used to predict the output based on the input given. To give an overview of ML algorithms, the following is presented.

Using Bayesian algorithms, data scientists is able to encode prior knowledge of what models should look like rather than what the data states. Clustering algorithms understand objects with similar characteristics and group them together in clusters. It is a fairly straightforward technique for unsupervised learning because the data is not labeled.

Branching structures are used by decision tree algorithm for demonstrating decision of the analysis. The model collects the possible outcome from each node of a decision tree and hence outcome of the model is formed from outcomes from all trees.

The most important supervised algorithms are

- K-nearest neighbors
- Linear regression
- Neural networks
- Support vector machines
- Logistic regression
- Decision trees and random forests

Though our objective is to develop a unified model, model development was started with individual material. This is because to confirm the effectiveness of the model development and also to learn the model development process with respect to the pattern of the data. In published literature, classifier and regressor algorithms were not combinedly used for ML model development. In the sense, they used either regression algorithm or classification algorithm in the model development. SVR algorithm was reported as a better for regression and neural network model was reported better for classification problem. This research was attempted to utilize both regressor and classifier in one model, which is a new concept known as meta based model. The output of the classifier was used as input to regressor. To evaluate the performance of models, we used Logistic regression and Extreme Boost (XGBoost) algorithms for classifier and SVR and XGBoost algorithms for regressor part of the models.

4.2 Dataset Description and Data Preparation

Datasets (Table 3.4 – Table 3.7) pertaining to all four materials were initially stored in in separate file. They have three input variables: Cutting speed (V_c), Feed Rate (f) and Depth of cut (ap) and an output variable: surface finish (μ) in micron. All these three inputs and the output data are nonlinear continuous data. As for as this problem is concerned, data cleaning was not required as only a small number of datasets in each material is available. And moreover, no null value was associated in the datasets.

One of the challenges in the model development is the level of imbalance in the dataset. Most of the researchers developed the models without considering the level of imbalance. Their focus was to evaluate the machining algorithms and to identify the suitable algorithm for the current datasets. But it could end up with low accuracy or with a false high accuracy because the model could only fit to the high-populated range of inputs and the output and perform very badly in the low-populated ranges.

To overcome the challenge of inaccuracy, and to improve the accuracy of prediction models, a strategy known as output discretization was adopted in this research. This strategy is to classify target variable using quantile distribution technique. If the total variation range of the deviation of surface roughness is divided in 3 intervals of the same size, there are x instances in the lower interval, y instances in the middle interval and z instances in the higher interval. Where x, y and z are decided based on the performance of the model.

Once the target output is classified into multi-class using distribution technique of Quantile regression, then the predicted multi-class output can be utilized as learning input parameter in second stage of model using regressor to predict the final result. This technique with Meta-learning algorithm is called stacked generalization. Stacking is a type of ensemble learning algorithm. Ensemble learning refers to machine learning algorithms that combine the predictions for two or more predictive models. Ensemble machine learning algorithm that uses meta-learning to combine the predictions made by ensemble members. In the meta based learning, experiences from multiple learning episodes are gained and utilized to improve its future learning performance. Herein highlighted to the readers that the model developed in our research is meta based ML model, even if not specified anywhere in the thesis.

4.3 Quantile Distribution

Quantile distribution method (0-100) was applied to split the data into the proper number of groups based on surface roughness (response variable). Based on the output from quantile distribution, the datasets of the respective material were divided into two groups. For example, Delrin material datasets were grouped into two: Group 1 that has 14 datasets ($\mu<=1.590$) and Group 2 that has 13 datasets ($\mu>1.590$ and $\mu<=1.88$). The choose of the datasets was not random, it was done by pareto principle. Table 4.1 shows grouping of data and data size used for training and validation.

Table 4.1: Output from Quantile Distribution

Material	Classes from Quantile percentage		Data Size	
	Group 1	Group 2	Training	Testing
Delrin	$\mu<=1.590$	$\mu>1.590$ and $\mu<=1.88$	21	6
PTFE	$\mu<=2.29$	$\mu>2.29$ and $\mu<=3.22$	21	6
PEEK	$\mu<=4.83$	$\mu>4.83$ and $\mu<=8.74$	21	6
PEEK/MWCNT	$\mu<=1.51$	$\mu>1.51$ and $\mu<=2.66$	21	6

4.4 Correlation Analysis

Foremost step in ML model development is to ensure the relationship between independent parameters and dependent parameter and identifying the significance of the parameters. Hence correlation between inputs at outputs was done using Analysis of Variance (ANOVA). The motivation of the correlation analysis done in this research was to check if the inputs are correctly selected, and to check the inputs have any influence on the output. The significance was checked as: (i) if the value of probability P <= 5%, the respective parameter is adequate and significant on response and (ii) if the value of P > 5%, the respective parameter is insignificant on response. Table 4.2 shows the output of correlation analysis.

It is concluded from the analysis that feed is the most significant parameter on surface roughness as for as individual material is concerned.

Table 4.2: Output from Correlation analysis

Parameter	DELRIN (Material 1)		PTFE (Material 2)		PEEK (Material 3)		PEEK/MWCNT (Material 4)	
	F_STATTISTIC	**P_VALUE**	**F_STATTISTIC**	**P_VALUE**	**F_STATTISTIC**	**P_VALUE**	**F_STATTISTIC**	**P_VALUE**
Speed	0.0772	0.7842	0.07715	0.784197	0.2679	0.6107	0.0531	0.8201
Feed	4.2564	0.0530	4.256423	0.053039	30.7619	0.000024	38.2657	0.000006
Depth of Cut	1.6819	0.2102	1.681983	0.21019	2.2352	0.1513	0.0637	0.8034

4.5 Evaluation Metric

Before explaining our ML model development, model evaluation metrics used in the current research is herein explained. In the process of ML model development, error is estimated which is otherwise called evaluation of residuals. It is done in the training phase of the model development. Difference in predicted and original responses is found as a numerical estimate, also called the training error. In this research, the goodness or accuracy of the meta based model was assessed by determining the largest sum of square (R^2) and the smallest Root Mean Squared Error (RMSE). During the training process, the performance of each model was monitored continuously and recorded with numerical results quantifying hypothesized relationships between variables. The prediction model with the lowest RMSE error is considered the best model.

4.6 Selection of ML Algorithms

Our meta based model will use two ML algorithms in a pipeline as explained below: classifier and regressor. Classifier is the one initially considers the experimental data and classifies them into reasonable and more accurate groups. Regressor is the one further considers the data in groups and predict the surface finish. It is depicted in Figure 4.2 below.

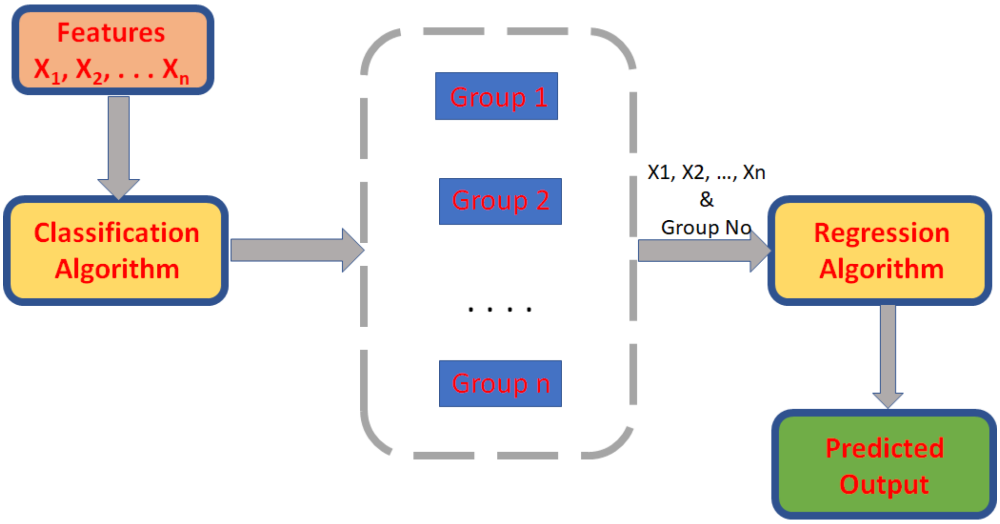

Figure 4.2: Illustration of working of Meta based ML model

XGBoost algorithm and Logistic regression algorithm were used for developing classifier part, while Support vector regressor (SVR) and XGBoost algorithm were used for developing regressor part of the meta based ML model.

4.6.1 Logistic Regression Algorithm

Logistic regression is a data-driven model, which is used to predict two values such as success/failure or yes/no. In this algorithm, data and relationship between one binary variable (dependent variable) and one or more nominal or ordinal or interval variables (independent variables) are defined. The formula of Logistic regression is as follows:

$$P(Y = 1|X_1, \ldots, X_P) = \frac{1}{1+e^{-(\beta_0 + \beta_1 X_1 + \ldots + \beta_P X_P)'}} \quad (1)$$

Where $P(Y = 1|X_1, \ldots., X_P)$ is the probability of being Non-pass product (P(Y = 1)) under the given manufacturing receipe $(X_1, \ldots.., X_P)$ the coefficients

$\beta_0, \beta_1, \dots, \beta_p$ are the effects of each explanatory variable estimated by maximum-likelihood from date[19]. The probability that Y=1 (P(Y=1)) depends on the value of X as shown in Figure 4.3.

