BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF MECHANICAL ENGINEERING
FAKULTA STROJNÍHO INŽENÝRSTVÍ

INSTITUTE OF MATERIALS SCIENCE AND ENGINEERING
ÚSTAV MATERIÁLOVÝCH VĚD A INŽENÝRSTVÍ

WEAR MODEL OF MATERIAL COUPLE AT HIGH TEMPERATURES
MODEL OPOTŘEBENÍ MATERIÁLOVÉ DVOJICE PRO VYSOKÉ TEPLOTY

MASTER'S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR   Bc. David Rychlý
AUTOR PRÁCE

SUPERVISOR   Ing. Karel Němec, Ph.D.
VEDOUCÍ PRÁCE

BRNO 2024

Assignment Master's Thesis

Institut:	Institute of Materials Science and Engineering
Student:	**Bc. David Rychlý**
Degree programm:	Materials Engineering
Branch:	no specialisation
Supervisor:	**Ing. Karel Němec, Ph.D.**
Academic year:	2023/24

As provided for by the Act No. 111/98 Coll. on higher education institutions and the BUT Study and Examination Regulations, the director of the Institute hereby assigns the following topic of Master's Thesis:

Wear model of material couple at high temperatures

Brief Description:

When regulating the turbocharger using the control mechanism, unlubricated contact of the kinematic components occurs, primarily at high temperatures. Various wear models can be employed to predict the wear rate of individual parts in contact during the work cycle. Nevertheless, these models require calibration based on the results of experiments conducted on test specimens during tribological tests. Experimental testing of material couples can be replicated using a multibody simulation of the tribological test bench, incorporating an applied wear model. This work assumes a basic understanding of operating in the GT–Suite environment. The wear model should accurately represent the tribological behavior of the friction couple of materials, aligning with experimental measurements under different operating conditions. The utilization of the wear model is expected in multibody systems when simulating the working cycle of the control kinematic mechanism. The model must be designed to be directly applicable in multibody systems.

Master's Thesis goals:

Analysis of existing computational wear models suitable for solving the given problem.

Design of a suitable wear model and an overview of the necessary empirical parameters of the model for use in a multibody system.

Calibration of the parameters of the selected model based on wear measurements of samples from the tribological test station.

Application of a wear model for a selected material couple in a contact used in turbocharger regulation.

Correlation of wear increase between multibody model simulation and tribological test bench.

Recommended bibliography:

Goryacheva, Irina Georgievna. Contact mechanics in tribology. Solid mechanics and its applications, vol. 61. Dordrecht: Kluwer, [1998]. ISBN 978-90-481-5102-8.

MILANESE, Enrico. Surface roughness evolution in adhesive wear processes. EPFL, 2020.

MENG, H. C.; LUDEMA, K. C. Wear models and predictive equations: their form and content. Wear, 1995, 181: 443-457.

YANG, L. J. An integrated transient and steady-state adhesive wear model. Tribology transactions, 2003, 46.3: 369-375.

Deadline for submission Master's Thesis is given by the Schedule of the Academic year 2023/24

In Brno,

L. S.

_____

prof. Ing. Ivo Dlouhý, CSc.
Director of the Institute

_____

doc. Ing. Jiří Hlinka, Ph.D.
FME dean

Abstract

The main goal of this thesis is to create a model for predicting the wear on material couple during high temperatures. This model should serve for prediction of wear behavior on kinematic mechanism on turbocharger. Predicting wear is difficult because of the very complex nature of processes involving physical, chemical, and mechanical phenomena, making it hard to predict. This thesis attempts to solve this problem using empirical relationships from literature and methods of machine learning using experimental data from Garrett Motion company. Both empirical relationships and machine learning approaches for predicting wear in tribology are thoroughly researched. The thesis critically evaluates the use of I.G.Goryacheva equation chosen as the best equation for adhesive wear, chosen as the primary mechanism of wear. Data from Garrett Motion serve as the base for machine learning algorithms. The best model is proposed – Support Vector Regression – and the best architecture is presented. For the best architecture, the thesis also includes feature importance analysis and search for optimal test-train split for dataset. Finally, the thesis includes recommendations for the next research and ways to improve the process of modelling the wear for this use case.

Abstrakt

Hlavním cílem práce bylo vytvořit model opotřebení materiálové dvojice za vysokých teplot pro predikci opotřebení kinematického mechanismu na turbodmychadle. Predikce opotřebení je komplexní proces, kterého se účastní mnoho chemických, fyzikálních a mechanických jevů, a proto není jednoduché jej analyticky počítat. Tato práce se vydává naproti tomuto problému pomocí empirických vztahů z literatury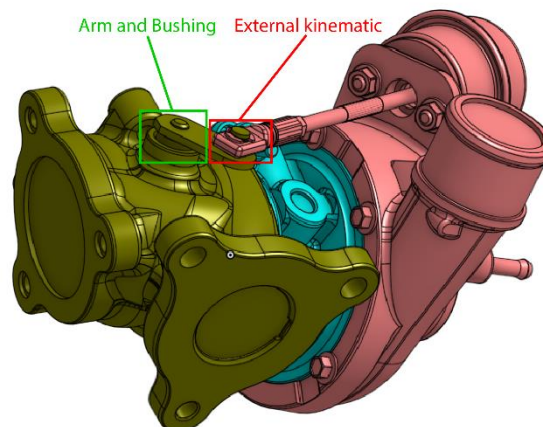 a aplikace principů a algoritmů strojového učení (machine learning) postaveného na datech z experimentu. Empirické vztahy i metody strojového učení jsou v práci podrobeny důkladné rešerši a aplikovány na dostupná data z firmy Garrett Motion. Práce kriticky hodnotí vhodnost vybraného empirického vztahu I.G.Goryachevy pro adhezivní opotřebení jako primárního mechanismu opotřebení a variuje koeficienty pro výzkum využití rovnice. Na data jsou aplikovány algoritmy strojového učení a je prezentována architektura nejlepšího modelu – Support Vector Regression – spolu s metrikami, analýzou důležitosti vlastností (feature importance) a optimálním rozdělení datasetu pro trénink a testování. Závěrem práce doporučuje směry dalšího výzkumu a způsoby, jak zlepšit proces modelování pro tento konkrétní problém.

Keywords

Tribology, Adhesive Wear, Empirical Wear Law, Goryacheva, Machine Learning, Regression

Klíčová slova

Tribologie, Adhezivní opotřebení, Empirické vztahy opotřebení, Goryacheva, Strojové učení, Regrese

Rozšířený abstrakt

Hlavním cílem práce bylo vytvořit model opotřebení materiálové dvojice za vysokých teplot pro predikci opotřebení kinematického mechanismu na turbodmychadle, viz obrázek 1. Predikce opotřebení je komplexní proces, kterého se účastní mnoho chemických, fyzikálních a mechanických jevů, a proto není jednoduché opotřebení analyticky počítat. Tato práce se vydává naproti tomuto problému pomocí empirických vztahů z literatury a aplikace principů a algoritmů strojového učení

Fig. 1 – Kinematické mechanismy na turbodmychadle

(machine learning) postaveného na datech z experimentu. Empirické vztahy i metody strojového učení jsou v práci podrobeny důkladné rešerši a aplikovány na dostupná data z firmy Garrett Motion. Práce kriticky hodnotí vhodnost vybraného empirického vztahu a jejím výstupem je nejlepší vhodný model strojového učení.

Opotřebení tvoří, spolu s třením a mazáním, hlavní pilíře vědeckého oboru tribologie. Náplní tribologie jsou zmíněné 3 pilíře a ostatní věci týkající se třecího kontaktu, tedy místa interakce dvou tuhých těles. Mohou zde nastat různé typy dějů. Fyzikální děje jako mechanické, elektrické, magnetické nebo teplo a také chemické nebo biologické děje. Tyto děje se poté manifestují v podobě třecí síly mezi tělesy. Vzhledem k tomu, že až jedna třetina energetických zdrojů na zemi je vynakládána na překonávání třecích sil, je studium těchto procesů prostřednictvím tribologie důležitou disciplínou ve strojírenství. [1][2]

Opotřebení je definováno jako progresivní ztráta materiálu z povrchu a jeho predikce je oproti tření obtížná. Jedná se o komplexní proces, jehož se zúčastní spousta fyzikálních, chemických a mechanických jevů. Nelze tak jednoznačně predikovat opotřebení a jednoznačně určit podobný koeficient opotřebení jako je koeficient tření. I přesto mnoho autorů věnovalo predikci opotřebení značné úsilí a tento obor se od svého počátku velmi posunul. [1][2]

V 40. letech 20. století začala hlavní vlna výzkumu empirické predikce opotřebení s hlavními autory jako Holm, Archard, Burwell, Strang, Rabinowicz a dalšími. Nejvýznamnějším poznatkem z této doby je Holm-Archardův vztah pro opotřebení, v rovnici (1), kde V je opotřebený objem, K je koeficient opotřebení, N je zatížení, L je vzdálenost a H je tvrdost. Koeficient opotřebení obvykle nabírá hodnot 10^{-3}-10^{-7} a je závislý na mnoha proměnných.

$$V = \frac{KNL}{H} \qquad (1)$$

Pro optimální proces modelování opotřebení je nutné zvolit primární typ opotřebení ze 7 hlavních typů opotřebení. V našem případě bylo zvoleno adhezivní opotřebení jako primární mechanismus. Adhezivní opotřebení nastává v případě tření mezi dvěma makroskopicky hladkými povrchy. Jelikož žádný povrch není dokonale hladký, tak kontakt nastává na vrcholcích nerovností. Tyto vrcholky jsou podrobeny vysokým silám na malých plochách a tíhnou k tomu lnout k sobě díky adhezním silám. Vznikají tak lokální „svary" na krátkou dobu, které se poté opět rozpojí. Díky deformačnímu zpevnění však nastávají i případy,

kdy svar je silnější než atomové síly v měkčím ze dvou materiálů. V tomto případě dochází k vytržení materiálu v nové rovině a materiál je následně vtlačen do druhého povrchu nebo uvolněn jako volná částice, která se pohybuje mezi povrchy. [19][20][21][22]

V oblasti empirických modelů adhezivního opotřebení byla provedena rozsáhlá rešerše, jejímž výstupem je nejvhodnější model opotřebení pro náš problém – vztah I.G. Goryachevy, v rovnici (2). Kde V je opotřebený objem, p je zatížení, v je rychlost, t je čas a K, α, β jsou koeficienty opotřebení. Tato rovnice bude využita pro predikci opotřebení na základě dostupných dat. [18]

$$V = Kp^{\alpha}v^{\beta}t \qquad\qquad\qquad (2)$$

Experimentální data byla poskytnuta firmou Garrett Motion, které tímto autor chce poděkovat. Pro predikci opotřebení a tření byl použit tribometr DN55, jeho princip znázorněn na obrázku 2. Vzorky v podobě jednoho „pinu" a dvou „platů" jsou vloženy do testovacího přípravku a celá komora je zahřáta na testovací teplotu. Na vzorky je aplikována přítlačná síla a pin koná cyklický pohyb nahoru a dolů. Experiment probíhá určitou dobu a následně jsou vzorky vyjmuty ven. Opotřebené vzorky jsou podrobeny analýze na laserovém profilometru a odebraný objem je spočítán. Ten je následně využit pro predikci pomocí modelů.
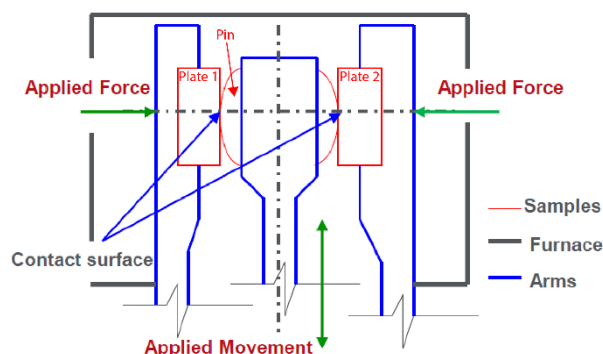

Fig. 2 - DN55 Tribometr

Nejprve je provedena průzkumná analýza dat a vizualizace a následně jsou spočítány hodnoty za různých variací koeficientů α a β pro určení koeficientu K. Spočítané výsledky jsou daleko od rozsahu koeficientu K uváděného v literatuře, tak je vyhotoven 3D graf závislosti K–α–β a následně také konvergence koeficientů α a β. Rovnice Goryacheva se ukázala býti méně vhodnou a autor práce doporučuje naměření více experimentálních hodnot a také větší variaci v nastavovaných parametrech.

Dalším využitým nástrojem pro predikci je strojové učení, v originále zvané „machine learning". Jedná se o metodu programování, která dává počítači schopnost učit se, aniž by byl specificky naprogramován. Jednoduchým příkladem je filtrování spam emailů. Bylo by velice obtížné definovat všechny jejich možné varianty. Je ovšem možné pomocí matematických algoritmů program „naučit" jak obecně vypadá, jaké má znaky, takový spam email a na základě toho program spočítá pravděpodobnost zdali je email spam nebo ne a zařadí jej do příslušné složky. Kromě klasifikačních úloh jako je spam filtr lze také pomocí regresních modelů predikovat hodnoty veličiny pomocí křivek. Jedná se o přístup založený na datech a proto je kvalita, kvantita a informační hodnota dat fundamentální pro kvalitní model. V našem případě se jedná o učení pod dohledem (supervised) – tedy máme k datům označení/výsledek. [35]

Pro naměřená data byla vypracována matice vlastností (feature matrix) v podobě veličin: materiálová dvojice, vzorek (pin/plate), teplota, poměr tvrdostí, koeficient tření. Predikovanou veličinou je opotřebený objem. Dataset sestávající se z 51 vektorů vlastností je rozdělen na trénovací a testovací set v poměru 80/20. Hrubá analýza dostupných modelů strojového učení je aplikována a jejím výsledkem jsou 4 modely, které jsou podrobeny procesu doladění (fine-tuning). Těmito modely jsou Support Vector Regression (SVR), Random Forest (RF),

Genetic Algorithm (GA) a Multilayer Perceptron Neural Network (NN). Po procesu doladění je vybrán nejlepší model – Support Vector Regression – s architekturou ukázanou v tabulce 1.