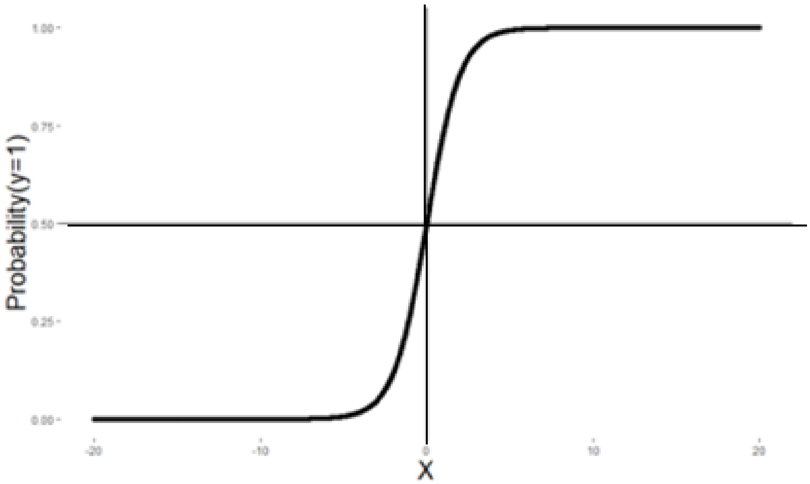

Figure 4.3: Graphical representation of a Standard Logistic Function

In Logistic regression, a threshold can be used to foresee the appropriate class for a data. The model groups the data into classes based on the set threshold. Increasing the order of polynomial and model fit are two significant considerations in the model development. The order of polynomial is increased to get the complex decision boundary, which can be linear or non-linear. The value of variance represented in the log odds (typically expressed as R^2) is increased when number of independent variables is increased. But, adding many more variables to the model may result overfitted model. Though many pseudo-R^2 values have been established, we must take intense caution for interpreting them, as they involve with many computational issues. It is the best to use any of the goodness of fit tests like Hosmer-Lemeshow. This commonly used goodness of fit is based on Chi-square test.

4.6.2 Extreme Gradient Boosting (XGBoost) Algorithm

Decision tree approach is a simple and easier method of modeling that interprets the features and conclude the response of the subject. This method has been used in statistics, data mining, and machine learning. The accuracy of decision tree is dependent upon the size of datasets, greater the amount of data available, higher the accuracy. Gradient boosting algorithm is a machine learning technique that can be used both in regression and classification problems. The term boosting refers to a family of algorithms that convert weak learners in the datasets into strong learners. It could understand that weak learners are slightly better than a random choice, while strong learners are perfect in performance. This approach can produce an ensemble predictive model from weak predictive models. In gradient boosting algorithm, gradient descent in function space is stage-wise used to construct the ensemble. The final model is a function taking input parameters as a vector of attributes $x \in R^n$ to get $F(x) \in R$. In $F_i(x) = F_{i-1}(x) + \gamma_i h_i(x)$, h_i is a function that models a single tree and $\gamma_i \in R$ is the weight associated with i^{th} tree. These two terms; function h_i and weight γ^i are learned during the training phase. Gradient boosting algorithm is more reliable and easier when compared to other machine learning algorithms.

Algorithm like linear regression has its number of degrees of freedom scaling with the number of features $O(M)$. It means that its ability to learn from the data plateau in the regime $N \gg M$, where *N* is the number of samples and *M* is number of features. The linear regression algorithm results low variance, but high bias. In the $N < M$ regime, L_1 regularization becomes necessary to learn the relevant features and zero-out the noise. A tree in its un-regularized form, has a low bias which can over fit the data to extreme, with depth of field scaling as $O(N)$, but it has a high variance (i.e., deep trees don't generalize well). But because a tree can reduce its complexity as much as needed, it can work in the regime $N < M$ by simply selecting the necessary features. A Random Forest is a low bias algorithm and the ensemble averages away the variance (but deeper trees call for more trees) and it doesn't overfit on the number of trees, so it is a lower variance algorithm. The homogenous learning that the trees tend to be similar and tends to limit its ability to learn more on much data.

XGBoost is one of the tree algorithms to mathematically formalize regularization in a tree. It is adapted to large data scales, as it has a low bias and high variance (due to the boosting mechanism). It is a parallelized and carefully optimized version of the gradient boosting algorithm. It has improved the training time by parallelizing the whole boosting process as depicted in Figure 4.4.

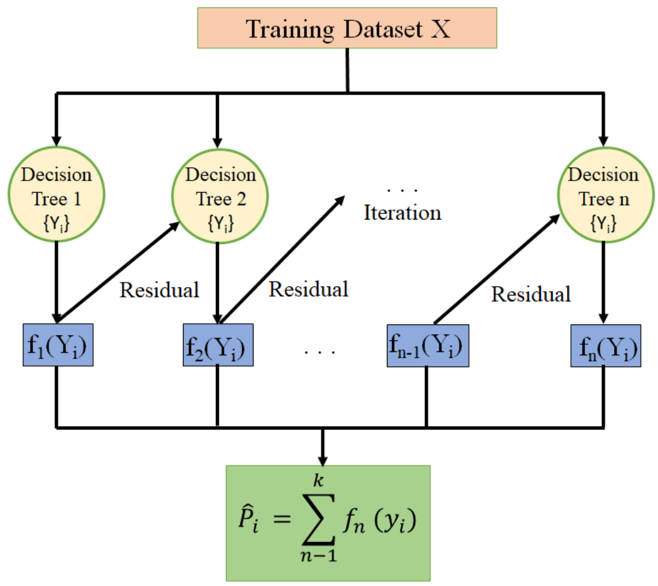

Figure 4.4: Flowchart of a XGBoost Model

Pseudo code for XGBoost algorithm is,

- Input: training set $\{X_i, Y_i\}$, a differentiate Loss Function $L(Y, F(x))$, number of iterations M.
- Initiate model with constant value: F_0 = γ (i.e., fits to actual value)
- Compute so-called Pseudo-residuals and fit a base learner (Ex. tree) to pseudo-residuals. That is training it using the train set $\{X_i, Y_i\}$,.
- Compute multiplier by solving the one-dimensional problem and followed by that update model
- Output of Model $F_m(X)$

As explained above, gradient boosting takes the training set and a loss function as inputs and the final trained model is gotten at the end of the algorithm by output $F_m(X)$.

4.6.3 Support Vector Machine

Support Vector Machine (SVM) is a popular machine learning tool for classification and regression. It is considered as a nonparametric technique as it relies on kernel functions. In SVM regression, the learning task is transformed to the minimization of the error function, defined through so called insensitive loss function ($\in$) which controls the accuracy of the regressor. In the development of Support Vector Regressor (SVR) model, the given dataset was firstly normalized. This method is known as Min-Max Normalization technique using standard scalar library available in scikit-learn. Data normalization is a part in transformation of data to obtain the same weight for all attributes of data. It makes the result of weighting has no dominant attribute. The reason to apply normalization technique in the current experimental machining datasets is that the data are discrete in nature, and metrics are handled using Euclidean distance whereas SVM is unable to apply directly. In the analysis using SVR, $\{X_i, Y_i\}_{i=1}^{N}$ is considered as a training set, such that $(X_i \in R^p)$ represents p-dimensional input vectors. The scalar measured output is denoted as $(Y_i \in R)$. The regression model can be constructed using a non-linear mapping function $\emptyset(*)$. The basic idea is to transform the feature space (also known as higher dimensional) and learn a linear regressor in the new space (known as non-linear data), using kernel function. This can do mapping implicitly.

The main objective is to construct a function $y = f(x)$, which in general represents the dependence of output y_i on the input of x_i. Thus, the form of function using kernel is,

$$y = W^T \emptyset(x) + b \tag{2}$$

Where W is known as weight vector and b is the bias. The function $\emptyset(*) : R^p \rightarrow R^h$ is mostly a non-linear function which maps the training data into a higher dimensional, possibly infinite, dimensional feature space. Kernel function fn allows us to compute inner products in R^h implicitly without using (or even knowing) $\emptyset(*)$.

In the current case, given training data $\{(X_1, Y_1) \ldots (X_n, Y_n)\}$, wants to find the best function to predict given *X*, such that $X_i \in R^m,\ Y \in R$. In basic cases of SVM, $Z = \emptyset(x) \in R^m$, especially learns the linear model in a transformed space, where $\emptyset(x)$ is an independent variable.

$$\sum L\,(Y_i, g(X_i, W)) \tag{3}$$

In general process of regression problem, finding the coefficient to further minimize the error loss is the significant activity for model development. In the above equation $g(X_i, W)$ represents the learning of the model, L denotes to be loss in the function and Y_i is the prediction. This is the general strategy of empirical risk minimization. Hence, the suitable loss function that allows using the kernel trick and gives better performance for the unseen data is to be chosen. In non-linear SVM, there is a method known as Gram matrix is used to find the optimal function $f(X)$ in transformed prediction space. The elements contained as $g_{bj} = G(X_i, Y_i)$ is the Gram matrix which is of course in n-by-n matrix. Each element g_{bj} is equal to the inner product of the predictors as transformed by $\emptyset$. The kernel function can be used to generate Gram matrix and subsequently the loss function**.** And finally, this non-linear SVM regression model finds the coefficient that minimizes the error.

Kernel functions such as Radial basis function (RBF) and Polynomial functions were attempted and evaluated. Each kernel function has some parameters that help to obtain the better performance for algorithm and model optimization.

In RBF, C and γ are the significant parameters in the below equation

$$K(X_n, X_i) = \exp(-\gamma||X_n - X_i||^2 + C), \tag{4}$$

In Polynomial, C, $\gamma, r\ and\ d$ are significant parameters in the below equation

$$K(X_n, X_i) = (\gamma(X_n, X_i) + r)^d, \tag{5}$$

It is found from the evaluation of these kernel functions that RBF results over-fitting of training data as these machining data are discrete in nature. The polynomial kernel that uses the quadratic condition results a good fitting of training data.

Pseudo code for SVR algorithm,

Step 1: Input of training set, $\{X_i, Y_i\}_{i=1}^{N}$

Step 2: Process of non-linear mapping function find $\emptyset(*)$ as solution of the optimization problem

Step 3: In max $\emptyset(*)$, using polynomial kernel function $K(X_n, X_i) = (\gamma(X_n, X_i) + r)^d$

Step 4: Subject to find the coefficient to minimize error loss, $\sum L\left(Y_i, g(X_i, W)\right)$

Step 5: Using the above technique non-linear in higher dimensional to be transformed, feature space in kernel function $y = W^T\emptyset(x) + b$

Step 6: $G(X_i, W) = (1 + X_i W)^q$, where (q = 2, 3,..), such that Output of weight vector W, kernel function to be implemented on Support Vector Regression.

4.7 ML Model Development for Individual Material

As stated in previous sections, ML model for individual material was started first. The training was started with 80% of data in the respective database and testing was done with 20% of data. Pycharm Community Edition 2022.3.2 was used for model development. For all model development, Intel(R) Core(TM) i7-7500U CPU @ 2.70GHz 2.90 GHz, 12.0 GB RAM was used. For classifier part, XGBoost algorithm and Logistic regression algorithm were used to develop the models. Models considered speed, feed, depth of cut and surface roughness as features.

4.7.1 Cross Validation Method (k-fold method)

The description and working of k-fold cross validation method is detailed below. The evaluation of residuals estimates the difference in predicted and original responses, but it does not indicate how well ML model will respond to real data or unseen data. The stability of the model in working with real data (unseen data) is to be checked, which is otherwise called validation of the model. It is to confirm whether model has not considered the noise data, but it has considered most of the pattens from the right data. This process in ML model development is called cross validation. Generally, data analyst chooses some percentage of datasets for training, some sets for testing and some sets for validation. When we allocate some sets for testing and validation, there is a risk at the model to miss out some patterns in the datasets and it may possess underfitting. But, if we use k-fold cross validation, it will use sufficient data for training and testing, and will not leave any data for both training and testing. It divides the data into k subsets and uses one of k subset for testing and k-1 subsets for training. It averages the error from k trails and finds the effectiveness of the model based on the averaged error. As it goes, every data point gets to be training set k-1 times and in a validation set exactly once. This method reduces unfairness of avoiding the data points as most of the data are used for fitting. It reduces variance as most of the data are also being used in validation. Besides, the effectiveness of the model is increased by interchanging the training and test sets.

K-fold cross validation method was applied in every part of our model development and hence ensured that all datasets were used both in training and testing.

Besides, Grid searching method was applied to tune the hyper parameters of each algorithm. After tuning hyper parameters in each iteration, performance of the model was measured using the largest sum of square (R^2) and the smallest Root Mean Squared Error (RMSE). This was continued iteratively till the maximum performance was achieved in each classifier model. Finally, the better classifier model was chosen to proceed in the next part of ML model development.

4.7.2 Regressor Model- A Final Part of Meta Based ML Model

The output of the classifier was used as a new feature (vector) for further model development. That means, speed, feed, depth of cut, surface roughness and group number were considered as features. Support vector regressor (SVR) and XGBoost algorithms were algorithms considered for model development and evaluation. Cross validation method and grid search method as explained in the previous section were also adopted, and hyper parameters were tuned iteratively till the performance of the model is in the acceptable range. After checking the performance metrics of each model, the best meta-based model was chosen. The results are presented and discussed in Chapter 5. Figure 4.5. depicts the meta based model pinning to individual material.

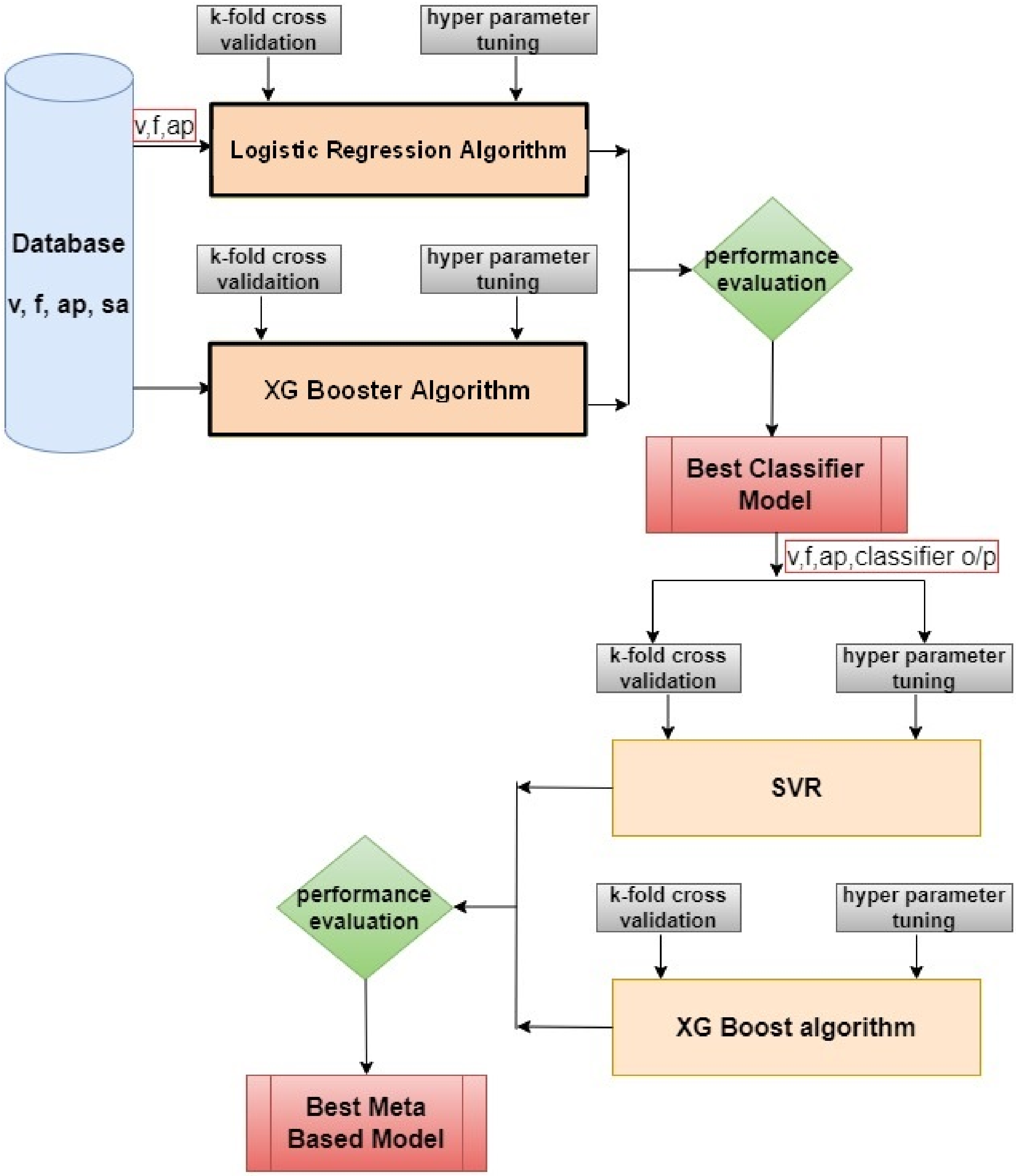

Figure 4.5: Illustration of meta based model development pinning to individual material

4.8 Unified Meta Based ML Model

After confirming our methodology, a unified meta based ML model was developed. The name unified is coined as this model is a holistic model that uses all four materials. In the sense, the database has the datasets of all four materials (4x27=108 datasets). The same procedure explained in section 4.7 was followed, but it used an additional feature (material). Classifier part used speed, feed, depth of cut, material number and surface roughness as input. The regressor part used five features such as speed, feed, depth of cut, material number, output from the classifier (group number) and surface roughness. Figure 4.6 depicts the model development. After developing models, the performance of them was measured in each iteration.

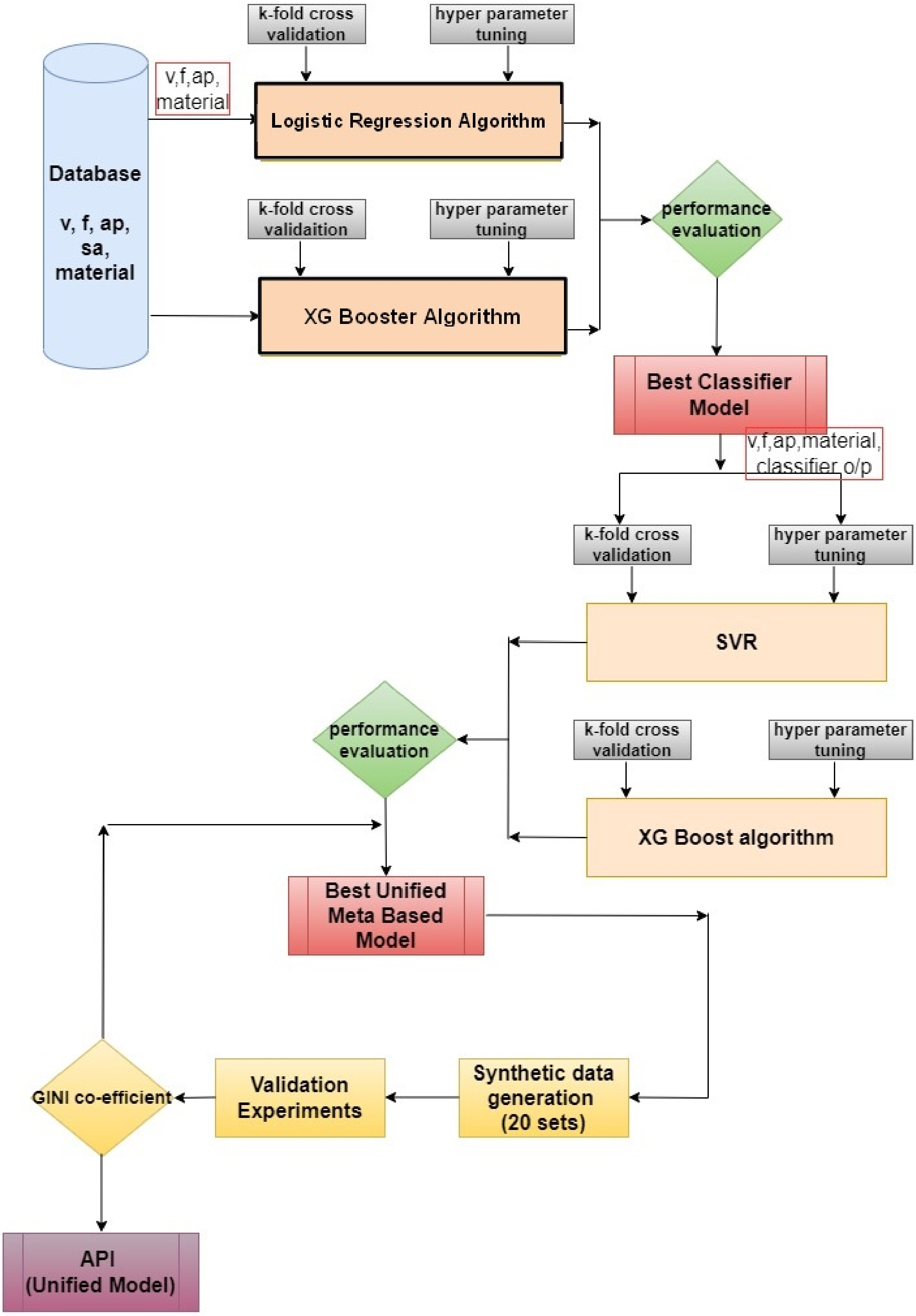

Figure 4.6: Illustration of unified meta based model development

4.9 Experimental Validation

Once the best meta based model was evaluated, it was used to generate 20 predicted results, which further taken to experimental validation. The reasons for experimental validation were: (1) to estimate the error in prediction (2) To use all these new datasets further model training. Once unified meta based ML model was finalized, application user interface (API) was developed for easy access and prediction.