Table 1 - architektura nejlepšího SVR modelu

Architektura		Vyhodnocení	
Nejlepší model – SVR	Hodnota	Metriky	Hodnota
C – Regularizační parametr [-]	1.3	R2 trénink [-]	0.351
Epsilon – šířka pásu [-]	0.1	R2 test [-]	0.429
Kernel [-]	rbf	Střední kvadratická chyba (MSE) [um3^2]	1.856e16
Gamma – kernel koeficient [-]	auto	Odmocněná střední kvadratická chyba (RMSE) [um3]	1.362e8
Stupeň Polynomial Features [-]	1		
Rozdělení datasetu [-]	80/20		

Pro výsledný SVR model je následně provedena analýza důležitosti vlastností (feature importance analysis), která je zobrazená v obrázku 3. Nejčastější veličina pro predikci v SVR modelu byl koeficient opotřebení a poté typ vzorku a teplota. Pro SVR model je také provedena analýza optimální hodnoty rozdělení datasetu s výslednou nejlepší hodnotou 80/20.
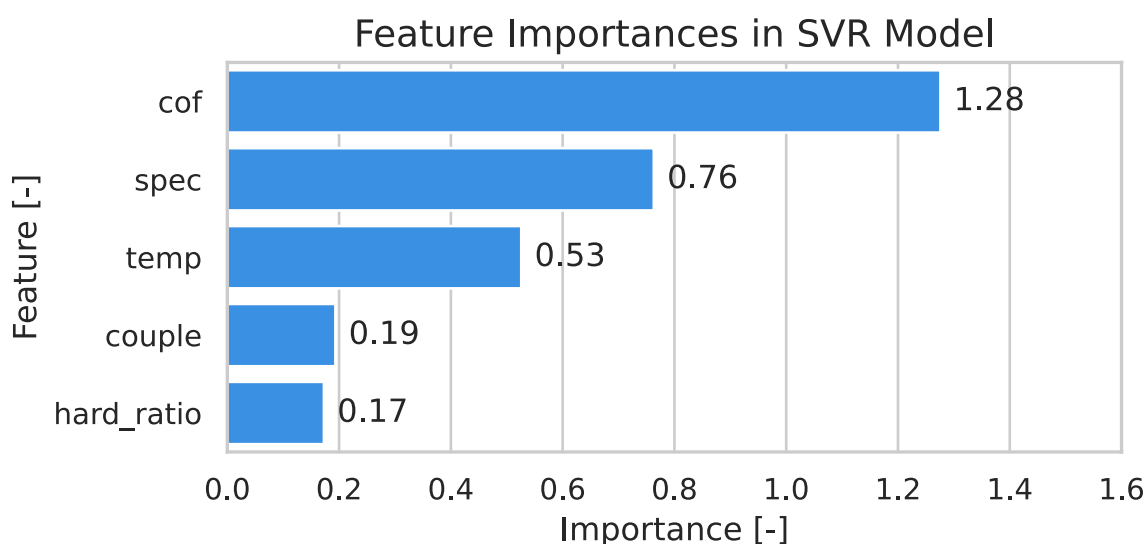

Fig. 3 – Analýza důležitosti vlastností pro nejlepší zvolený model – SVR

Pro konkrétní případ opotřebení kinematického mechanismu na turbodmychadle, experimentálně simulovaném na tribometru DN55, byla provedena predikce zvoleným empirickým modelem Goryacheva a také pomocí metod strojového učení. Empirický model Goryacheva se ukázal být méně vhodným vhledem k nedostatku informací a komplexnosti procesů opotřebení a tato skutečnost byla ověřena variací parametrů. Na základě dostupných dat bylo vyzkoušeno mnoho modelů strojového učení, vybrány 4 nejlepší a po jejich optimalizaci byl určen nejlepší model – Support Vector Regression s architekturou a výsledky prezentovanými výše. Pro zlepšení kvality modelu a predikce autor doporučuje sběr více experimentálních dat a větší variaci experimentálních parametrů. Dalším potenciálním směrem je rozšíření matice vlastností o nové veličiny z experimentu.

RYCHLÝ, David. Wear model of material couple at high temperatures [online]. Brno. 2024. Available at: https://www.vut.cz/studenti/zav-prace/detail/157310. Master Thesis. Brno University of Technology, Faculty of Mechanical Engineering, Institute of Material Science and Engineering. Supervisor Karel Němec.

RYCHLÝ, David. Model opotřebení materiálové dvojice pro vysoké teploty [online]. Brno, 2024. Dostupné z: https://www.vut.cz/studenti/zav-prace/detail/157310. Diplomová práce. Vysoké učení technické v Brně, Fakulta strojního inženýrství, Ústav materiálových věd a inženýrství. Vedoucí práce Karel Němec

I hereby declare that I wrote my Master Thesis „Wear model of material couple at high temperatures" all by myself under the direction of my Academic supervisor, Ing. Karel Němec Ph.D. and Garrett Motion supervisor Ing. Jindřich Šenekl. All used sources are listed in the used literature chapter.

In Brno 19.5.2024   David Rychlý

Table of contents

1. Introduction
1.1. Tribology

Tribology is a discipline that deals with the processes and phenomena occurring during the friction interaction of solids. The subject of the tribological study is *friction contact* – region of interaction of bodies in contact. At friction contact, processes of physical nature like mechanical, electrical, magnetic and heat can occur. Processes can also be of a chemical and biological nature. These processes manifest themselves in the form of *Friction force* i.e. resistance to the relative displacement of bodies, which form the contact. In the modern world, up to one third of the world's energy resources is being spent on overcoming friction forces. The friction as a phenomena, and its study via tribology, is therefore undeniable and very important. [1]

Another manifestation of processes occurring during the contact interaction is the wear. Wear is a progressive loss of material from surfaces due to its fracture in friction interaction showing up in gradual change of the dimensions and shape of the contacting bodies. Wear of material can change the properties and functionality of given machine parts or even lead to machine failure. Study of wear, its causes and elaboration of methods for improvement of wear resistance is therefore a very important sub-field in the discipline of tribology. Wear can be slight or extremely rapid, depending on the various variables. In most cases, wear is detrimental. It can cause increased clearances between the moving components, unwanted freedom of movement and loss of precision. This can lead to vibration, increased mechanical loading which increases wear again. Even small material loss due to wear can cause complete failure of large and complex machines. Wear can also be used for desirable purposes such as grinding and polishing. This is the case when wear is the mechanism of material removal. [1][2]

Tribology can be considered applied science because it has goals like diminishing of the energy losses, effect of friction and wear on the environment and the increase of machine life. The successful solution to these problems lies in the deep understanding of nature of friction and wear. Tribology as a discipline has evolved on the basis of mechanics, physics, chemistry and other sciences. Unfortunately, the results obtained in these individual fields cannot be applied directly because of complexity of the tribological processes. Tribological processes are complicated, interconnected and they involve multiple scales and hierarchical levels. They therefore must be considered using results of different scientific disciplines simultaneously. [1]

As mentioned previously, friction is one of the main topics of the tribology. Friction plays a central role in the performance of many mechanical systems. In some cases, low friction is desirable and essential. This is the case of operation of joints for example. Hinges on doors, human hip joints or bridge supports demand low friction forces. In other cases, the goal is to absolutely eliminate friction. This is done in applications such as bearings and gears, when lowering the friction can increase overall efficiency of those mechanical parts. Opposite to that, there are applications where controlled friction is essential to dissipate kinetic energy and transfer torque such as brakes and clutches. Another good example of use of high friction is friction between the road surface and the tire of a vehicle. [2]

The best solution for reducing the friction, and often also wear, is lubrication of the system. Therefore, studies of lubrication are very closely related to the studies of wear and friction. If we do not use artificial lubricant, we still must consider lubrication caused by the

surrounding atmosphere, especially oxygen and water vapour in the air. They often play a similar role as the lubricant. [2]

Two solid surfaces are placed together and contact between them will generally occur only in isolated parts of the nominal contact area, not the whole surface. This is later explained as a "true contact area". The reason for contact being only in some parts between the surfaces is the surface roughness of them. Even with the small value of surface roughness in well-polished surfaces, the surface is not ideally flat. True contact area is therefore independent of the surface roughness. In contact between the bodies, there are therefore localized regions of forces. These regions of forces are responsible for the friction. The difference between nominal contact area A_n and true contact area A_t is visually shown in Figure 4. [2][3]

Fig. 4 - Nominal contact area vs True contact area [4]
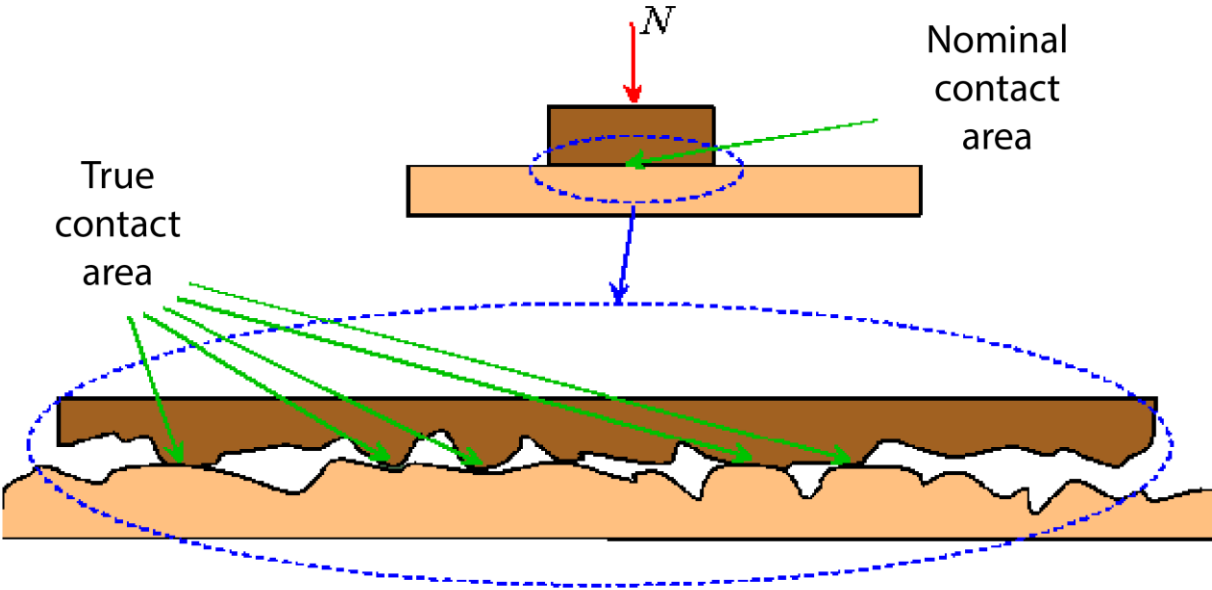
The size of the true contact area is not dependent on the size of the nominal area or surface roughness, but rather load. Burwell and Strang [5] show this schematically in Figure 5. The higher the applied load, the more asperities are in contact that make the true contact area. [5]
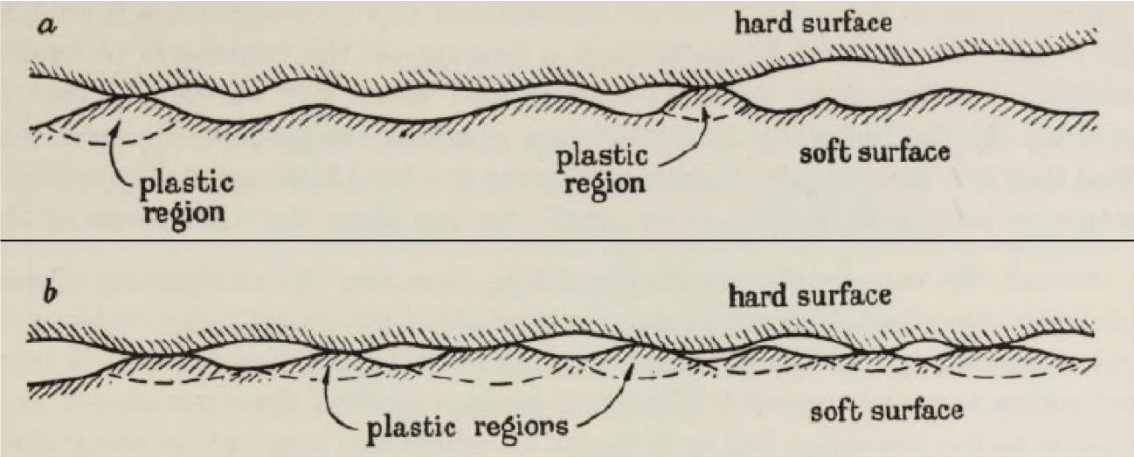

Fig. 5 - Surfaces in contact during a) low applied load b) high applied load [5]

Friction is a force and may be defined as *resistance encountered by one body in moving over another*. This definition is broad and embraces two important classes of relative motion: sliding and rolling as shown in Figure 6. It can be beneficial to make distinctions between sliding friction and rolling friction, but they are not mutually exclusive. Meaning, that even "pure rolling" nearly always involves some sliding. [2]
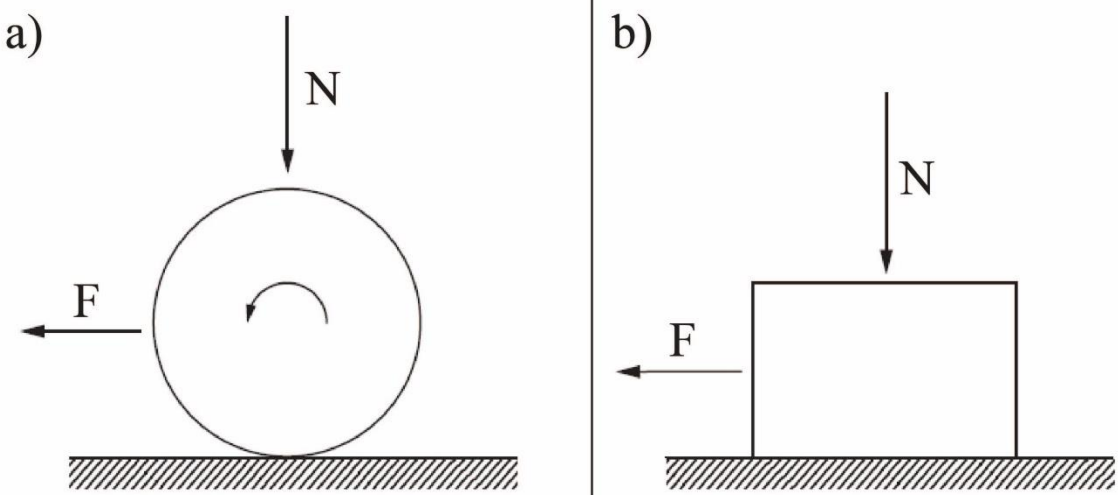

Fig. 6 - Schematically shown a) rolling and b) sliding [2]

Considering ideal rolling or ideal sliding, in both cases tangential force F is needed to move the upper body over the stationary surface. Also considering the normal load N, these two forces can be combined into ratio, which is known as *coefficient of friction*. Coefficient of friction (CoF) is usually denoted μ and is described in equation (3).

$$\mu = \frac{F}{N} \quad \rightarrow \quad F = \mu N \tag{3}$$

The magnitude of the frictional force is described by the value of the coefficient of friction. Values of the coefficient of friction can vary by multiple orders. It can be very small, such as 0.001 in lightly loaded rolling bearing, meaning the friction force is low. This is the preferred case when designing the bearings. On the other hand, it can be very big, such as 10 for two identical clean metal surfaces sliding in vacuum. In this case, the friction force is large. These are the extreme cases. For the case of sliding in air and the absence of lubricant, considering most common materials, the value of the coefficient of friction is in the range from 0.1 to 1. It is also important to mention, that coefficient of friction μ is not a fundamental property of pair of surface, because strong frictional forces can be experienced even without a normal load if the surfaces are clean and have an intrinsic adhesive capability. [2][6]

When a tangential traction is applied to a system already in a state of plastic contact, the plasticity theory predicts that the junction area will grow as the two surfaces slide against each other. Complete welding of surfaces is usually prevented by the controlling influence exerted by the interfacial contaminating layers. Even a small degree of contamination can reduce the shear strength of the interface sufficiently to discourage continuous growth of the bonded area. This means that coefficients of friction μ tend to remain finite.