CHAPTER 5.0

RESULTS AND DISCUSSION

In the process of developing a unified meta based ML model, supervised algorithms such as SVR, XGBoost, Logistic regression algorithms were involved. Meta based ML models for individual material and unified meta based ML model were trained, tested and experimentally validated for their efficacy. This section presents and discusses the output of each model.

5.1 Individual Model Performance

Table 5.1 details the accuracy of classifier models for each material. It is concluded from the results that Logistic Regression model is the better model in all cases. Because it predicts probabilities of each observation in each class based on thresholds. It does not predict classes directly. It uses trade-off concerns or errors like number of false positives with respect to the number of false negatives. This is essential because each concern or error would affect other type of the error.

Table 5.1: Performance Metrics from Classifier models

Delrin (Material 1)	**Accuracy**		**F1-Score**	
	Train	Test	Train	Test
Logistic Regression model	71.4	66.6	73	67
XGB model	52.3	33.3	69	50
PTFE (Material 2)	**Accuracy**		**F1-Score**	
	Train	Test	Train	Test
Logistic Regression model	71.4	66.6	73	67
XGB model	52.3	33.3	69	50
PEEK (Material 3)	**Accuracy**		**F1-Score**	
	Train	Test	Train	Test
Logistic Regression model	94.3	91.8	92	89
XGB model	52.3	48.2	69	48
PEEK/MWCNT (Material 4)	**Accuracy**		**F1-Score**	
	Train	Test	Train	Test
Logistic Regression model	90.2	83.33	91	86
XGB model	52.3	50	69	67

Receiver Operating Characteristic (ROC) Curves and Precision-Recall curves are two diagnostic tools used in binary classification models. They help to interpret the probabilistic forecast of the model. Relative trade-off between the number of true positives and false positives based on thresholds are given by ROC Curves. Relative trade-off between number of true positive and the positive predictive value based on different probability thresholds are interpreted by Precision-Recall curves. ROC curves graphically present the performance of the classifier. Figure 5.1 – Figure 5.4 shows Precision-Recall curve and ROC curves for each individual material model.

For individual material, Meta based models used four variables (speed, feed, depth of cut and surface_roughness_class) in which the variable surface_roughness_class is the output of Logistic Regression classifier model. It is concluded from our training and testing that, XGB model is the better model in which R^2 is close to100 in all cases. Table 4.2 details the accuracy of meta based models for each material.

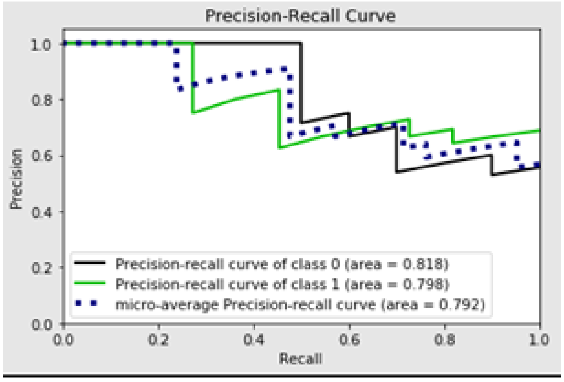

Figure 5.1a: Precision-Recall curve for Delrin models

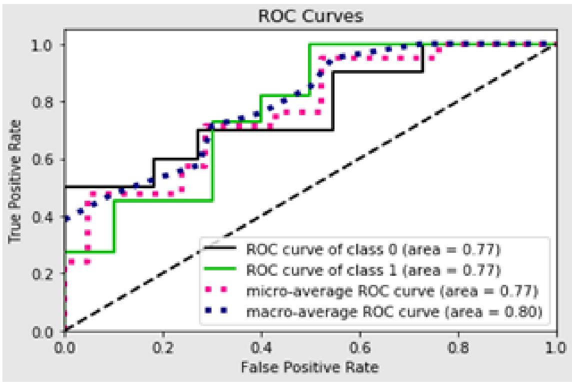

Figure 5.1b: ROC curves for Delrin models

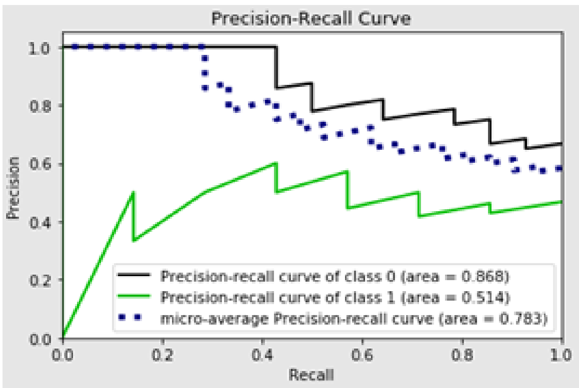

Figure 5.2a: Precision-Recall curve for PTFE models

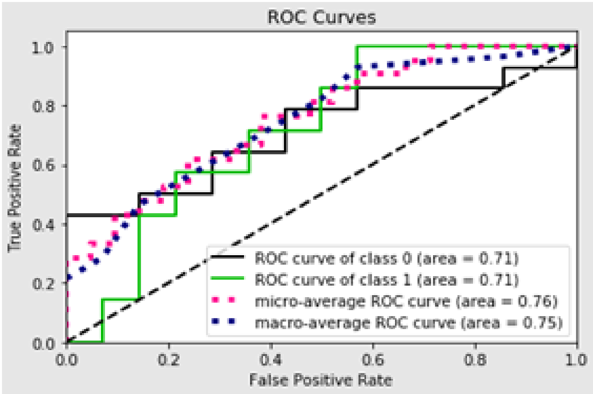

Figure 5.2b: ROC curves for PTFE models

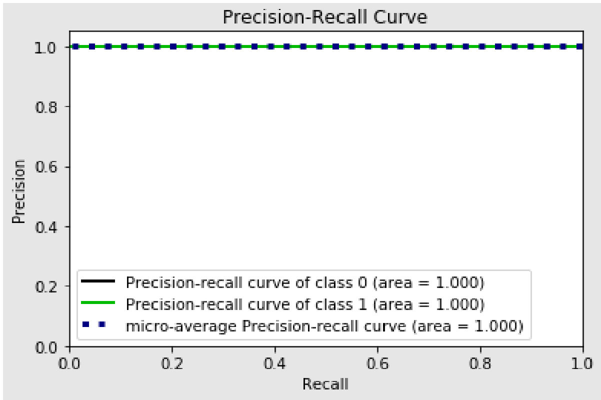

Figure 5.3a: Precision-Recall curve for PEEK models

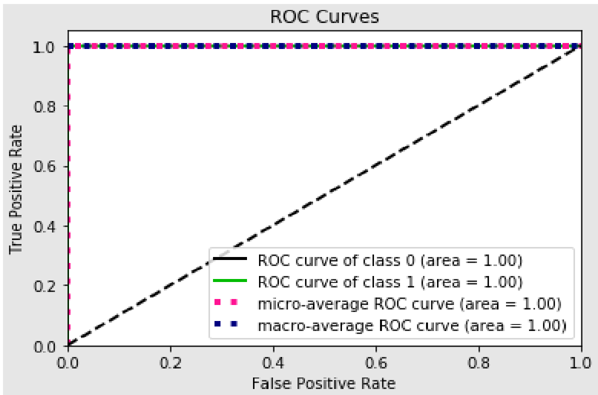

Figure 5.3b: ROC curve for PEEK models

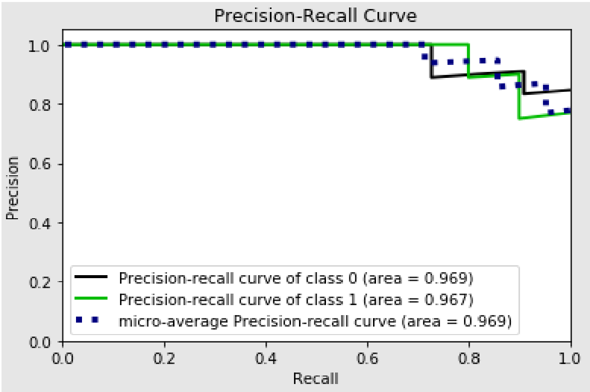

Figure 5.4a: Precision-Recall curve for PEEK/MWCNT models

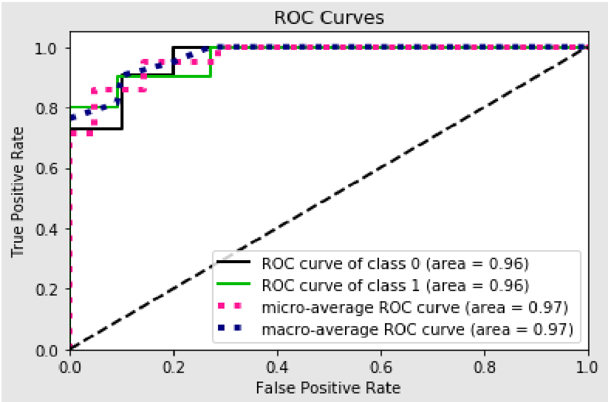

Figure 5.4b: ROC curves for PEEK/MWCNT models

Table 5.2: Performance metrics of meta based model for individual material

Delrin (Material 1)	R^2		RMSE	
	Train	Test	Train	Test
SVR	79.7	74.2	0.089	0.095
XGB	99.67	97.21	0.031	0.039
PTFE (Material 2)	R^2		RMSE	
	Train	Test	Train	Test
SVR	82.6	79.1	0.067	0.078
XGB	99.95	90.5	0.031	0.039
PEEK (Material 3)	R^2		RMSE	
	Train	Test	Train	Test
SVR	82.8	74.3	0.091	0.096
XGB	99.55	96.93	0.029	0.041
PEEK/MWCNT (Material 4)	R^2		RMSE	
	Train	Test	Train	Test
SVR	81.2	78.9	0.077	0.071
XGB	99.91	97.31	0.024	0.029

5.2 Unified Meta Based ML Model Performance

Once the meta based model for each material was developed, the same procedure was followed to develop a unified meta based ML model. This model considered speed, feed, depth of cut and material as input to the classifier model and used the output of the classifier too in the regressor model. The performance of both SVR and XGB model was verified using R^2 and RMSE values. For unified model, it is found from ANOVA that feed and material are most significant parameters as shown in Table 5.3.

Table 5.3: ANOVA result for Unified meta based ML model

ANOVA Result (Feature Significance)

	F_STATISTIC	P_VALUE
Speed	1.062051	0.305706
Feed	31.501198	2.51E-07
Depth of Cut	0.655936	0.420287
Material	3.821137	0.053936

Table 5.4 shows the accuracy of classifier in unified meta based ML model, where we see XGB model has the higher accuracy of 90% in training and 86.3% in testing. Figure 5.5a & Figure 5.5b shows Precision-Recall curve and ROC curves for unified meta based ML model. Table 5.5 shows the overall model performance metrics of unified meta based ML model, where we find XGB has resulted almost 100% accuracy in training and 98.86% in testing. This is different from the meta based model of the individual material. For individual model, Logistic regression algorithm and XG Boost algorithm were the best in classifier and regressor respectively. But for unified meta based model, XG Boost algorithm was the best for both classifier and regressor. In the past, XGB algorithm was attempted for different applications as discussed below. Ting Hu and Ting Song (2019) used XGB algorithm for forecasting analysis, where the grades of students were used. Ahmedbahaaaldin Ibrahem Ahmed Osman et al (2021) used XGB algorithm for modeling ground water level in a particular state in Malaysia. They claimed that this algorithm outperformed other regression algorithm and neural network algorithm. Zhang P et al (2022) used XGB algorithm for under sampled data, while SVM-SMOTE was used for over sampled data. They used public open-source data for evaluating their model. Montomoli J et al (2021) developed XBG model to predict increase or decrease in Sequential Organ Failure Assessment (SOFA) after 5 days of any COVID patient is admitted into ICU. Su,W et al (2023) used XGB algorithm to develop knowledge tracing model for online education. In all these research, XGB was found the best performer.

The reason for XGB to out perform in the current research is that XGB would be more accurate when it is trained by high volume of data. Since unified meta based model used about 108 datasets in the initial training itself, it was able to give more accurate results accordingly.

Table 5.4: Performance of classifier in Unified meta based ML model

	Accuracy		F1-Score	
Model	Train	Test	Train	Test
Logistic Regression	81	72	83	70
XGB	90	86.3	89	82

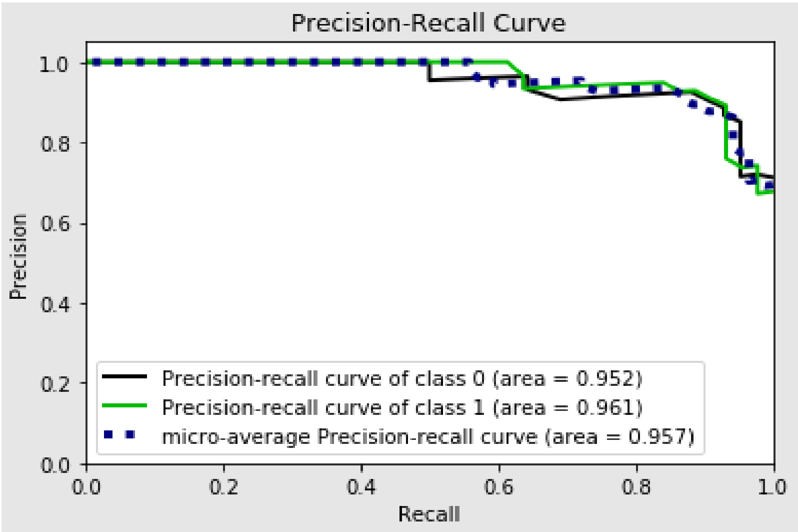

Figure 5.5a: Precision-Recall curve for Unified Meta Based ML model

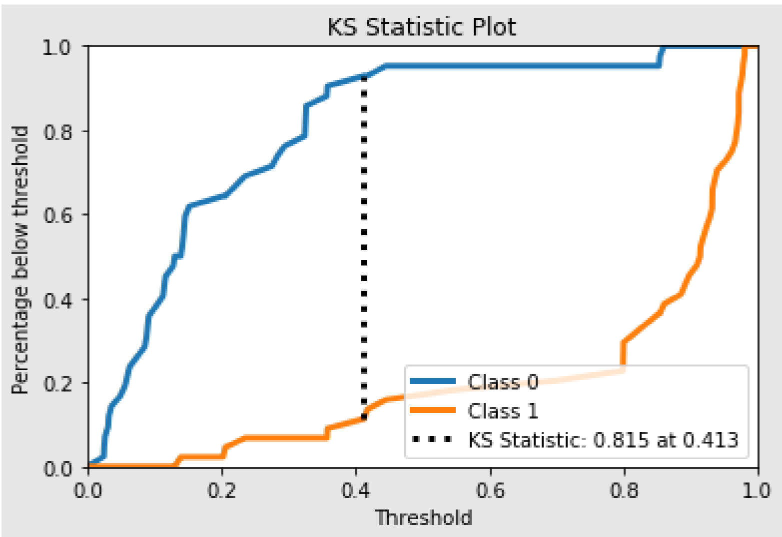

Figure 5.5b: KS Statistic plot for Unified Meta Based ML model

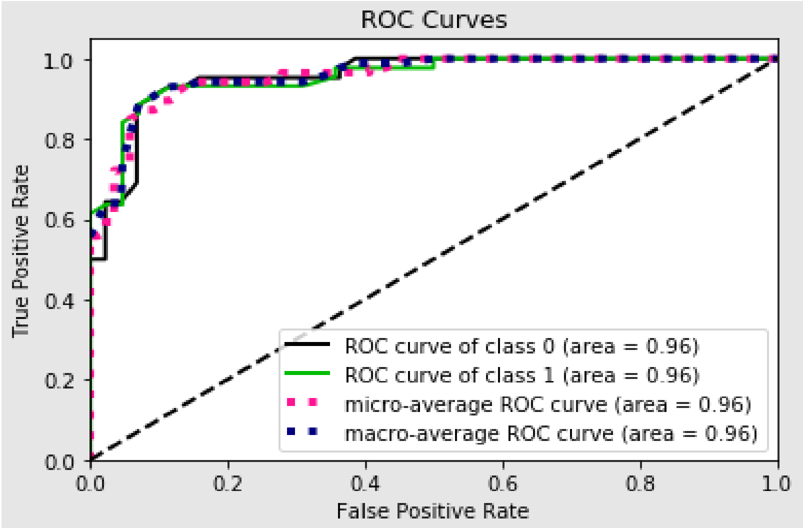

Figure 5.5c: ROC curves for Unified Meta Based ML model

Table 5.5: Performance metrics of unified meta based models

Meta Model	**R^2**		**RMSE**	
	Train	Test	Train	Test
SVR	85.1	77.4	0.083	0.0905
XGB	99.98	98.86	0.0178	0.023

5.3 Experimental Validation

After the best unified model was confirmed (XGB meta based model), it was run to generate twenty predicted values. These predicted values were corresponding to some independent parameters. Independent parameters were anonymously selected. Validation experiments were further conducted with these independent parameters and their corresponding surface roughness was measured. Table 5.6 shows the experimental values and predicted values. On plotting a scatter plot of these values (shown in Figure 5.6), it is seen that the predicted values are very close to the actual values.

Table 5.6: Predicted values and validation experiment results

SNo	Input parameters			Material	Surface Roughness (True Value or experimental value)	Surface Roughness (Predicted Value)
	Cutting speed, V_c (mm/minute)	Feed Rate, *f* (mm/revolution)	Depth of Cut, *ap* (mm)			
1	80	0.4	0.6	1	2.85	2.89
2	95	0.5	1.5	1	1.73	1.75
3	130	0.3	2	1	1.70	1.72
4	125	0.4	0.6	1	2.62	2.65
5	160	0.4	0.6	1	2.43	2.43
6	80	0.4	0.6	2	5.05	5.10
7	95	0.5	1.5	2	4.43	4.44
8	130	0.3	2	2	4.11	4.19
9	125	0.4	0.6	2	4.80	4.82
10	160	0.4	0.6	2	4.11	4.13
11	80	0.4	0.6	3	3.55	3.55
12	95	0.5	1.5	3	2.97	2.96
13	130	0.3	2	3	2.90	2.89
14	125	0.4	0.6	3	3.29	3.30
15	160	0.4	0.6	3	3.01	3.02
16	80	0.4	0.6	4	3.55	3.55
17	95	0.5	1.5	4	2.92	2.96
18	130	0.3	2	4	2.87	2.89
19	125	0.4	0.6	4	3.31	3.30
20	160	0.4	0.6	4	2.99	3.02

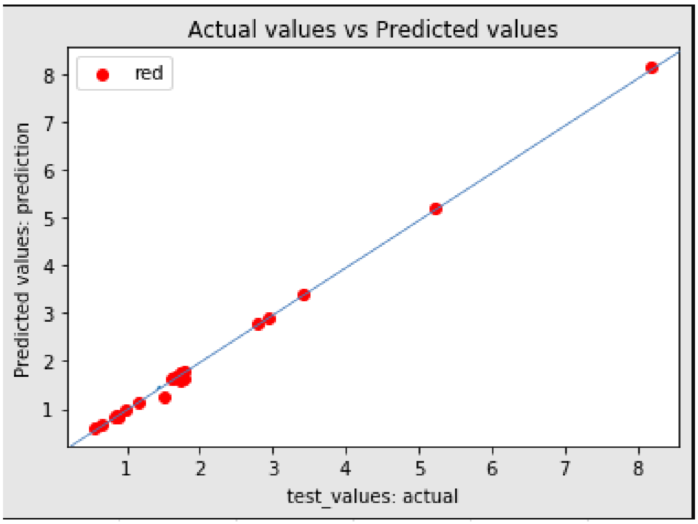

Figure 5.6: Scatter plot showing Actual vs Predicted values

The reason for higher efficacy in Extreme Gradient Boosting model is that it is a gradient descent type of an algorithm. In each round, it takes the current ensemble that it has and computes a gradient, i.e. a direction in which the model can improve (actually the direction of improvement is the opposite side of the gradient but let's put that aside). With this direction in hand, it trains a tree to predict it and adds it to the gradient. Therefore, each additional tree tries to get the model closer to the target and reduce the bias of the model rather than the variance. It is concluded from the experimental validation that no more further training with new datasets is required.