Laws of friction

Over the course of years, few laws of friction were presented. Two empirical Laws of Sliding Friction are named after the person who stated them in 1699 – Guillaume Amonton. However, the first mentions and description of the laws date back to 1493, more than 200 years earlier, by the famous Leonardo da Vinci. In addition to the first two empirical Amontons laws, the third law is sometimes mentioned, which is often attributed to Coulomb in 1785. [2][7][8]

The laws of friction may be stated as follows:

1. The friction force is proportional to the normal force (equation 1) (Amontons' first law)
2. The friction force is independent of the nominal contact area (Amontons' second law)
3. The friction force is independent of the sliding velocity (Coulomb's law)

The laws have varying reliability and must be taken with caution, but they provide good and useful summaries of empirical observations. The first law, described by equation (1) is obeyed by most metals and many other materials. Exceptions are usually seen in polymers and materials with a very low elastic modulus. The second law is nowadays also well attested for most materials with, again, the exception of polymers. The third law is less well founded and based on common observation, that the frictional force to initiate sliding is usually greater than the force to maintain the sliding and therefore the coefficient of static friction μ_s is greater than the coefficient of dynamic friction μ_d . It was found that once the sliding is onset, the coefficient of dynamic friction is nearly independent of sliding velocity. [2]

Wear laws

Compared to the "well-known" laws of friction, our knowledge about the laws of wear is much smaller. Technical literature is full of reported data on wear, but these data are the results of experiments in the industry which simulate specific operations in service and usually complicated and uncertain compositions. This results in no general laws of wear that are widely applicable and no "wear coefficient" in sense such as Coulomb's coefficient of friction or Hooke's Young modulus. As a result of this, dimensional analysis cannot be applied to wear problems. Meaning that we cannot conclude rationally anything about wear from scaled experiments and we also cannot rationally estimate the wear occurring in an actual machine. Nevertheless, some authors have found empirical relationships that gave us a brief notion of how wear laws can look like. Let us take a brief look into the history of wear laws and its discoveries. [5]

In 1860, German mathematician Karl Theodor Reye presented his hypothesis that the volume of the removed debris due to wear is proportional to the work done by friction forces. This is called "Reye's hypothesis" or "energy dissipative hypothesis". Reye did experiments with pins and considered all the various ways in which the mechanical work consumed by friction is converted. He then proposed that a correct theory of frictional wear should be in accordance with all the factors that govern the conversion of frictional work into heat, electricity, deformation, shocks etc. The hypothesis is formulated mainly with observations made during wear, because other components of friction work are so rapidly propagated, that they will virtually remain stationary. That means the individual rubbing elements the effects can be observed at their points of origin. In sliding contact in pin, there is assumption that the size of wear in one direction refers to the frictional work in the same direction. This wear is proportional to the part of the total frictional work that is expended in overcoming the frictional resistance. Given the uniformity of material, lubrication and temperature for all points

on the pin surface, the ratios between various factors influencing pin friction are the same for the entire lifetime in each direction referring to the friction surface. Therefore, the total frictional work, and consequently the wear, is proportional to the total normal force applied to the relative projection of the friction surfaces. Reye proposed this written equation: "For a given pin, the normal displacement in any direction equals the same share of the total frictional work caused by the normal pressure, and it forms the product with its relative velocity in the direction perpendicular to the friction surfaces." [9][10]

In 1946, Ragnar Holm proposed a semi-empirical theory of galling wear based on the fact that the true contact area A_t between rubbing surfaces is given by equation (4), where N is the normal load and p_m is a flow pressure of the softer of the two surfaces. The physical mechanism being the plastic deformation on the local asperities on the softer surface. [11][12]

$$A_t = \frac{N}{p_m} \tag{4}$$

Another assumption made by Holm is that every encounter of two atoms of the opposite surfaces in true contact area has a certain statistical probability that one of the atoms will be pulled out of its parent surface. After counting all the encounters during the sliding and with the use of average interatomic spacings, Holm produced a simple wear volume equation (5). Where V is the wear volume, k is the probability of removing an atom (also as Z in literature) and L is sliding distance. [11][12]

$$V = kA_t L \tag{5}$$

If we insert equation (4) into equation (5), we get equation (6). Divided by true contact area A_t we obtain similar equation (7) with wear depth h and average normal stress P.

$$V = \frac{kNL}{p_m} \tag{6}$$

$$h = \frac{kPL}{p_m} \tag{7}$$

This theory by Holm assumes that the material is removed as individual atoms. Burwell and Strang [12] put this into question and proved that it's probably not generally valid. They proved it with electron microscopy examination of the wear particles removed from two hardened steel surfaces rubbing together under a lubricant, as shown on Figure 7. The size of the individual specks or grains in the Figure, the large black regions probably being its agglomerates, is 5-10 nm. This is quite small regarding the macroscopic point of view, but on the atomic scale, it consists of millions of atoms. The softer the material is, the particles are found larger. Burwell and Strang nevertheless began experimental procedure based on the findings of Holm to see to what extent equations (6) and (7) hold. [12]
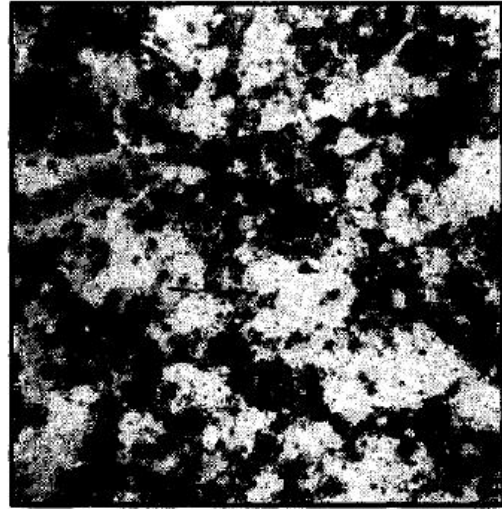

Fig. 7 - Electron micrograph of wear particles from hardened steel surfaces (25,OOOX) by Burwell and Strang [12]

Holm, Bowden, Tabor and others also proposed the relation between friction force F and true contact area A_t in equation (8) with *s* being the average shear strength of the local adhesion between the surfaces. [12]

$$F = s \cdot A_t \tag{8}$$

Dividing the equation (8) by equation (4) gives us the familiar expression of the friction coefficient μ in equation (3). Both *s* and p_m are plastic properties of the concerned materials and may be expected to have same ratio for both hard and soft materials. [12]

$$\mu = \frac{F}{N} = \frac{s}{p_m} \tag{9}$$

As mentioned before, Burwell and Strang did experiments in 1952 and 1958 to understand wear and hopefully propose a wear law and started their test with validating to which extent equations (6) and (7) proposed by Holm are valid. They had a simple wear machine with relatively soft pin rubbing the surface of a smooth hardened steel disk. The measured value was the wear of pin. Conventional depth wear rate W_d, which is rate of material removal as in equation (10), where h is depth of material removed from the soft pin/rider (i.e. wear depth) and L is sliding distance, is constant for a constant normal stress.

$$W_d = \frac{h}{L} \tag{10}$$

Instead, Burwell and Strang deemed it appropriate to plot wear rate against the corresponding stress P because it correlates more with equation (7), therefore plotting quantity h/LP against average normal stress P for two different materials as shown in Figure 8 and discussed the obtained curve with its two distinct regions divided by the approximate value of 1/3 of hardness. [5]
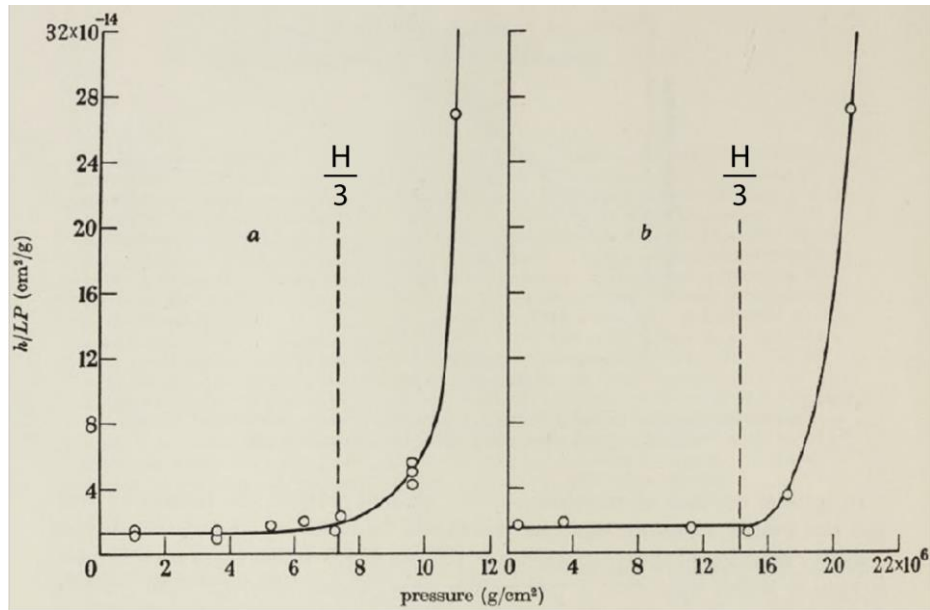
6

Fig. 8 – Wear in form of h/LP plotted against pressure P on the surface [5]

For low stresses, up until approximately 1/3 of hardness, the plot shows that quantity h/LP is constant as predicted by the equation (7). For stresses higher than 1/3 of hardness, the quantity h/LP increases sharply and completely abandons behavior according to equation (7). This shape of the curve was obtained with lubricant and without it, so the breakdown of the liquid film cannot be a reason for the sharp increase of quantity h/LP. However, breakdown of some sort of dry film, such as oxide or other material, is not out of question.

A detailed explanation of the high stress region and sharp rise of h/LP is not yet possible, but some things seem to be definite and will be shortly explained. Wear particles, due to certain reasons proposed by Burwell and Strang [5], cannot escape the contact and therefore produce more wear. The formation of the secondary wear particles is doing self-acceleration of the wear process. When particles are worn off the surface and they cannot escape the contact of two surfaces, they produce more wear. This can lead to some extreme cases such as forming the avalanche of wear material. [5][12]

The threshold value being 1/3 of hardness is probably due to the fact, that it is approximately the same value as the yield strength of completely work-hardened surface, such as wearing surface we are considering, according to Burwell and Strang. Because wear, especially adhesive in this case, is based on plastic deformation of asperities, the exceeding of the yield strength is necessary to induce such plastic deformation. When the asperity and its base exceeds the yield strength, the material starts to flow. [5][12]

Burwell and Strang therefore experimentally proved Holm's equation but the wear being constant only up to the pressures of 1/3 of hardness and with the proposition, that the relation should be applied to asperities, rather than atoms as Holm's proposed.

In 1953 J.F. Archard wrote his famous article about modelling wear named "Contact and Rubbing of Flat Surfaces" in which he proposes the theory of wear law based on the collision of two asperities as shown in Figure 9. [13]
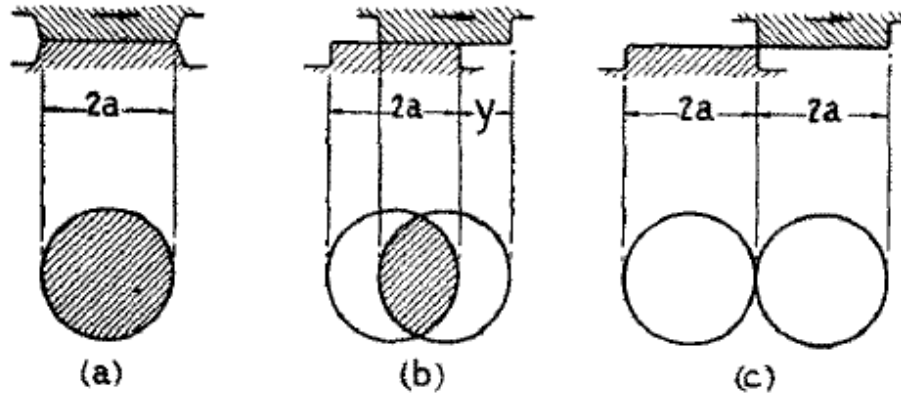
Idealized representation of single contact in sliding surfaces. (a) Maximum contact area of radius *a*. (b) After sliding through distance *y*. (c) After sliding through a distance 2*a*; contact area just reduced to zero.

Fig. 9 - Single contact in sliding surface by Archard [13]

Considering asperity contact as shown in Figure 9 we consider that the highest load the asperity can carry is shown by equation (11) where A is area, H is hardness and a is radius of the asperity. The total load N carried by all asperity contacts is then shown in equation 12 as a sum of all areas in contact. [13]

$$dN = A * H = \pi a^2 H \tag{11}$$

$$N = \sum dN = H \sum \pi a^2 \quad therefore \quad \sum \pi a^2 = \frac{N}{H} \tag{12}$$

Equation (12) shows us, that the area which is carrying the load is given by load N and hardness H. Also, wear particle can be formed in the shape of the half-sphere, hence its volume V_p is calculated in equation (13). [13]

$$V_p = \frac{\frac{4}{3}\pi a^3}{2} = \frac{2}{3}\pi a^3 \tag{13}$$

Not all contacts give rise to wear particles, so Archard assumed only a fraction θ does. In the original Archard article referred to as k, but for better understanding now, we use θ. Wear rate dW_w i.e. average volume worn away per sliding distance 2a is therefore shown in equation (14).

$$dW_w = \theta \frac{V_p}{2a} = \frac{\frac{2}{3}\pi a^3}{2a} = \theta \frac{\pi a^2}{3} \tag{14}$$

The total wear rate is then the sum of all asperity collisions as shown in equation (15).

8

$$W_w = \sum dW_w = \frac{\theta}{3} \sum \pi a^2 \tag{15}$$

If we insert equation (12) into equation (15), we obtain equation (16) and write wear rate as wear volume V over sliding distance L.

$$W_w = \frac{V}{L} = \frac{\theta}{3} \cdot \frac{N}{H} \tag{16}$$

Since fraction of particles worn away represented as constant $\theta/3$ is mathematical construct based on the choice of spherical particle, we can simplify it so wear coefficient $K = \theta/3$ and arrive at Archard's wear law in equation (17) which mirrors Holm's equation (6) [13]

$$\frac{V}{L} = K \cdot \frac{N}{H} \quad \rightarrow \quad V = \frac{KNL}{H} \tag{17}$$

Archard considered more mechanisms of deformation and shape but proposed a model which to his best knowledge copied reality, which is that particles are removed from surface as lumps by the plastic deformation. Equation (17) assumes that wear particles have hemispherical shape and have same radius. As said before, it is similar to Holm's equation and also replaces Holm's concept of material removal by atoms to the material removal by wear particles. Archard also proposes conclusions after his study as follows:

A) The wear rate is proportional to load and is independent of the nominal area of contact
B) If the K and H (or p_m) remains constant, the wear rate is independent of the speed of sliding.
C) The theoretical value of wear rate is independent of the model used to represent the surfaces

This findings made by Holm, Burwell, Strang, Archard and many others led to the most universally used model of wear today despite its shortcomings such as usage only for steady-state wear in this basic form, not considering the running-in phase (which is described in next paragraph) or the fact that wear coefficient K is very likely to be dependent on many different factors and is hard to predict. The law is known under many names in literature such as "Archard wear law", "Holm-Archard wear law", "Holm wear law" or "Reye–Archard–Khrushchov wear law". [13]

In 1952, Burwell and Strang [5] proposed two qualitative conclusions about wear. Firstly, that wear decreases with increasing the hardness of the rubbing surfaces, which corresponds with the propositions of Archard and Holm. Secondly, that wear increases with the distance of travel, but generally not linearly. Many exceptions to these two rules can be presented, but most of the time, the relationship between wear volume and distance of travel is as shown in Figure 10. [1][5]

The so called "wear curve" typically consists of 3 stages. "Running-in", "Steady state" and "Catastrophic or Severe wear" stage. Running-in stage is more described in next paragraph.
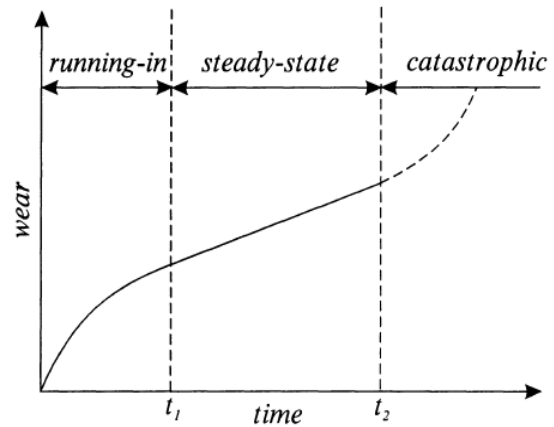

Fig. 10 - The most qualitative relation between distance and volume of material worn. Divided into 3 main stages "running-in", "steady state" and "catastrophic" or "severe" wear [5]

Steady state is characteristic by its linear nature of wear dependence on time. Most of the wear laws presented, such as Archard Law, are the best description of this linear stage. The Catastrophic stage is characteristic by its sharp increase in wear. In this stage, working conditions become drastically worse and the pair of materials in contact is quickly worn out. [1][5][14]

The behavior of wear in Figure 10 in the "running-in" stage can be investigated using Figure 8. When the machine is freshly assembled, even the best prepared surfaces will bear load only on an edge or a corner or a local high region. Therefore, the true contact area is quite small and the normal load is high. High load on small area results in high stress. This high stress is in the high-stress region in Figure 8. Because of this, the initial depth wear rate is extremely high and the original wear area is rapidly enlarged. At this stage, caution is necessary to avoid catastrophe by the self-accelerating process of wear by the primary wear particles that cannot escape the contact. [5]

There are generally three possibilities that can happen during running-in phase. The first possibility is when the surfaces are made initially near-enough to conform and the rapid wear brings their entire areas into mating, the run-in can be considered successful. When the stage of mating entire areas of the two given surfaces is reached, load is carried by bigger area, stresses should drop to the low-stress region on Figure 8 and the wear rate becomes smaller. This wear rate is often very small, especially compared to the initial running-in wear, that it may be negligible. It is often said that when the surfaces have reached the run-in condition, the wear ceases. This is sometimes ascribed to a "glazed" or otherwise specially conditioned surface. The best way of running-in the surfaces together is to simply operate in the low-stress region of the curve on Figure 8. [5]

The second possibility is when the two surfaces are not made to conform. In this case, the initial wear area grows only gradually and the average stress remains in the high-stress region for a long time or even indefinitely. This leads to the wear during whole life of the part and no ceasing of the wear as in the first possibility. [5]

The third possibility occurs when the growth of the wear area is sufficient to enable the stresses to drop into the low-stress region and therefore into the low-wear rate region, but the wear area

10

is still only a fraction of the designed area. The result of this is a surface not functioning properly such as gas or oil seal in the case of piston rings. Rings pump the oil, but their wear rate is so low that they will never wear completely in a normal lifetime and the seals against the leakage of the transported medium will not form. In the case of the piston rings, the surface is purposefully designed to be rough. This serves as an insurance that the surface stresses stay in the high-stress region, therefore in the high wear rate region until the ring is well in contact over most of its face and serves as the seal for the transported medium as it should. [5]

Wear coefficient K

When we take look at the equation (17), the Archard law in the form V = KNL/H, the important part of calculating wear is the wear coefficient K. It is a non-dimensional constant which physically represents the probability that during contact of two surfaces at asperity, the sizeable wear particle is produced. It is considered primarily for the softer surface of the two, since it's the one wearing faster. For harder surface, K is generally a 1/3 of the K of softer surface. [15]

As mentioned before, the wear coefficient K physically represents the probability of producing a wear particle. In 1958, Rabinowicz [16] presented a criterion for the critical size of wear particle to come loose and supported it with experimental data. Rabinowicz criterion was confirmed in 2016 using quasi-molecular simulations by Aghababaei, Warner and Molinari (AWM) in [17]. [16][17]

Criterion is based on elastic energy. To detach a particle from the surface, elastic energy stored in the fragment needs to overcome the work of adhesion between the fragment and the surface to which it is attached. As explained more closely in [16], from the energy balance, we can extract critical diameter of wear particle d_c as shown in equation (18). E is the Young modulus, W is work of adhesion of the system and σ_{yf} is yield strength of fragment material. If the particle is larger than critical diameter d_c, it comes loose.

$$d_c = \frac{30EW}{\sigma_{yf}^2}$$

(18)

Rabinowicz in 1981 [15] and many other authors conducted experiments to determine wear coefficients K for different material couples. Rabinowicz chose to aim his interest at metals with other lattice than HCP, no soft metals and metals with low oxidation rate. He did so with running experiments at N = 1kg, distance L calculated based on speed v = 0.01 m/s and time varied from 1 hour up to 96 hours. Hardness varied for each material and wear volume was determined by weighing the specimen before and after the experiment on precise scales and calculating the volume worn away. K is therefore calculated as shown in equation (19). Rabinowicz multiplied Hardness by 3 such as shown in equation (15) with θ but equations usually have this form. [15]

$$K = \frac{VH}{NL}$$

(19)

He ran experiments with a group of materials consisting of Al, Ag, Cu, Nb, Ni, Mo with various level of lubrication and based on their metallurgical compatibility – the extent to which the materials are alike and are expected to form strong bonds. Experiments are more closely explained in [15] and values for coefficient K as shown in Figure 11. are presented. The nature of the experiment focuses on adhesive wear, which is the most investigated form of wear since its most difficult to overcome and isn't entirely possible. More on the topic of adhesive wear and other types of wear is in chapter 1.2. As we can see, the K has tendency to decrease when we use metallurgically incompatible materials and assure good lubrication. [15]

	Identical	Compatible	Intermediate	Incompatible
Symbol	◯	⊙	◗	●
Unlubricated	5×10^{-3}	2×10^{-3}	4×10^{-4}	5×10^{-5}
Poor lubricant	10^{-3}	4×10^{-4}	8×10^{-5}	10^{-5}
Good lubricant	10^{-4}	4×10^{-5}	8×10^{-6}	10^{-6}

Fig. 11 - Wear coefficient K values based on different metallurgical compatibility and lubrication [15]

Archard himself in 1961 made experiments as explained in [18] and arrived at values of K as shown in Figure 12. We can see that the range of K values is spread from 10^{-3} to 10^{-7} order, which is quite a big dispersion. [18]

Wear coefficient K is hard to predict because it is dependent on more things than just the physical quantities mentioned in Archard equation. With wear being a complex phenomenon caused by variety of mechanism, as will be explained in chapter X, such as adhesion, abrasion and these mechanisms also interacting with each other, wear coefficient is difficult to predict universally or even for each isolated mechanism of wear. Wear coefficient is also highly sensitive to the test conditions in which they are assumed such as parameters of the experiment or the temperature, lack of standardized test methods, statistical variability, material properties and possible even more.

TABLE I. Wear rates, coefficients of friction and values of K. Load 400 g; speed 180 cm/sec. Rings are hardened tool steel except where stated otherwise.

Material	Wear rate 10^{-10} cm³/cm	Coefficient of friction μ	K
Mild steel on mild steel	1570	0.62	7.0×10^{-3}
60/40 leaded brass	240	0.24	6.0×10^{-4}
P.T.F.E.	20	0.18	2.5×10^{-5}
Stellite	3.2	0.60	5.5×10^{-5}
Ferritic stainless steel	2.7	0.53	1.7×10^{-5}
Polythene	0.3	0.65	1.3×10^{-7}
Tungsten carbide on tungsten carbide	0.02	0.35	1.0×10^{-6}

Fig. 12 - Wear coefficient K values measured by Archard [18]

1.2. Wear mechanisms

As mentioned before, wear is a progressive loss of material from surfaces due to multiple reasons. With friction, wear is one of the main topics of the tribology field. However, it is still one of the least understood phenomena in tribology. This is partially because wear is influenced by many complex processes, such as contact, plasticity, crack nucleation and propagation, chemical reactions, material mixing, material transfer between surfaces, lubricants and formation of surface layers. [19][20]

In general, there are considered 7 mechanisms of wear: [19][21]

1. Abrasive from contact with hard sharp granular materials
2. Abrasive from embedded particles
3. Adhesive
4. Fretting
5. Cavitation erosion
6. Particle erosion
7. Fatigue of surfaces

In the beginning of the process of modelling of the wear, it is necessary to choose the main wear mechanism. This is the reason why we must understand all of them and make an informed decision.

1.2.1. Abrasive wear from contact with hard sharp granular materials

Abrasive wear is characterized by material removal from the surface by a cutting action or by a process of multiple indentation from abrasive particle, as seen in Figure 13. This process can be either controlled and intended or not. Controlled use is beneficial and done in processes like grinding, polishing or filing. When the process is not controlled, e.g. randomly occurring in machine operation, it becomes a problem. Surfaces damaged by abrasive wear can show multiple magnitudes of damage. It can be fine scratching when damaged lightly or deep gouges in the surface which can seriously impart functionality of a given machine part. Real life example is seen in Figure 14. [19][22]
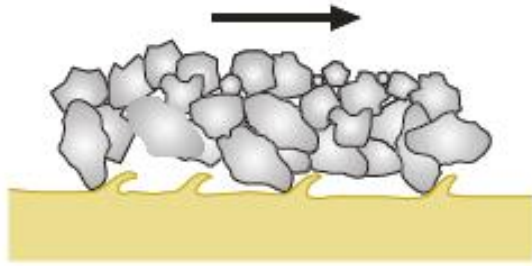

Fig. 13 - Principle of abrasive wear from moving contact with hard sharp granular materials [19]
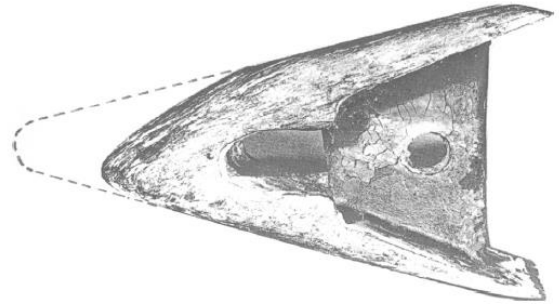
Fig. 14 – Example of abrasive wear from moving contact with hard sharp granular materials [22]

1.2.2. Abrasive wear from embedded particles

Abrasive wear can also occur in two smooth surfaces, if one of them has hard embedded particles in its surface. The debris leaving the material surface is different for different materials. Ductile materials, such as steel, can have spiral shaped debris. On the other hand, very hard materials tend to have debris in the form of chips. This is because of a local brittle fracture of the material. We can see the principle and real life case in Figures 15 and 16. [19][22]
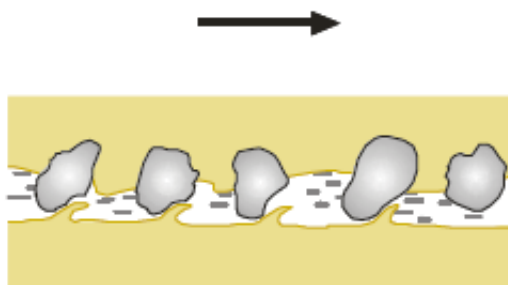

Fig. 15 - Principle of abrasive wear from hard sharp particles trapped between moving surfaces [19]

Fig. 16 – Example of abrasive wear from hard sharp particles trapped between moving surfaces [22]

14

1.2.3. Adhesive wear

Adhesive wear is the mechanism occurring when two macroscopically smooth surfaces rub against each other. Generally, and as previously mentioned, each surface has a certain surface roughness. Therefore no surface is ideally flat. When two such surfaces make contact, rubbing occurs at the high spots as seen in Figure 17. These local areas experience concentrated contact loads and interactions and they tend to adhere to each other. This leads to material being dragged away along the surface. Surfaces damaged by adhesive wear mechanism show polished surface with fine flakes of wear debris and spots of "torn out" material. This can be very well seen in Figure 18. [19][22]
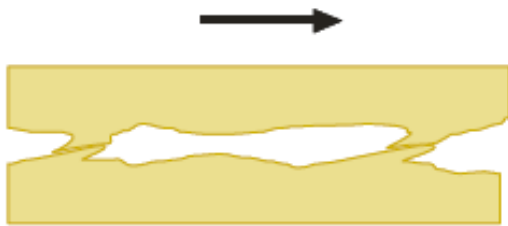

Fig. 17 - Principle of adhesive wear from the rubbing together of relatively smooth surfaces [19]
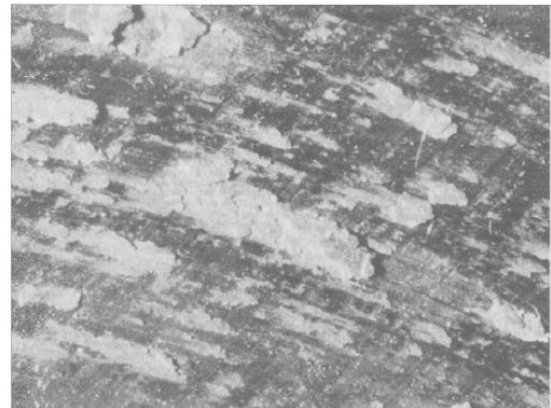
Fig. 18 – Example of adhesive wear from the rubbing together of relatively smooth surfaces [22]