5.4 User Interface

One of the objectives of this research is to develop a user interface (application program interface) which operators can access and estimate the surface roughness before they take up the turning operation. For successfully work through this, Python and its packages such as Pandas, numpy, scikit-learn, XGBOOST and Gradio were used.

Gradio is a GUI library that allows to create customizable GUI components for Machine Learning model. Application program interface (API) was coded and embedded with our unified meta based ML model. It is presently stored in local server. It has slider input controls and text box controls for input values. The user can either input the value directly or use the slider to assign the value. Upon entering values for three independent variables, the developed unified meta based model will result the corresponding predicted response on the screen. Indeed, the interfaceclass in Gradio will be initiated with the following three parameters; show_sentence() - a function to trigger, user input in text form, and output text. The function named launch() will be used to output the result on the screen. The function provides a way for Gradio to get input from users and pass it on to the ML model, which will then process it and then pass it back to Gradio to show the predicted output. Figure 5.7 shows a sample screen where input was fed, and the corresponding surface roughness was predicted.

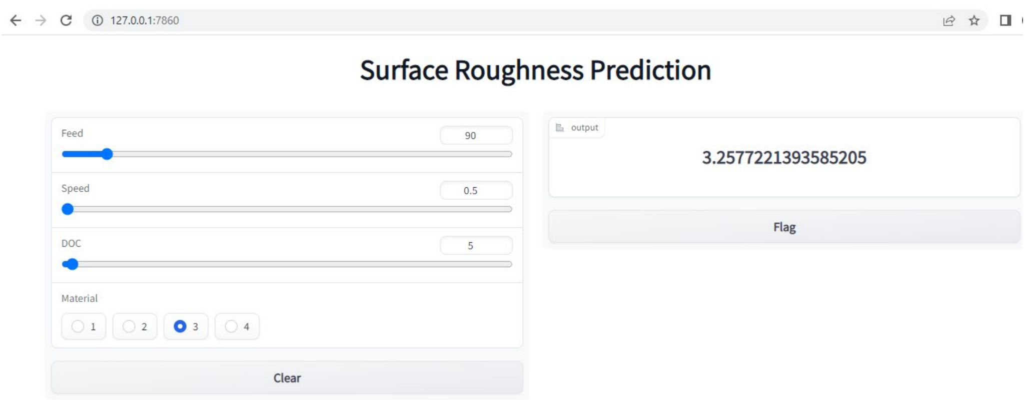

Figure 5.7 A sample screen from API

CHAPTER 6.0

CONCLUSION

6.1 Conclusion of the Research Work

Polymeric materials are now seen as a vibrant alternative to metals because of its high weight-strength ratio. In last two decades, fiber reinforced materials have been adopted in transport, aerospace, biomedical and many more industrial applications. Considering its potentiality, this research was attempted to investigate and develop a unified meta based machine learning model for turning of different polymeric materials. Meta based model is the one utilizing both classification algorithm and regression algorithm to achieve the predicted output. The classification algorithm was used first to group the datasets which further used by regression algorithm to derive the predicted output.

Polyoxymethylene (POM), Polytetrafluoroethylene (PTFE), Polyether ether ketone (PEEK) and multiwall carbon nano tubes (MWCNT) reinforced PEEK (PEEK/MWCNT) are significant polymeric materials used in industrial applications and house-hold items and hence they were considered for this research. Firstly, design of experiments (DoE) for each material was prepared, which otherwise known as L_{27} design matrix. The design matrix had Speed (V_c), feed (f), depth of cut (ap) as input parameters and surface roughness (R_a) as response parameter (output). Cylindrical rods of 10 mm in diameter and 500 mm in length (purchased from the local supplier in Malaysia) were turned using CNC turning centre (Model: Sprint 16TC Fanuc 0i T Mate Model C) according to the design matrices. The experiments were carried out with servo super cut coolant 32t as per advice from the supplier. High carbon cemented carbide tool (Grade CNMG 120408 QM) was used in all experiments. The specification of the tool was Rhombie shape, insert angle=80°, tolerance=±0.13, insert size=12mm, insert thickness=4.76mm, insert clearance=0º. Surface roughness of the machined sample was instantaneously measured with Mitutoyo make surf tester. Three trials were done in each setting and the average of surface roughness is recorded. In total, 27 datasets for each of four materials were acquired from the experiments. All these experimental results were stored in a separate .csv file.

Model development was started with individual material, and then to unified model referring to all materials. To overcome the challenge of inaccuracy, and to improve the accuracy of prediction models, a strategy known as output discretization was adopted and hence this model is called meta based model. This strategy is to classify output in different levels and then use them in regression to derive the output response. As for as this problem is concerned, data cleaning was not required as we had a small number of datasets in each material, and no null value was associated in the datasets. Quantile distribution (0-100) was applied to each dataset and splitted the data into the proper number of groups based on surface roughness(response variable). Based on the output from quantile distribution, the datasets of the respective material was divided into two groups. For example, Delrin material datasets were grouped into two:14 datasets ($\mu<=1.590$) and 13 datasets ($\mu>1.590$ and $\mu<=1.88$). The choose of the datasets was not random, it was done by pareto principle. After this, correlation between inputs at outputs was done using Analysis of Variance

(ANOVA). This analysis was done to check if the inputs were correctly selected, and to check the inputs had influence on the output. It was found from ANOVA that feed is the most significant parameter on surface roughness.

After conducting the basis of feature extraction and correlation analysis, model development was started as discussed below. Developed meta based model has two main parts: classifier and regressor. Classifier is the one initially taking the experimental data and classify them into reasonable and more accurate groups. Regressor use the output from the classifier and predict the surface finish. For classification, XGBoost algorithm and Logistic regression algorithm were investigated. The training was started with 80% of datasets. Cross validation method (k-fold method) was then adopted to validate the model. Grid searching method was used further to tune the hyper parameters for each algorithm. The goodness or accuracy of the model was assessed by determining the largest sum of square (R^2) and the smallest Mean Squared Error (MSE). During the training process, the performance of each model was monitored continuously and recorded with numerical results quantifying hypothesized relationships between variables. The prediction model with the lowest RMSE error was considered the best model. It was found from these results that Logistic Regression model is the better to be used as classifier as it is very flexible to predict probabilities of an observation in each class. Once classifier model was confirmed, the output of the classifier was added to the database as a new feature.

Now, with four independent features (including output of classifier), further model development was continued. Support vector regressor (SVR) and XGBoost algorithm were used to complete classifier-regressor model (meta based model). Cross validation method and grid search method were also adopted, and hyper parameters were tuned iteratively till the performance of the model is in the acceptable range. After checking the performance metrics of each model, the best meta-based model was chosen. It is concluded from training and testing that, XGB model was the better model for each material. In this way, meta based model for each material was developed separately.

After investigating the model development for each material, a unified meta based ML model (a model for all four polymeric materials) was developed following the same procedure as discussed in previous paragraphs. It used speed, feed, depth of cut, material for classifier model. Interestingly from the investigations that XGB model is the best model working great in both classification and regression. It resulted almost 100% accuracy in training and 98.86% in testing. After confirming the best model, a group of predicted results was generated from the prediction model and validated experimentally. The deviation from predicted results and experimental results were checked and found very negligible. In this way, the unified meta based model was developed and validated.

Finally, Application program interface (API) was developed with the final unified meta based model which industry can use in its production line for achieving high productivity.

6.2 Recommendation for Future Work

The efficacy of the developed unified model is almost 100% and hence it can be used in industries. This model would help the industries to get higher productivity and profit. Indeed, there are many more thermosetting plastics and thermoplastics in the usage. To name a few, Polyurethane (PU), Epoxy, Polyethylene (PE), Polyester, Polypropylene, and Nylon are very commonly used plastics. Many more reinforced polymer composites are also available. For example, carbon fiber reinforced plastics, glass fiber reinforced plastics are highly used in automotive industries, airbuses and space applications. These polymer composites are very competent materials that serve equal to metals. Hence using of these polymer composites give almost equal mechanical properties with low density. Polymer research is vibrant these days because of the importance of polymers in bio medical applications too.

As this model has utilized only four different polymeric materials, this unified model cannot be used for predicting the surface finish of other polymeric parts. Hence, the scope of this research must be further extended to many other significant polymeric materials. When new materials are added, the learning and testing must be done again, though methodology is same as used in this research.

Besides, many other ML algorithms can also be attempted and evaluated. For example, deep learning can be attempted as it uses multi neural network layers in the model. The accuracy of the model may be improved in this way. Also, user interface (UI) may be deployed to mobile devices and cloud environment, so that operators can access API through their mobile phone or remotely.

REFERENCES

A. Chabbi,, M. A. Yallese, M. Nouioua, I. Meddour, T. Mabrouki, and François Girardin, (2017). Modeling and optimization of turning process parameters during the cutting of polymer (POM C) based on RSM, ANN, and DF methods, Int. J. Adv. Manuf. Technol., vol. 91, no. 5-8, pp. 2267-2290.

Aggogeri, F.; Pellegrini, N.; Tagliani, F.L. Recent Advances on Machine Learning Applications in Machining Processes. Appl. Sci. 2021, 11, 8764. https://doi.org/10.3390/app11188764

Ahmedbahaaaldin Ibrahem Ahmed Osman, Ali Najah Ahmed, Ming Fai Chow, Yuk Feng Huang, Ahmed El-Shafie, Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia, Ain Shams Engineering Journal,Volume 12, Issue 2,2021,Pages 1545-1556.

Ahmed Yaseer, Heping Chen, Machine learning based layer roughness modeling in robotic additive manufacturing, Journal of Manufacturing Processes, Volume 70, 2021,Pages 543-552

Alajmi, M.S.; Almeshal, A.M. Prediction and optimization of surface roughness in a turning process using the ANFIS-QPSO method. Materials 2020, 13, 2986.

Andreas Schumacher, Selim Erol, Wilfried Sihn, A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises, Procedia CIRP, Volume 52, 2016, Pages 161-166

B. Lela & D. Bajić & S. Jozić, (2008). Regression analysis, support vector machines, and Bayesian neural network approaches to modeling surface roughness in face milling," International Journal Advanced Manufacturing Technology, 42:1082–1088.

Baicun Wang, Fei Tao, Xudong Fang, Chao Liu, Yufei Liu, Theodor Freiheit,Smart Manufacturing and Intelligent Manufacturing: A Comparative Review, Engineering, Volume 7, Issue 6, 2021, Pages 738-757.

Benkedjouh, T., Medjaher, K., Zerhouni, N. et al. Health assessment and life prediction of cutting tools based on support vector regression. J Intell Manuf 26, 213–223 (2015). https://doi.org/10.1007/s10845-013-0774-6

Bustillo, A., Pimenov, D.Y., Mia, M. et al. Machine-learning for automatic prediction of flatness deviation considering the wear of the face mill teeth. J Intell Manuf 32, 895–912 (2021). https://doi.org/10.1007/s10845-020-01645-3

Çaydaş, Ulaş, Ekici, Sami. (2012). Support vector machines models for surface roughness prediction in CNC turning of AISI 304 austenitic stainless steel. Journal of Intelligent Manufacturing vol. 21. 1-12. 10.1007/s10845-010-0415-2.

Cherukuri, Harish, Perez-Bernabeu, Elena, Selles, Miguel, Schmitz (2019), Machining Chatter Prediction Using a Data Learning Model. Journal of Manufacturing and Materials Processing. 3. 45. 10.3390/jmmp3020045.

D. Lazarevića, M. Madića, P. Jankovića, and A. Lazarevićb, (2012). Cutting parameters optimization for surface roughness in turning operation of Polyethylene (PE) using Taguchi method, Tribol. Ind., vol. 34, no. 2, pp. 68-73.

Dong-Hyeon Kim, Thomas J. Y. Kim, Xinlin Wang, Mincheol Kim, Ying-Jun Quan, Jin Woo Oh, Soo-Hong Min, Hyungjung Kim, Binayak Bhandari, Insoon Yang, Sung-Hoon Ahn, (2018). Smart Machining Process Using Machine Learning: A Review and Perspective on Machining Industry, International Journal of Precision Engineering and Manufacturing-Green Technology, Vol. 5, No. 4, pp. 555-568

Duo, A., Basagoiti, R., Arrazola, P.J. et al. The capacity of statistical features extracted from multiple signals to predict tool wear in the drilling process. Int J Adv Manuf Technol 102, 2133–2146 (2019). https://doi.org/10.1007/s00170-019-03300-5

E. Natarajan, V. Kaviarasan, W. H. Lim, S. S. Tiang and T. H. Tan, (2018). Enhanced Multi-Objective Teaching-Learning-Based Optimization for Machining of Delrin, in IEEE Access, vol. 6, pp. 51528-51546. doi: 10.1109/ACCESS.2018.2869040

Elango Natarajan., Varadaraju Kaviarasan., Wei Hong Lim., Sew Sun Tiang., S. Parasuraman., Sangeetha Elango (2020). Non-dominated sorting modified teaching–learning-based optimization for multi-objective machining of polytetrafluoroethylene (PTFE), Journal of Intelligent Manufacturing, DOI:https://doi.org/10.1007/s10845-019-01486-9

Elango, S., Gnanamuthu, E.M.A., Durairaj, R. et al. Machine learning model for discrete, nonlinear datasets from machining operation: an industrial need to have high quality. Multiscale and Multidiscip. Model. Exp. and Des. (2022). https://doi.org/10.1007/s41939-022-00118-9

Elango, S., Varadaraju, K., Natarajan, E., Abraham Gnanamuthu, E. M., Durairaj, R., & Mariappan, P. (2022). PTFE in wet and dry drilling: Two-tier modeling and optimization through ANFIS. Mathematical Problems in Engineering, 2022 doi:10.1155/2022/4812470

Grzenda, M.; Bustillo, A. Semi-supervised roughness prediction with partly unlabeled vibration data streams. J. Intell. Manuf. 2019, 30, 933–945.

Gülçür, M., Whiteside, B. A study of micromanufacturing process fingerprints in micro-injection moulding for machine learning and Industry 4.0 applications. Int J Adv Manuf Technol 115, 1943–1954 (2021). https://doi.org/10.1007/s00170-021-07252-7

Gupta, A.K.; Guntuku, S.C.; Desu, R.K.; Balu, A. (2015). Optimisation of turning parameters by integrating genetic algorithm with support vector regression and artificial neural networks. Int. J. Adv. Manuf. Technol., 77, 331–339

Gupta, Amit Kumar, (2010) Predictive modelling of turning operations using response surface methodology, artificial neural networks and support vector regression, International Journal of Production Research, vol.48: 3, 763 — 778, DOI: 10.1080/00207540802452132

Harald Bauer, Felix Brandl, Christopher Lock, Gunther Reinhart, Integration of Industrie 4.0 in Lean Manufacturing Learning Factories, Procedia Manufacturing, Volume 23, 2018, Pages 147-152

Hui, Yang, Mei, Xuesong, Jiang, Gedong, Tao, Tao, Pei, Changyu, Ma, Ziwei. (2019). Milling Tool Wear State Recognition by Vibration Signal Using a Stacked Generalization Ensemble Model. Shock and Vibration. pg. 1-16. 10.1155/2019/7386523.

J. Abdul Shukor, S. Said, R. Harun, S. Husin, and Ab. Kadir, (2016). Optimising of machining parameters of plastic material using Taguchi method, Adv. Mater. Process Technol., vol. 2, no. 1, pp. 50-56.

J.S. Pang, M.N.M. Ansari, Omar S. Zaroog, Moaz H. Ali, S.M. Sapuan, (2014)Taguchi design optimization of machining parameters on the CNC end milling process of halloysite nanotube with aluminium reinforced epoxy matrix (HNT/Al/Ep) hybrid composite, HBRC Journal, Volume 10, Issue 2,Pages 138-144, ISSN 1687-4048, https://doi.org/10.1016/j.hbrcj.2013.09.007.

Jia, S.; Wang, S.; Lv, J.; Cai, W.; Zha ng, N.; Zhang, Z.; Bai, S. Multi-Objective Optimization of CNC Turning Process Parameters Considering Transient-Steady State Energy Consumption. Sustainability 2021, 13, 13803. https://doi.org/10.3390/su132413803

Kai Guo, Zhenze Yang, Chi-Hua Yu, Markus J. Buehler, Artificial intelligence and machine learning in design of mechanical materials, Mater. Horiz., 2021,8, 1153-1172.

Karthik, Rao M. C., Rashmi L. Malghan, Fuat Kara, Arunkumar Shettigar, Shrikantha S. Rao, and Mervin A. Herbert. 2021. "Influence of Support Vector Regression (SVR) on Cryogenic Face Milling" edited by F.

Lusquiños. Advances in Materials Science and Engineering 2021:1–18. doi: 10.1155/2021/9984369

Kaviarasan V, Venkatesan R, Natarajan E. Prediction of surface quality and optimization of process parameters in drilling of Delrin using neural network. *Progress in Rubber, Plastics and Recycling Technology*. 2019;35(3):149-169. doi:10.1177/1477760619855078

Kayaroganam, P., Krishnan, V., Natarajan, E., Natarajan, S., & Muthusamy, K. (2021). Drilling parameters analysis on in-situ al/b4 c/mica hybrid composite and an integrated optimization approach using fuzzy model and non-dominated sorting genetic algorithm. Metals, 11(12) doi:10.3390/met11122060

Kilickap, E.; Yardimeden, A.; Çelik, Y.H. Mathematical Modelling and Optimization of Cutting Force, Tool Wear and Surface Roughness by Using Artificial Neural Network and Response Surface Methodology in Milling of Ti-6242S. Appl. Sci. 2017, 7, 1064. https://doi.org/10.3390/app7101064

Kim, DH., Kim, T.J.Y., Wang, X. et al. Smart Machining Process Using Machine Learning: A Review and Perspective on Machining Industry. Int. J. of Precis. Eng. and Manuf.-Green Tech. 5, 555–568 (2018). https://doi.org/10.1007/s40684-018-0057-y

Koon Meng Ang, Elango Natarajan, Nor Ashidi Mat Isa, Abhishek Sharma, Hameedur Rahman, Richie Yi Shiun Then, Moath Alrifaey, Sew Sun Tiang, Wei Hong Lim, Modified teaching-learning-based optimization and applications in multi-response machining processes, Computers & Industrial Engineering, Volume 174,2022,108719

Krishnakumar, P, Krishnaswamy, Rameshkumar, K I, Ramachandran (2015). Tool Wear Condition Prediction Using Vibration Signals in High Speed Machining (HSM) of Titanium (Ti-6Al-4 V) Alloy. Procedia Computer Science, 50. 10.1016/j.procs.2015.04.049.

Kuntoğlu, M.; Acar, O.; Gupta, M.K.; Sağlam, H.; Sarikaya, M.; Giasin, K.; Pimenov, D.Y. Parametric Optimization for Cutting Forces and Material Removal Rate in the Turning of AISI 5140. Machines 2021, 9, 90. https://doi.org/10.3390/machines9050090

Laddada S, Si-Chaib MedO, Benkedjouh T, Drai R. Tool wear condition monitoring based on wavelet transform and improved extreme learning machine. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2020;234(5):1057-1068. doi:10.1177/0954406219888544

Lei, Z., Zhou, Y., Sun, B. et al. An intrinsic timescale decomposition-based kernel extreme learning machine method to detect tool wear conditions in the milling process. Int J Adv Manuf Technol 106, 1203–1212 (2020). https://doi.org/10.1007/s00170-019-04689-9

Lin, W.-J.; Lo, S.-H.; Young, H.-T.; Hung, C.-L. Evaluation of Deep Learning Neural Networks for Surface Roughness Prediction Using Vibration Signal Analysis. Appl. Sci. 2019, 9, 1462.

Liu H, Lin H, Mao X, Jiang X, Liu Q, Li B. Surface roughness optimal estimation for disc parts turning using Gaussian-process-based Bayesian combined model. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2019;233(11):4032-4046. doi:10.1177/0954406218809116

Lu, X., Hu, X., Wang, H., Si, L., Liu, Y., and Gao, L., (2016). Research on the Prediction Model of Micro-Milling Surface Roughness of Inconel718 Based on SVM, Industrial Lubrication and Tribology, Vol. 68, No. 2, pp. 206-211.