1.2.4. Fretting wear

Fretting is a particular form of adhesive wear. It occurs during the contact of materials if there are small oscillatory movements present. Principle presented at Figure 19. Oscillatory movements can be either intended or not. Intended movements are for example in gear couplings. Not intended movements can arise from many of reasons, for example from the deflection of machine components with clamped joints or press fits. The product of the fretting is usually fine powdered and oxidized wear debris. Fretting usually results in surface damage and roughening of the surfaces in contact. Fretting plays a huge role in the life cycle of the wire ropes. The reason for this are the relative deflections between the individual steel wires in the rod. These deflections arise due to changes in the load on the wire or during bending the wire around a pulley. Example of fretting wear damaged surface can be seen on Figure 20 [19][22]
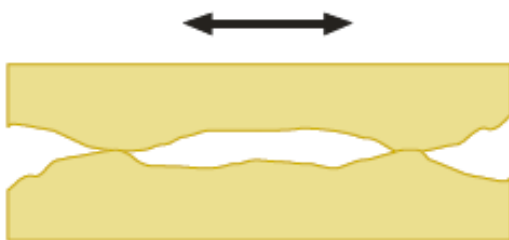

Fig. 19 - Principle of fretting wear from the small oscillatory movements between two smooth surfaces [19]

Fig. 20 – Example of fretting wear from the small oscillatory movements between two smooth surfaces [22]

1.2.5. Cavitation erosion

Cavitation erosion as a process is the removal of material from the surface by the high speed impact of a liquid. Cavitation erosion occurs on the components that are subject to low transient fluid pressures on their surface. The scheme of this erosion formation can be seen in Figure 21. For example ships propellers. Collapse of a low pressure vapour bubbles make intense local impact on the surface of a component. How cavitation erosion damage appears in reality can be seen in Figure 22. [19][22]

Fig. 21 – Principle of cavitation wear caused by collapse of the low pressure vapour bubbles [19]
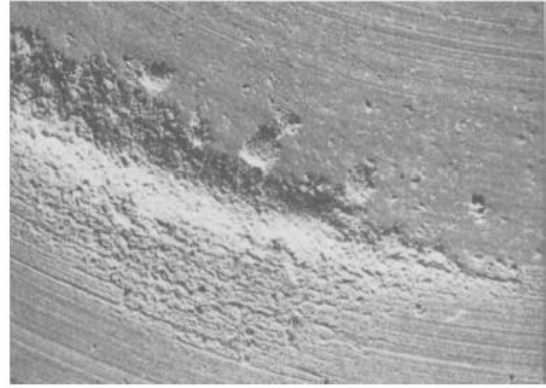

Fig. 22 – Example of cavitation wear caused by collapse of the low pressure vapour bubbles [22]

1.2.6. Particle erosion

Particle erosion is the removal of material from the surface caused by a stream of hard particles carried in fluid flow as schematically shown in Figure 23. Particles can be directed at the surface purposefully as for example in shot blasting processes or it arises incidentally. Incidental particle stream affecting the surface is for example in the pipelines or in the components operating with sand or other particle. Surface heavily damaged by particle erosion is shown on Figure 24. [19][22]
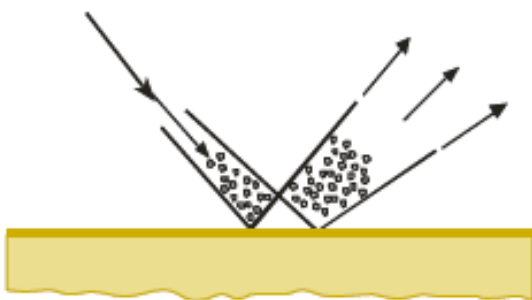

Fig. 23 – Scheme of particle erosion from hard particles in a stream of fluid [19]

Fig. 24 – Example of particle erosion from hard particles in a stream of fluid [22]

1.2.7. Fatigue of surfaces

Wear of the surface can be caused also by fatigue of surfaces. This occurs when fatigue cracks in the surface join to create loose particles. Surface fatigue can occur due to two different mechanisms. Either contact stress fatigue mechanism or thermal stress fatigue mechanism. Contact fatigue usually occurs in rolling contacts where the passage of another component, for example ball or roller, over the surface causes alternating tensile and compressive stresses as shown in Figure 25. These alternating stresses can create fatigue cracks. Surface damaged by contact stress fatigue is shown in Figure 27. The mechanism of thermal stress fatigue arises from the transient heating and cooling of surfaces. Especially when this effect is combined with surface frictional forces, for example in clutch plates or heavily loaded plain bearings. Another severe case of surface fatigue is when the studied surface is in contact with a very hot material, for example molten metal. Example of the thermal fatigue can be seen in Figure 26. [19][22]
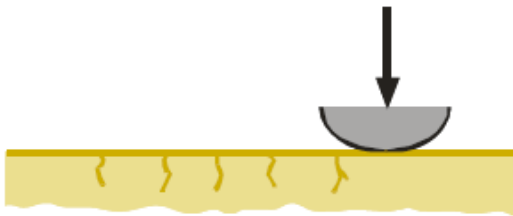
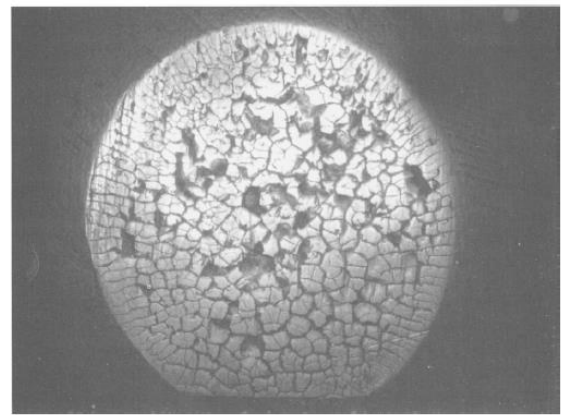*Fig. 25 – Scheme of release of particles due to the fatigue of the surface [19]*

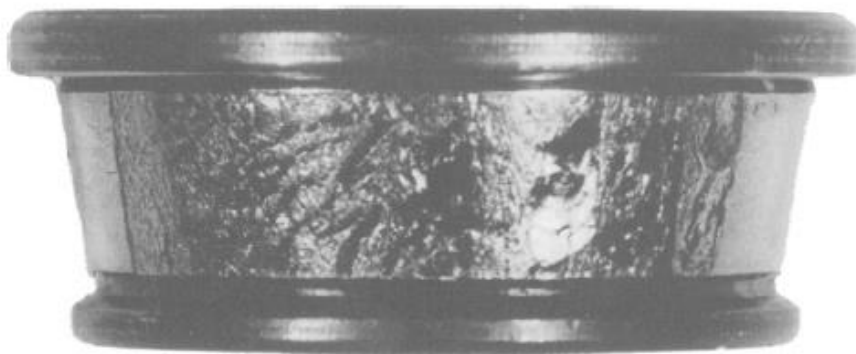*Fig. 26 – Example of thermal stress fatigue [22]*

Fig. 27 - Example of contact stress fatigue [22]

1.3. Adhesive wear

Adhesive wear was chosen as the primary mechanism for the given task of modelling the wear in turbocharger actuator mechanism in order to simplify the modelling. It will therefore be described more profoundly in this section. The reason for this is deeper understanding and therefore better solving of the given problem. Mechanism of adhesive wear is briefly described in section 1.2.3 and the author recommends reading that section first, because this section is a follow-up.

As mentioned before, adhesive wear occurs when two smooth surfaces rub against each other. Surfaces have a certain surface roughness and the high spots tend to adhere to each other. The nature of this adhesion is the creation of atomic/molecular bonds same as creation of bonds between the atoms/molecules in material. When two surfaces "meet" and make contact, chemical bonds may be created. When the surfaces move, the bonds sometimes stay put and "pull" the material out because its more energetically favorable. This leads to certain parts of material being torn out of the surface. [23]

Out of all 7 wear mechanisms, adhesive wear is the one least avoidable and remains even when all other types of wear are eliminated. When we carefully control conditions during laboratory experiment, such as eliminating all corroding materials as well as all abrasives in the environment, it is never possible to completely avoid small amount of wear due to galling. [24]

As mentioned before, load is carried by the high spots of material, which creates the concentrated stresses in local points. It has been experimentally shown by Bowden and Tabor [7] in 1950 using electrical conductance in contacts, that area of these "high spots" is rarely greater than 1/100 of the apparent area of study and often as small as 1/10 000 of it. Because the true contact area is very small, it takes no more than a few kilograms of load to create high local pressures. Usually, these high local pressures exceed the yield point of the softer of the two materials. When we combine these high local pressures with the relative sliding motion, minute welds are being formed at each of the local areas. During continued sliding, welds are sheared. Welds often shear on their original surface of contact because the weld is not strong enough to pull the other material out. Due to work-hardening of the surface, some of the welds are as strong or stronger than the softer of the two base materials.
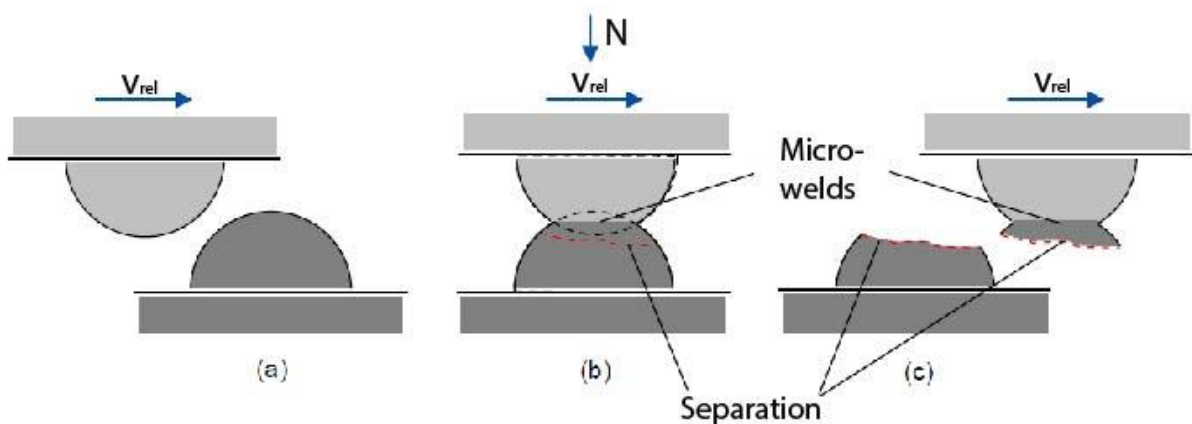

Fig. 28 - Adhesive wear mechanism (a) before contact; (b) during contact; (c) after contact [25]

This leads to these strong welds shear not at their original surface of contact, but at the surface of the base materials, usually the softer one. This results in some of the softer material being transferred to the other, harder, surface. Figure 28 shows this schematically. [7][25]

Aghababaei, Warner and Molinari in 2016 also conducted simulations regarding the adhesive wear-mechanism at the asperity level. They observed both gradual smoothing by plastic deformation and fracture-induced debris formation. Schematically shown on Figure 29. [17]
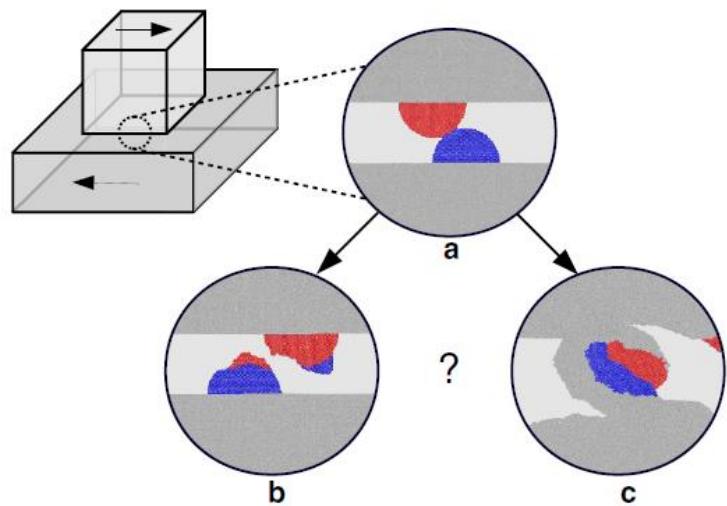

Fig. 29 – Two possible asperity-level adhesive wear mechanisms by [17]. (a) adhesive interaction between two asperities (b) gradual smoothing mechanism by plastic deformation (c) fracture-induced debris formation mechanism observed in AFM wear experiments. The red and blue coloring of atoms is artificial for better visualization.

This transferred material may either be adhered to the latter surface and be worked into it or it can be knocked loose and become loose wear particle. According to Burwell [24], if any of those two options happens, the softer surface of the two is subject to wear damage. In contrast, Neale [6] defines wear only in the case when the particle is broken away, not when it is transferred to the other surface. Because such transferred material often resides on a surface and may even go back to its original surface. According to Neale [6] the creation of the bonds (also called "cold welds") is only the first stage of the adhesive wear and by itself is not directly responsible for the wear. After that, some secondary mechanism is required to break the particle away from the surface and therefore cause wear as Neale describes it. The particles frequently form groups and break away as a single entity. There are many explanations for this phenomenon and Neale mentions the one where the elastic energy exceeds the surface energy and therefore causes break-away. This corresponds with Rabinowicz wear criterion [16] mentioned earlier. This final stage of the wear process is greatly affected by the environment. [6][7][24]

Neale suggests looking upon the adhesive wear system as being in a state of dynamic equilibrium with its environment. Continuous sliding and the exposure of fresh surfaces is not able to go indefinitely. This situation is usually stabilized by the healing reaction of the air or other active components of the surrounding fluid. Rupturing and healing processes are therefore in balance. This balance can be upset by the change of the operating parameters such as speed of sliding. When the change of the balance occurs, the contact of two surfaces can abruptly be changed from low to high wearing stage. E.g. increasing the sliding speed not only reduces the time available for the healing processes, but also generates more heat by friction in contact, which may accelerate chemical reactions or desorb weakly bound adsorbents. [6]

In 1984 Rabinowicz characterized three general types of adhesive wear. Namely galling as severe wear, moderate wear and burnishing. Severe galling wear occurs primarily in metals with a high degree of metallurgical compatibility when the surfaces are clean or poorly lubricated. Wear coefficient for galling is in range 10^{-2} - 10^{-4} and wear particle sizes is approximately 200-20 μm. Severe adhesive wear in non-metal materials occurs very rarely. Moderate wear occurs when the surfaces are less compatible, well lubricated or pressure between the surfaces is low. For moderate wear, wear coefficient is typically in range 10^{-4} – 10^{-7} and wear particle size ranges from $20 – 2$ μm. The mildest form of adhesive wear, called burnishing, is encountered in special circumstances. Such circumstances can be e.g. high incompatibility of metal couple or one of the surfaces has HCP lattice, in low pressure and with good lubrication. When this happens, surfaces take on a burnished appearance, hence its name. Wear coefficient is in range 10^{-6} – 10^{-8} and no sizeable wear particles are observed. Transition between these adhesive wear regimes often occurs abruptly. [26]

Even if the wear process starts off as adhesive mechanism, the break-away of the particles is involved and therefore the mechanism can change to abrasive. As mentioned before, abrasive wear is scratching or ploughing caused by hard particles in the softer surface. In most cases of adhesive wear, the wear debris is formed as oxides, which are generally hard and therefore cause abrasion to the surfaces. If the surfaces are subject to a small oscillatory motion, the fretting wear mechanism also helps the buildup of the debris between the surfaces. The debris later serves as the buffer between the two surfaces which slows down the wear rate. [6]

The most convenient demonstration method of adhesive wear is by means of radioactive tracers. In the technique of radioactive tracers, one of the surfaces in contact is made radioactive and the other surface is not. After this initial preparation, surfaces are rubbed together in experiment and then the surface without initial radioactivity is observed. Subject of the observation is the examination for evidence of radioactivity. Examination can be done either with a Geiger counter (electronic device used for detecting and measuring ionizing radiation) or by placing a photographic film in contact with the non-radioactive surface and allowing

Fig. 30 - Autoradiograph of two friction tracks induced in lubricated conditions [24]

the particles of the transferred radioactive material to be revealed. This revelation can be seen in Figure 30 of autoradiograph of two friction tracks which occurred during rubbing in lubricated conditions. [24]

Nature of the surfaces and the ambient conditions can profoundly affect the amount of adhesive material transfer. Regarding the nature of the surfaces, adhesive wear is greatest in the couple of same or similar materials, for example steel against steel. It is therefore recommended that we design a material couple that has contact between a metal and non-metal material, such as mineral, plastic or an oxide coating. [24]

The reason why we never run two members of the same metal is explained by Holm in [21]. When we run two members of the different metals in the contact, the specific friction force ψ (friction force F per true contact area A) , the pressure and consequently also the contact area are prescribed by the strength of the softer member and the harder member prescribes the structure of the surface. If the structure of the surface of the harder member is smooth, the contact surface between two members remains smooth without interlocking. However, if we run the same metal on both members, there is no "harder" or "softer" member. Both have the same properties. Therefore the pressure p and specific friction force ψ can produce plastic deformation in both members, not just the softer member. The contact surface of this same metal contact will be in labile state and have a wavy structure. This wavy structure of contact surface will lead to interlocking of the couple and therefore high wear. [21]

Another important fact to consider is the cleanliness of the surfaces. In environments with lower pressures, e.g. at high altitudes or in vacuum, and if the surfaces have been heated to drive off any adsorbed or oxide films, the material couple will stick together in the slightest pressure. In this case, separating the materials is very difficult, which indicates the strong nature of the welding tendency, when all interfering materials i.e. atoms are removed. When the same material couple is put into an experiment in air (not vacuum), we would not be aware of any adhesion if we haven't done the radioactive tracers measurement. Adhesion strength is greatly reduced by reducing the surface interactions at the surface and therefore even a few ppm of atoms of air decreases the adhesive strength. Thick films of liquids or gases reduce μ even more because it is easier to shear into a fluid film than to shear solid to solid contact. Adhesion strength is reduced either by the presence of contaminants or deliberately applied fluid film e.g. air, water or lubricant. [23][24]

From these two observations, we can infer that adhesive wear is most pronounced between similar materials and in vacuum conditions. Additionally, a crucial aspect to consider is the temperature dependency of adhesive wear. As temperatures rise, adhesive wear tends to accelerate. This phenomenon is attributed to the atomic-scale formation of welds, which can be likened to inter-atomic bonding, also explained as surface chemical reactions. In cases where materials are heated close to their melting points, inter-atomic bonds tend to weaken, while the bonds in the resulting cold welds tend to become stronger. Consequently, material galling becomes more severe. This underscores the tendency for low melting point materials, such as glass, to experience more pronounced galling compared to high melting point materials, such as ceramics. [24]

The temperature in question is not only the temperature of the environment, but also the heat created by friction. Heating by friction is a self-regenerating process. Self-regenerating process meaning that when initial adhesive transfer of material produces a high spot on the surface, the high spot participates in the friction process and therefore increases it. With increased friction, an increase in the heating occurs and the temperature rises. Rise of the temperature leads to more severe galling and creation of more material transfers which create high spots because the adhesive force strength rises with temperature. Also, in metals, with rising temperature the yield strength decreases and ductility increases, which leads to an increase in the true contact area. This is therefore a cycle, a self-regenerating process. When this self-regenerating process is not checked, it can rapidly lead to a large scale welding or freezing the surfaces together. In this case, surfaces become essentially one piece. If we want to separate this welded one piece of material couple, large scale damage is inevitable. [24][27]

To reduce adhesive wear damage, we apply lubricants, whose primary function in reducing the adhesive wear is to insulate the two surfaces from each other. This insulation is done by introducing a layer of foreign material, lubricant, which will prevent the welds from forming. Adhesive wear can be reduced greatly using proper lubrication, such as the most effective oiliness or E-P lubricants, but it cannot be eliminated completely. This incomplete elimination means that there are always a few welds present, which can be the starting points for large scale galling if the conditions become sufficiently severe. [24]

1.4. Models of adhesive wear

Various approaches can be used to model or predict the wear. It can be Empirical Models that are based on experimental data and observation of wear under specific conditions, Finite Element Method for simulating the contact and interaction between the surfaces on macroscopic scale, Molecular Dynamics for understanding the behavior of contact at atomic and molecular scale, Fracture Mechanics in case of wear processes initiating and propagating the cracks or Thermal Mechanics to study the coupled effects of mechanical loading and heat generation. Each of those methods has its uses, advantages and shortcomings that can be found in literature. However, they are not the subject of this thesis as it focuses solely on empirical models such as Archard Wear Law to develop efficient wear model using various statistical or machine learning methods to predict wear occurring in turbocharger actuator mechanism.

As mentioned before, adhesive wear is the least avoidable wear type and by many authors also chosen as the main mechanism when considering the wear. In previous chapters, the difficulty of empirical "modeling" or "predicting" of the adhesive was stated with main factors of influence being the complexity of the wear process (contrary to our simplification to 7 distinctive wear mechanisms), temperature and pressure effects, chemical processes and many more facts. This chapter focuses only on empirical equations, which are supposed to predict specifically the adhesive wear mechanism and therefore ignoring 6 other types of wear mechanisms presented in chapter 1.2 in dry conditions.

Apart from the most noted and used law i.e. Archard Law, many other authors proposed various equations with various variables to best predict the wear of materials. Most notable of these equations found in available literature were put together in Table 2 to choose the most suitable one for our case. Experimental setup and more specifics can be found in the respective articles of the authors cited.

Table 2 - Empirical laws for adhesive wear by various authors

Author	Wear volume V [μm^3] Wear depth h [μm] Weight loss Δw [g]	Wear coefficient K [-]	Note
Archard [13]	$V = \dfrac{KNL}{H}$	$K = \dfrac{VH}{NL}$	
Goryacheva [1] (integrated from wear rate)	$V = Kp^{\alpha}v^{\beta}t$	$K = \dfrac{V}{p^{\alpha}v^{\beta}t}$	α, β = coefficients

Rhee [28]	$\Delta w = K F^a v^b t^c$	$K = \dfrac{\Delta w}{F^a v^b t^c}$	Δw = weight loss a,b,c = coefficients
Yang [29] proposed wear coefficient K_s put into Archard Law	$V = \dfrac{K_s N L}{3H}$ $\quad$ where Ks $\rightarrow$	$K_s = \dfrac{3HC}{PL}(1 - e^{-BL})$	K_s = steady state wear coefficient B, C = coefficients e = Euler number 2.718...
Bhattacharyya [30][31]	$V = KAL$ $\quad$ where $K \rightarrow$	$K = B(1 + e^{Cp})$	B, C = coefficients e = Euler number 2.718...
Varenberg [32] extension for Archard law based on bearing curve ratio	$h = -\dfrac{1}{a}\ln\left((1 + e^{ab})e^{aKpvt} - e^{ab}\right)$	$K = \dfrac{1}{apvt}\ln\left(\dfrac{e^{-ah} + e^{ab}}{1 + e^{ab}}\right)$	Extension of Archard Law for the running-in stage of wear h = wear depth a, b = constants
Mishina [33]	$V = K \cdot \left(\dfrac{p * L}{\sigma_y}\right)$ Where K $\rightarrow$	$K = \dfrac{1}{3} \cdot \dfrac{n}{\lambda} \cdot \left(\dfrac{b}{a}\right)^3$	n = number of wear elements generated at the junction (physical) lambda = chemisorption rate (chemical) a, b = statistical parameters of junctions and asperities σ_y = yield stress
Queener [34]	$\Delta w = \beta(1 - e^{-nL}) + KL$	$K = \dfrac{\Delta w - \beta(1 - e^{-nL})}{L}$	Δw = weight loss β = coefficient of maximum contribution to wear by transient mechanism n = coefficient

As the most suitable equation for our purposes of modelling the adhesive wear, the Goryacheva [1] equation has been chosen as shown in equation (20) for wear volume and equation (21) for wear coefficient K.

$$V = Kp^{\alpha}v^{\beta}t \qquad (20)$$

$$K = \frac{V}{p^{\alpha}v^{\beta}t} \qquad (21)$$

The Goryacheva equation contains experimental parameters pressure p, speed v, time t, measured value of wear volume V and coefficients K, α and β. This equation will be employed in the Methodology chapter to see if it can empirically predict the wear based on experimental parameters available for our use case.

1.5. Machine learning
1.5.1. Introduction to machine learning

Machine learning (ML) has been around for more than 30 years in various forms, solving various problems in various fields and is one of the major reasons why today state of the art software can do the "magic" things it does. Things like speech recognition, ranking web search results, recommending you videos on platforms, filtering spam emails, recognition of images and many more. But what exactly is machine learning?

In definition given in 1959 by Arthur Samuel: *"Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed."* which simply explains the basic premise of ML. Instead of explicitly programming the system if the email is spam or not, which would be impossible if you wanted to cover every possible option of spam email, you can feed the model huge number of examples of what are spam emails and what are not spam emails. Based on mathematical algorithms, the program "learns" how spam email generally looks like and then, based on probability, it decides if the incoming email is spam or not and forwards it to the appropriate folder in your mailbox. [35]
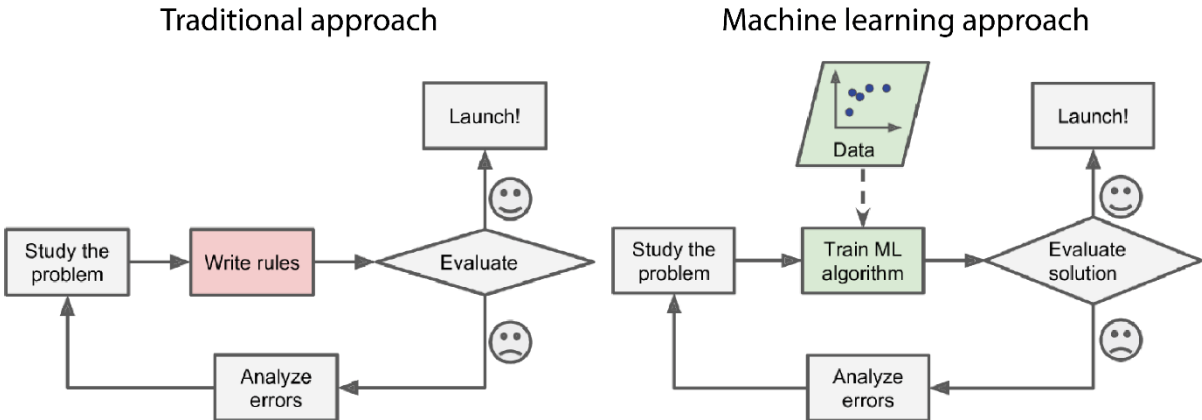

Fig. 31 - Traditional vs machine learning approach [35]

As briefly mentioned in the previous paragraph, machine learning is fundamentally data driven i.e. algorithms learn patterns and relationships from data to make classifications (spam or not spam class assigned to each email in training data), predictions (e.g. value of certain feature, such as wear volume or price of stock in stock market) or decisions (mechanism to approve whether or not the person should have approved the credit card application). Data therefore plays big role in machine learning methods and we will give it more attention later in this chapter. [35][36]

Machine learning approaches offer many advantages and disadvantages which we need to remember in order to maximize the potential of the methods in our use case. ML is good for problems that would need lots of fine tuning or long lists of rules (e.g. spam filter), Complex problems for which traditional approaches yield no solution, Fluctuating environments and offers scalability and adaptability for large datasets producing predictions with great accuracy. On the other hand, as with every method, ML has also shortcomings. Results are dependent on the quality and quantity of labeled data fed into the model, models may overfit or underfit (will be explained later), some models offer difficult interpretability (e.g. We don't know what exactly is happening inside Neural Networks hidden layers) and they often require significant computational resources. With these advantages and disadvantages in mind, we can deploy the ML methods on our use case and hopefully make quality assumptions about the use case. [35][36]

1.5.2. Categories of machine learning

Machine learning systems can be classified into categories by different criteria. Such as if human supervision is present or not, if the system is able to learn incrementally or not and whether learning is instance-based or model based. Instance-based learning is based on measuring similarities between new data points and existing data points, for example if certain number of closest neighbors to the new data point represent certain class, the new data point will also probably represent this class. Model based learning is about capturing underlying patterns and relationships in data such as predicting regression curve by which the data is fitted. For our purposes, machine learning will require no need to learn incrementally and learning will be model-based since our data doesn't have specified classes and we will rather predict values by regression and other algorithms. Dividing machine learning by human supervision is the most important topic and will be given separate attention below. [35][36]

The first category regarding human supervision in ML is supervised learning where the human supervision is present in the form of labeling the training data. For example the image of a dog is presented with the corresponding label "dog". When the algorithm makes prediction about new data, it is "told" if it succeeded and labeled new unseen dog picture as "dog" or failed, labeling dog picture as for example "cat". Another example of labeled dataset is our use case – we have labeled values such as temperature, hardness ratio, CoF or wear volume and certain value for each feature. In broad terms, there is a supervisor to give the algorithm insight on how much a decision is good or bad, which will be more specified in the metrics paragraph. Typical supervised learning tasks are classification (spam or not spam) and regression (predicting target numeric value). Most used algorithms in this category are Linear Regression (LR), Logistic Regression (LogR), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Random Forest (RF) and Neural Networks (NN). This category will be of greatest interest in this thesis. [35][36]