Lu, Z.; Wang, M.; Dai, W. Machined Surface Quality Monitoring Using a Wireless Sensory Tool Holder in the Machining Process. Sensors 2019, 19, 1847.

M. Kaddeche, K. Chaoui, and M. A. Yallese, (2012). Cutting parameters effects on the machining of two high density polyethylene pipes resins, Mech. Ind., vol. 13, no. 14, pp. 307-316.

M. R. Panda, S. K. Biswal, and Y. K. Sharma, (2016). Experimental analysis on the effect of process parameters during CNC turning on nylon-6/6 using tungsten carbide tool, Int. J. Eng. Sci. & Res. Technol. vol. 5, no. 4, pp. 79-84.

Markus Brillinger, Marcel Wuwer, Muaaz Abdul Hadi, Franz Haas, Energy prediction for CNC machining with machine learning, CIRP Journal of Manufacturing Science and Technology, 2021, Volume 35, Pages 715-723

Mirifar, S.; Kadivar, M.; Azarhoushang, B. First Steps through Intelligent Grinding Using Machine Learning via Integrated Acoustic Emission Sensors. J. Manuf. Mater. Process. 2020, 4, 35.

Montomoli J, Romeo L, Moccia S, Bernardini M, Migliorelli L, Berardini D, Donati A, Carsetti A, Bocci MG, Wendel Garcia PD, Fumeaux T, Guerci P, Schüpbach RA, Ince C, Frontoni E, Hilty MP; RISC-19-ICU Investigators. Machine learning using the extreme gradient boosting (XGBoost) algorithm predicts 5-day delta of SOFA score at ICU admission in COVID-19 patients. J Intensive Med. 2021 Oct 22;1(2):110-116. doi: 10.1016/j.jointm.2021.09.002. PMID: 36785563; PMCID: PMC8531027.

N. Hamlaoui, S. Azzouz, K. Chaoui, Z. Azari, and M. A. Yallese. (2017). Machining of tough polyethylene pipe material: surface roughness and cutting temperature optimization, Int. J. Adv. Manuf. Technol., vol. 92, no. 5-8, pp. 2231-2245.

N. K. Vuong, Y. Xue, S. Liu, Y. Zhou and M. Wu, "Predicting Surface Roughness and Flank Wear in Turning Processes," 2020 IEEE International Conference on Prognostics and Health Management (ICPHM), 2020, pp. 1-6, doi: 10.1109/ICPHM49022.2020.9187056.

Nasir, V., Sassani, F. A review on deep learning in machining and tool monitoring: methods, opportunities, and challenges. Int J Adv Manuf Technol 115, 2683–2709 (2021). https://doi.org/10.1007/s00170-021-07325-7

Natarajan, E.; Markandan, K.; Sekar, S.M.; Varadaraju, K.; Nesappan, S.; Albert Selvaraj, A.D.; Lim, W.H.; Franz, G. Drilling-Induced Damages in Hybrid Carbon and Glass Fiber-Reinforced Composite Laminate and Optimized Drilling Parameters. J. Compos. Sci. 2022, 6, 310. https://doi.org/10.3390/jcs6100310

Ortt, R., Stolwijk, C. and Punter, M. (2020), "Implementing Industry 4.0: assessing the current state", Journal of Manufacturing Technology Management, Vol. 31 No. 5, pp. 825-836. https://doi.org/10.1108/JMTM-07-2020-0284

Palanikumar, K., Nithyanandam, J., Natarajan, E., Hong Lim, W., & Sun Tiang, S. (2022). Mitigated cutting force and surface roughness in titanium alloy-multiple effective guided chaotic multi objective teaching learning based optimization. Alexandria Engineering Journal, doi:10.1016/j.aej.2022.09.029

Pan, Y.; Kang, R.; Dong, Z.; Du, W.; Yin, S.; Bao, Y. On-line prediction of ultrasonic elliptical vibration cutting surface roughness of tungsten heavy alloy based on deep learning. J. Intell. Manuf. 2020.

Paweł Twardowski, Maciej Tabaszewski, Martyna Wiciak – Pikuła, Agata Felusiak-Czyryca, Identification of tool wear using acoustic emission signal and machine learning methods, Precision Engineering, Volume 72, 2021,Pages 738-744

Peng Wang, Ziye Liu, Robert X. Gao, Yuebin Guo, Heterogeneous data-driven hybrid machine learning for tool condition prognosis, CIRP Annals, Volume 68, Issue 1, 2019, Pages 455-458

Ruholla Jafari-Marandi, Mojtaba Khanzadeh, Wenmeng Tian, Brian Smith, Linkan Bian,From in-situ monitoring toward high-throughput process control: cost-driven decision-making framework for laser-based additive manufacturing, Journal of Manufacturing Systems,Volume 51,2019,Pages 29-41

S, R., Palanikumar, K., Boppana, S. B., & Natarajan, E. (2022). Analysis of chip formation and temperature measurement in machining of titanium alloy (ti-6Al-4V). Experimental Techniques, doi:10.1007/s40799-021-00537-2

Sara Karam, Piera Centobelli, Doriana M. D'Addona, Roberto Teti, Online Prediction of Cutting Tool Life in Turning via Cognitive Decision Making, Procedia CIRP,Volume 41,2016,Pages 927-932

Sebastian Schorr, Matthias Möller, Jörg Heib, Dirk Bähre, In-process Quality Control of Drilled and Reamed Bores using NC-Internal Signals and Machine Learning Method, Procedia CIRP, Volume 93, 2020, Pages 1328-1333

Shah, M., Vakharia, V., Chaudhari, R. et al. Tool wear prediction in face milling of stainless steel using singular generative adversarial network and LSTM deep learning models. Int J Adv Manuf Technol 121, 723–736 (2022). https://doi.org/10.1007/s00170-022-09356-0

Sri Kolla, Meysam Minufekr, Peter Plapper, Deriving essential components of lean and industry 4.0 assessment model for manufacturing SMEs, Procedia CIRP, Volume 81, 2019, Pages 753-758

Su, W., Jiang, F., Shi, C. *et al.* An XGBoost-Based Knowledge Tracing Model. *Int J Comput Intell Syst* 16, 13 (2023). https://doi.org/10.1007/s44196-023-00192-y

Tabaszewski M, Twardowski P, Wiciak-Pikuła M, Znojkiewicz N, Felusiak-Czyryca A, Czyżycki J. Machine Learning Approaches for Monitoring of Tool Wear during Grey Cast-Iron Turning. Materials (Basel). 2022 Jun 20;15(12):4359. doi: 10.3390/ma15124359. PMID: 35744419; PMCID: PMC9230642.

Ting Hu and Ting Song (2019) Research on XGboost academic forecasting and analysis modelling, Journal of Physics: Conference Series, Volume 1324, DOI 10.1088/1742-6596/1324/1/012091

Tobias Wagner, Christoph Herrmann, Sebastian Thiede, Industry 4.0 Impacts on Lean Production Systems,Procedia CIRP, Volume 63, 2017, Pages 125-131

Tommaso Gallo, Chiara Cagnetti, Cecilia Silvestri, Alessandro Ruggieri, Industry 4.0 tools in lean production: A systematic literature review, Procedia Computer Science, Volume 180, 2021,Pages 394-403

Tran, MQ., Liu, MK. & Tran, QV. Milling chatter detection using scalogram and deep convolutional neural network. Int J Adv Manuf

Technol 107, 1505–1516 (2020). https://doi.org/10.1007/s00170-019-04807-7

Van-Hai Nguyen ,Tien-Thinh Le , Hoanh-Son Truong,Minh Vuong Le, Van-Luc Ngo,Anh Tuan Nguyen , and Huu Quang Nguyen, Applying Bayesian Optimization for Machine Learning Models in Predicting the Surface Roughness in Single-Point Diamond Turning Polycarbonate, Mathematical Problems in Engineering, Volume 2021, https://doi.org/10.1155/2021/6815802

Xin-Cheng Cao, Bin-Qiang Chen, Bin Yao, Wang-Peng He, Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification, Computers in Industry, Volume 106, 2019, Pages 71-84

Yiğit M. Arisoy & Tuğrul Özel (2015) Machine Learning Based Predictive Modeling of Machining Induced Microhardness and Grain Size in Ti–6Al–4V Alloy, Materials and Manufacturing Processes, 30:4, 425-433, DOI: 10.1080/10426914.2014.961476

Yuqing Zhou, Bintao Sun, Weifang Sun, A tool condition monitoring method based on two-layer angle kernel extreme learning machine and binary differential evolution for milling, Measurement, Volume 166,2020

Žarko Ćojbašić, Dalibor Petković, Shahaboddin Shamshirband, Chong Wen Tong, Sudheer Ch, Predrag Janković, Nedeljko Dučić, Jelena Baralić, Surface roughness prediction by extreme learning machine constructed with abrasive water jet, Precision Engineering, Volume 43,2016,Pages 86-92

Zhang P, Jia Y, Shang Y. Research and application of XGBoost in imbalanced data. International Journal of Distributed Sensor Networks. 2022;18(6). DOI:10.1177/15501329221106935

Zurkovic, Z., Cukor, G., Brezocnik, M. et al. (2018). A comparison of machine learning methods for cutting parameters prediction in highspeed turning process", J Intell Manuf 29, 1683–1693. https://doi.org/10.1007/s10845-016-1206-1

APPENDIX – A

Publications

- Elango, S., Gnanamuthu, E.M.A., Durairaj, R. et al. (2022). Machine learning model for discrete, nonlinear datasets from machining operation: an industrial need to have high quality. Multiscale and Multidiscipline Modeling, Experiments and Design. 5, 279–291 https://doi.org/10.1007/s41939-022-00118-9 (Scopus indexed)
- Sangeetha Elango at al. (2022). Extreme Gradient Boosting Regressor Solution for Defy in Drilling of Materials, Advances in Materials Science and Engineering, Volume 2022, Article ID 8330144 https://doi.org/10.1155/2022/8330144 (WoS and Scopus indexed; Q2 in SCR)
- Sangeetha Elango at al. (2022). PTFE in Wet and Dry Drilling: Two-Tier Modeling and Optimization through ANFIS, Mathematical Problems in Engineering, Volume 2022, Article ID 4812470, https://doi.org/10.1155/2022/4812470 (WoS and Scopus indexed; Q2 in SCR)