On the other hand of human supervision, there is learning without human supervision, called unsupervised learning, where training data is not labeled. It is therefore up to the algorithm to find the labels and define them. Model is therefore forced to learn the underlying structure of dataset and relationships between the data. Typical unsupervised learning tasks are clustering - detecting groups of similar features via e.g. K-Means Clustering, anomaly and novelty detection via e.g. One-class SVM or Isolation Forest (IF) and visualization, dimensionality reduction via Principal Component Analysis (PCA). There is also a combination of supervised and unsupervised learning, called semisupervised learning, which operates with partially labeled data and utilizes combination of algorithms from both categories. Unsupervised and semisupervised learning are in large part not suitable for our use case and therefore will be given no further attention in this chapter unless they seem effective to use. [35][36]

There is a third category of machine learning, regarding human supervision, called reinforcement learning. Reinforcement learning differs from the previous 2 categories. The learning system can observe the environment, select and perform actions. In return, it gets rewards for good actions or penalties for bad actions. The goal of the model is to find the best strategy to get the most reward over time. Imagine a robot placed in a maze and tasked with finding its way to the exit. The robot receives positive rewards for reaching the exit, negative rewards for hitting walls or obstacles, and zero rewards for moving without progress. Using reinforcement learning algorithms, the robot learns to navigate the maze efficiently by exploring different paths and learning from past experiences. [35][36]

1.5.3. Data and Feature engineering

As a popular saying in data science and machine learning goes *"Garbage in, garbage out"* the importance of good quality and quantity of data is the most vital part of any machine learning prediction. If we have poor quality data to enter, the result will also be poor. Your dataset needs to be representative of the cases you want to generalize the model to. If not, you will receive results based on the not representative sample which will be different. You should minimize the errors in data, remove outliers and noise so the system can detect underlying patterns regarding the general behavior. It has been proved that different machine algorithms from simple to complex perform almost identically well given enough data, as is shown in [37][38]. It is therefore in our best interest to obtain a lot of data with different values that are representative of the general behavior. Data measurement also needs to be done properly, so we do not introduce some bias already due to the nature of the measuring method for instance or with the methods of data scraping from online sources. This is called sampling bias. To summarize, the three main problems we want to avoid are the low quantity of data, non-representative data, and poor quality data. [35][36]

When we measure the data for machine learning, we want to include data from more than one parameter for training. Let's demonstrate this with the example of predicting equipment failure. We take multiple parameters into consideration, such as operating time, number of cycles, temperature, load or humidity. These parameters are called features and the process of selecting a good set of features is called *feature engineering*. Feature engineering involves feature selection, feature extraction and creating new features by gathering new data. Feature selection is when we choose features to put into the model such as operating time and other beforementioned when we think they are relevant to a problem and we can base our predictions on them. We don't necessarily need to use all the features obtained from the measurement, e.g. the information what color the equipment is if it's not affecting the mechanical properties of the equipment. Feature extraction is about combining existing features to produce a more

useful feature such as when we have hardness values of two surfaces in contact, we can combine them and create hardness ratio between two surfaces. If we see that some other features can be added, we can create a new method of data collection from the experiment and add the feature to our dataset. Feature engineering also has methods to produce *feature importance* which tell us the relative importance of each feature in making a prediction, i.e. which features the model performance is based on mostly, assigning low score to un-important features and high score to important features. For different models are different techniques to calculate feature importance. In linear models, the coefficients assigned to each feature indicate their importance, the larger, the better. In decision tree based models, as Random Forest (RF) or Gradient Boosting Machines (GBM), feature importance is based on how frequently the feature is used to split to split the data across all decision trees in the ensemble. Other methods are *Permutation importance* and *SHAP* or also possibly Principal Component Analysis (PCA) which is primarily used as dimensionality-reduction technique but can indirectly highlight the importance of feature. [35][36][39]

1.5.4. Holdout validation, Generalization error

These features form the so-called "Feature matrix" usually noted X, which contains all the feature vectors x. Corresponding to each feature vector is target value. All target values make so-called "target vector", usually noted y. Supervised ML algorithm is then given input in form of feature matrix X and corresponding target vector y. [35][36]

To train a successful ML model, you must train the model on seen data and test the performance of the model on unseen data to see if the model generalizes well for unseen data, not just for training data. This process is called *holdout validation* and the result metric of performance on unseen data is called *generalization error*. Holdout validation is usually done by splitting the dataset into "*training set*" and "*test set*" as shown in Figure 32. This is usually done in ratio 60/40 or 80/20 with former being train set and latter test set. Sometimes, the split is performed into 3 different sets – third one being *"validation set"* which is used for tuning the hyperparameters or choosing the best model. A hyperparameter is the parameter of model that is not learned by the model itself in the process, but rather setting that controls the learning process. Examples of hyperparameter are the *learning rate* controlling the step size at which the model parameters are updated during training in Gradient Descent (GD) methods or *number of hidden layers* in a neural networks, both of which will be explained in dedicated sub-chapter later in this chapter.
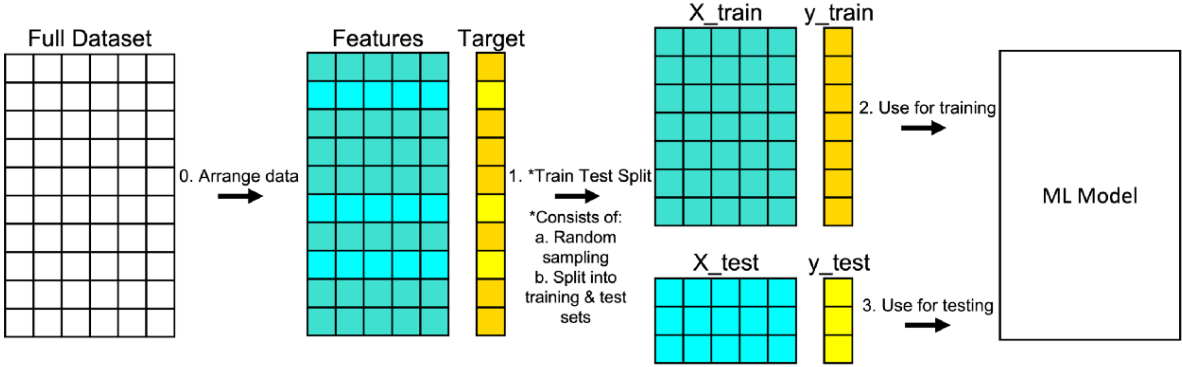

Fig. 32 - Splitting dataset into training and testing set [40]

The model is given X matrix from training set (X_train) with corresponding target vector y (y_train) and learns underlying patterns and relationships i.e. what in X_train "leads" to the values in y_train. Then the model is given X matrix from test set (X_test) and predicts target vector values based on what it learned from the X_train and y_train. Model obtains predicted vector of target values (y_pred). This predicted vector y_pred is then compared with real values of initial withheld target vector values (y_test). The model is then evaluated using various metrics to see how well it can predict values of unseen data. Metrics by which models are evaluated will be discussed later. [35][36][40]

Two different problems with generalizing the model on the unseen data can happen. Either *overfitting* or *underfitting* the data. Remember that the model is trained on the train set and tested using test set. We will show the two problems on the plotted curve, which was created as 2nd degree polynomial with noise. Therefore, 2nd degree polynomial regression should be best fit to regress the data and too small degree polynomial and too high degree polynomial will be used to show overfitting and underfitting. [35][36]
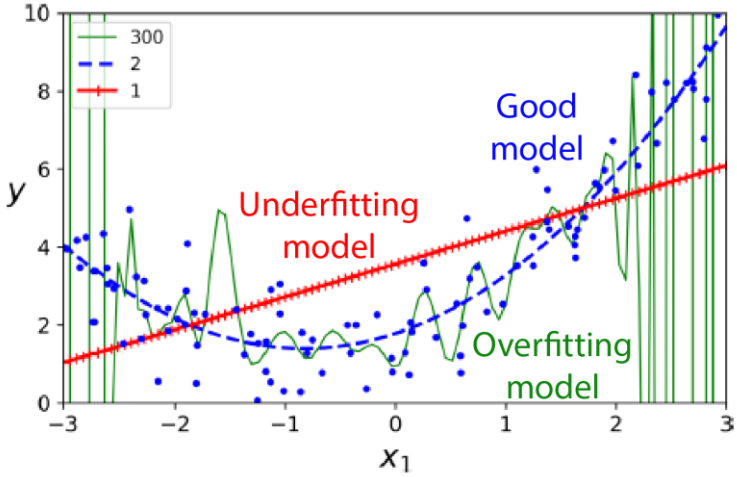

Fig. 33 - Overfitting, Underfitting and Good model of regression [35]

Simply put, if the model is too complex, such as high degree polynomial depicted in green in Figure 33, it fits the data well in the range of -2 to 2, but for bigger range, the polynomial goes into high values and predictions based on it would be very bad. This is clear sign of overfitting. We can recognize the problem of overfitting when the model performs well on the training data but poorly on unseen test data. Possible solutions to overfitting are to simplify the model such as selecting model with fewer parameters or applying *regularization* to take away some degrees of freedom from the model, gathering more training data or reduce noise in training data (fix errors and remove outliers). [35][36]

On the other hand, when we look at the linear regression curve in red in Figure 33, we can see the opposite problem. The linear model does not fit the training data well and would predict nonsense when it would be deployed on the test set. It is basically too simple to learn the underlying structure of the data, meaning that reality is more complex than this chosen model. Underfitting can be recognized when the model performs poorly both on the training set and test set. Underfitting can be fixed by selecting more powerful model with more parameters, feeding better features to the algorithm, or reducing the constraints on the model assigned via regularization. [35][36]

Models' generalization performance and hence also overfitting and underfitting can be detected initially by visual cues, such as our example with polynomials, or by more precise methods. Most notable of these methods are the use of *learning curves* to plot performance of model as a function of dataset size or training duration to see if model performs badly on the test data (overfitting) or badly on both train and test set (underfitting) and the use of *Cross-validation*. Usually performed as k-fold cross validation, with k being usually 5-10, the method parts the

dataset into multiple "folds" and iteratively trains the model on k-1 folds while testing it on the remaining one fold. This process involves iterating so each fold is once the test set and gives comprehensive evaluation of models' generalization performance over different test sets (folds). [35][36]

1.5.5. Performance metrics

As with every method in engineering and technology, some metrics to quantify the performance and properties of the method or its parts is necessary. Out thesis focuses on regression tasks, so the according metrics will be provided in detail. Most notable of them are *mean absolute error, mean squared error, root mean squared error, coefficient of determination (r-squared)* and *adjusted r-squared*. These metrics will be used later in this thesis to compare the performance of models created in the methodology chapter. Regarding classification as the second important task performed in supervised ML, most important metrics are *accuracy, precision, recall and F1 score* among others, but these are not important in the context of our use case, so they will be given no more attention in this thesis. [35][36]

Mean absolute error (MAE)

Calculates average distance (average absolute difference) shown in Figure 34 between the values predicted by model and the actual values. Average error is easy to count and straightforward to interpret but is less sensitive to outliers than MSE and RMSE. Calculation of MAE is shown in equation (22). y_i is actual value, $\hat{y}$ is predicted value and n is number of values. [41][42]
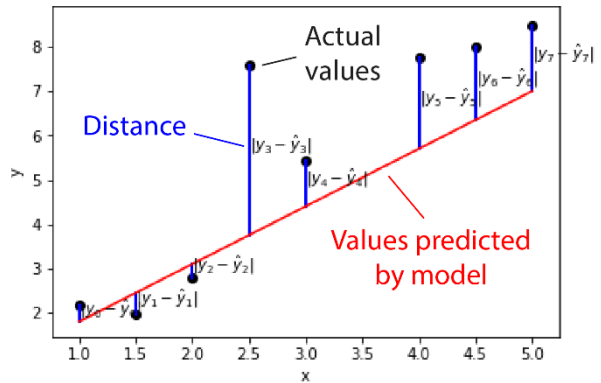

Fig. 34 - mean absolute error calculation schematic [41]

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y_i}|}{n} \tag{22}$$

Mean squared error (MSE)

Calculates average squared distance (average of squared differences) between predicted values and actual values. Calculation of MSE is similar to MAE, but every distance is squared before summing up as shown in equation (23). MSE is more sensitive to outliers because their larger distance from predicted value becomes even larger by squaring it. Units of MSE are the squared units of target variable y, which is less intuitive than to interpret. [43]

$$MSE = \frac{\sum_{i=1}^{n}|y_i - \hat{y_i}|^2}{n} \tag{23}$$

Root mean squared error (RMSE)

To provide a more interpretable measure of the mean squared error unit, the square root is made. This metric is called root mean squared error and shares similar properties as MSE but is easier to interpret. It is simply square root of MSE as shown in equation (24). [41][44]

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}|y_i - \widehat{y_i}|^2}{n}} \tag{24}$$

Coefficient of determination (R2)

Coefficient of determination, also called R-squared, represents the proportion of variance in the target variable that is explained by the model. Measuring how well the model fits the observed data. R2 values range from 0 to 1, with 1 being perfect fit and zero being no fit at all. Therefore, higher R2 suggests better model. R2 does not provide information about the direction or magnitude of errors. R2 is calculated from the sum of squared residuals (SSR) and total sum of squares (TSS) as shown in Figure 35 and equation (25). y_i is the actual value, $\hat{y}$ is predicted value, $\bar{y}$ is mean value and n is number of values. If we have good fit, as we can see in blue SSR, the SSR value is small, making the SSR/TSS ratio small and therefore we subtract small part from 1. 1 minus small SSR/TSS is close to one and that means good fit. [45]
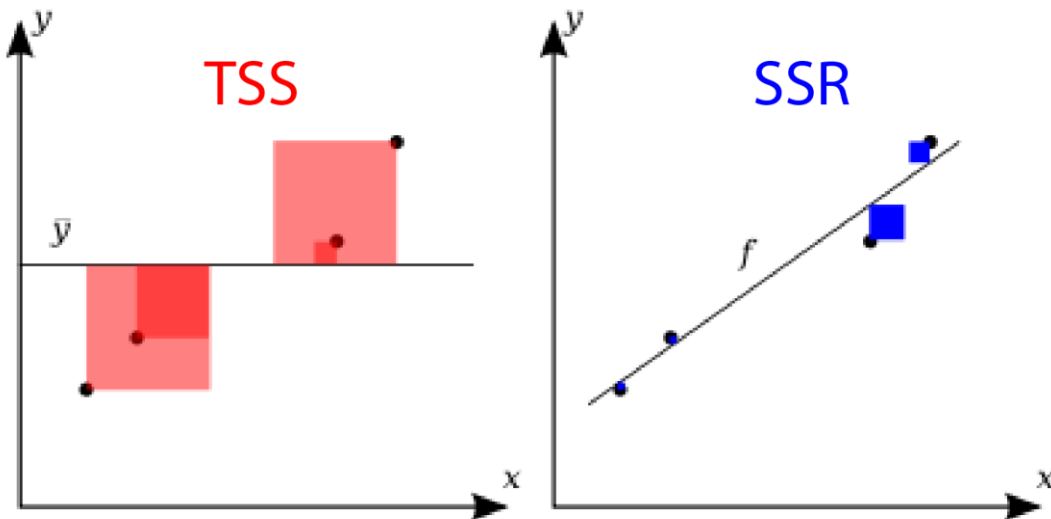

Fig. 35 - R2 schematic of total sum of squares and sum of squared residuals [45]

$$R2 = 1 - \frac{SSR}{TSS} = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{25}$$

Adjusted coefficient of determination (Adjusted R2)

Coefficient of determination can be also adjusted for number of terms we have $(x_1, x_2, ..., x_n)$. This is performed because classic R2 is automatically upgraded with each added term x_i but this addition does not necessarily enhance the model. Most used correction for calculating adjusted R2 is the correction by Mordecai Ezekiel taking into account the degrees of freedom of model and its variables as shown in equation (26). R2 is the classic coefficient of determination, n is number of values (sample size) and p is the total number of explanatory variables in the model. [45]

$$\overline{R2} = 1 - (1 - R2)\frac{n-1}{n-p-1} \tag{26}$$

Explained variance score (EVS)

EVS calculates the ratio between variance of the difference between the true and predicted values to the variance of true values. The resulting score ranges between -inf and 1 with score of 1 indicating perfect match and score 0 is indicating that the model does not perform better than predicting the mean of the true values. It is similar metric as R2, but the two are not the same. R2 focuses rather on overall goodness of the fit of model and EVS on model's predictive ability. Calculation of EVS is shown in equation (27). y is actual value, ŷ is predicted value. [46]

$$EVS = 1 - \frac{var(y - \hat{y})}{var(y)} \tag{27}$$

These metrics will be used to evaluate proposed machine learning models in the methodology chapter of this thesis and their performance will be compared based on them. Most important supervised learning models for regression task, hence our use case, are described in following sections.

1.5.6. Support Vector Machine

Support vector machine (SVM) is versatile powerful ML model for linear and non-linear classification, regression, and outlier detection. For classification tasks, SVM finds the line with greatest margins (largest street possible) that divides the data and sorts it into binary classes. The name of the method is derived from the *support vector* which is the edge of the street shown in Figure 36 as dashed line. Support vector regression (SVR) uses contrary approach – instead of trying to fit largest possible street between two classes while limiting margin violations as classification task SVM does, SVR tries to fit as many instances as possible on the street while limiting the margin violations. I.e. find the street that has the most instances on it and not many instances off it. This is schematically shown in Figure 36 since SVR is the primary concern for our use case regarding the SVM algorithms. The width of the street in SVR model is controlled by hyperparameter ε. Changin the value of ε hyperparameter is also shown in Figure 36. The higher the ε, the wider the street. [35][47]
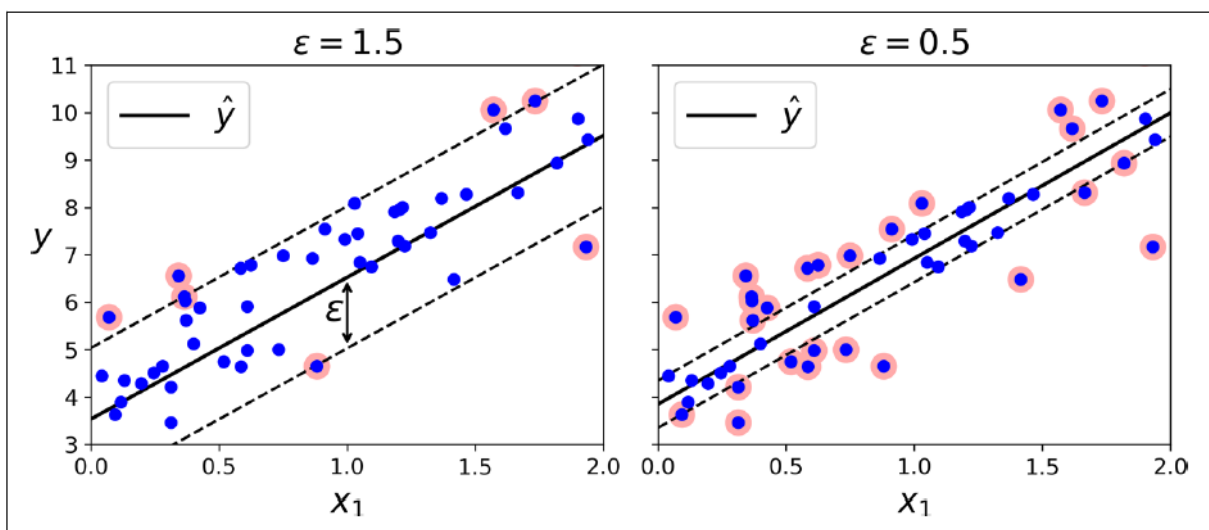

Fig. 36 - Principle of SVR and changing the ε hyperparameter for width of the street [35]

SVR supports both linear and nonlinear regression. Adding more training instances within the margin (on the street) does not affect the model's predictions, hence the model is *ε-insensitive*. In this thesis, we will use scikit-learn LinearSVR (linear regression) and SVR (polynomial regression) models. Both LinearSVR and SVR require scaling of the dataset, in our thesis done via scikit-learn MinMax Scaler. Time complexity of LinearSVR in big-O notation is roughly $O(m \cdot n)$ and time complexity of SVR is $O(m^2 \cdot n)$ to $O(m^3 \cdot n)$. [35][47]

For our use case, SVR from scikit-learn will be used. Defined parameters are:

- C – Regularization parameter. Strength of regularization is inversely proportional to C and penalty is squared L2.
- Epsilon – Width of the street
- Kernel – Kernel type used in algorithm – function that transform the data into a higher-dimensional space
- Gamma – Kernel coefficient for rbf, poly and sigmoid kernels

1.5.7. Random Forests

Random Forest (RF) is an ensemble method based on the combination of Decision Trees (DT). Decision Trees are ML algorithms for both classification and regression tasks. Decision Trees present a flowchart-like structure in which each node represents a decision point or test on the feature, as shown in Figure 37. If the test result is True, the outcome moves to corresponding node to next layer or to the other in False result. Initially, the entire dataset is put in as Root node, then multiple Internal nodes act as decision points which the value is sent through and in the end Leaf Nodes are the nodes for regression values which predict the target value. So, at each point, the node uses some criteria to decide which direction send the value next until it comes to the terminal Leaf node and presents the value a predicted. [35][48]
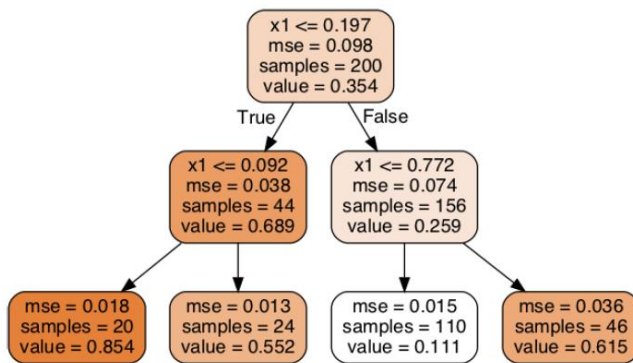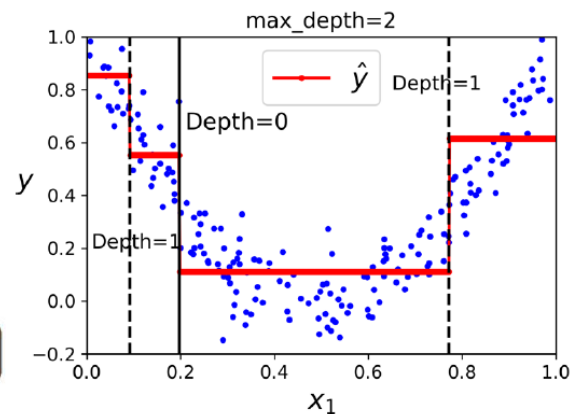

Fig. 37 - Decision Tree Algorithm [35]

Fig. 38 - Decision Tree fitting curve - average of region [35]

Random Forest (RF) is an ensemble method that builds upon the principles of Decision Trees to enhance predictive performance and robustness. It operates by constructing a multitude of Decision Trees during training time and outputting the class that is the mode of the classes (classification) or the mean prediction (regression) of the individual trees. The fundamental concept behind Random Forests is to introduce randomness into the model building process to create a diverse set of trees. This is achieved through two main mechanisms: bootstrapping and feature randomness. Bootstrapping involves generating multiple training datasets by sampling with replacement from the original dataset, ensuring each tree is trained on a different subset of data. Feature randomness selects a random subset of features for each split in the trees, promoting diversity among the trees and reducing the likelihood of overfitting. By aggregating the predictions of these diverse trees, Random Forests mitigate the variance inherent in single Decision Trees, leading to improved accuracy and generalization. Additionally, Random Forests provide an internal estimate of error, known as out-of-bag (OOB) error, and a measure of feature importance, aiding in the interpretability of the model. [35][48]

For our use case, RandomForestRegressor from scikit-learn will be used. Defined parameters are:

- N estimators – Number of decision trees in the forest
- Max depth – Maximum depth of tree for prediction
- Min samples split – Minimum number of samples required to split an internal node
- Min samples leaf – Minimum number of samples to be at a leaf node

- Min weight fraction – The minimum weighted fraction of the sum total of weights required to be at a leaf node
- Max leaf nodes – Maximal number of leaf nodes
- Min impurity decrease – A node will be split if this split induces a decrease of the impurity equal or greater than this value

1.5.8. Genetic Algorithm

Genetic Algorithm (GA) or Genetic Programming (GP) is a bio-inspired AI approach based on natural selection – the process that drives biological evolution, where only the fittest individuals survive and reproduce. It is used for both constrained and unconstrained optimization problems. GA repeatedly modifies a population of individual solutions to converge towards the best result. At each step, the GA selects individuals from the current population (parents) to produce the next generation (children). With each successive generation, the population evolves toward an optimal solution. In the context of GA, the population consists of a variety of potential solutions to the problem at hand, which are often represented as mathematical models or encoded as chromosomes. With each

Fig. 39 - Genetic algorithm [49]

generation, only the fittest individuals – those that best meet the criteria defined by the fitness function – are more likely to survive and contribute to the next generation. This is schematically shown in Figure 39. [49][50][51]
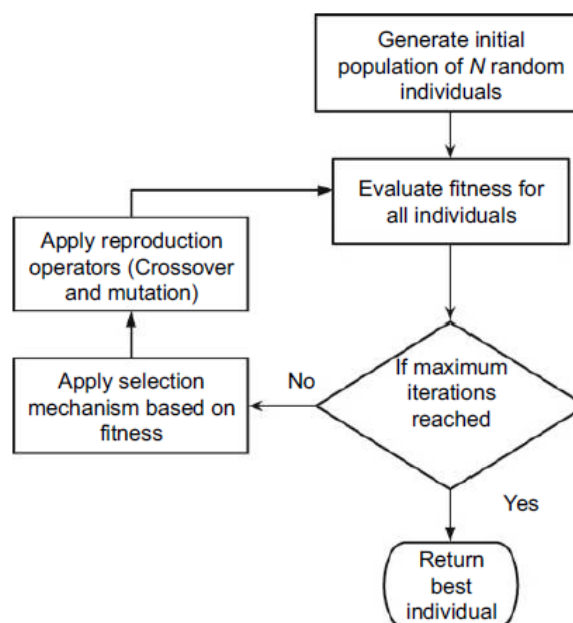
The creation of a new generation is governed by three main operators: Selection, Crossover, and Mutation:

- Selection: This operator selects the fittest individuals from the current population to act as parents for the next generation. These individuals are often referred to as the *elite*
- Crossover: Also known as recombination, this operator combines genetic information from two parent individuals to create one or more offspring. This mimics biological reproduction and allows for the mixing of traits.
- Mutation: This operator introduces random changes to individual parents, creating variation in the offspring. Mutation helps maintain genetic diversity within the population and prevents premature convergence on suboptimal solutions.

The algorithm begins with the creation of a random initial population. Subsequent generations are created based on the rules specified above. Each generation is evaluated using a *fitness function,* which scores how well each individual mathematical model solves the problem. The individuals with the highest fitness scores are more likely to be selected as parents (elite) for the next generation. [49][50][51]

The GA process continues iterating through generations until stopping criteria are met. These criteria are either reaching the set number of generations, reaching the early stopping threshold of fitness function i.e. the fitness function is "good" enough to stop iterating or other criteria. [49][50][51]

For our use case, Symbolic Regressor from GPlearn will be used. Defined parameters are:

- Population size: The number of mathematical models evaluated in each generation.
- Generations: The number of iterations the algorithm will run.
- Tournament size: The number of top models that compete to become part of the next generation.
- Stopping criteria: The fitness function threshold below which the model stops evolving.
- Max samples: The fraction of data drawn from the feature matrix to evaluate each model.
- Parsimony coefficient: A constant that penalizes large programs by adjusting their fitness function to be less favorable to selection.
- Function set: The set of functions used to build and evolve mathematical models, ranging from basic functions like addition, subtraction, multiplication, and division, to more complex functions like logarithms, sine, cosine, and tangent.

1.5.9. Artificial Neural Networks

Introduced in 1943 by McCulloch and Pitts, Artificial neural networks (ANN) or just Neural Networks (NN) are ML models inspired by the networks of biological neurons found in our brain. They are versatile, powerful, and scalable which makes them ideal for tackling large and high complex ML tasks such as classifying billions of images, powering speech recognition or beating the world champion at the game of Go. Since 1943, development of NNs has gone through different phases of development speed and today age is well supporting for them. There is large amounts of data on the internet, massive increases in computing power since 1990s to this day and also major fundings. In 2024, with AI boom already in motion, mainly because of Large Language Models (LLMs), it makes only sense that NNs might be beneficial for this thesis and its predictions. [35][52]

In the human brain, a neuron is a specialized cell that processes and transmits electrical and chemical signals, enabling communication between different parts of the nervous system. ANNs were inspired by this and *Artificial Neuron (AN)* is computational unit that processes input data using weighted connections and activation functions to produce output signals, mimicking the behavior of biological neurons. The simplest AN proposed by McCulloch and Pitts in 1943 has one or more binary inputs (on/off) i.e. one or more connections to other neurons, and one binary output. This serves well for logical operations. [35]

In 1957, Frank Rosenblatt invented a different form of AN called *threshold logic unit (TLU)* also sometimes called *linear threshold unit (LTU)* in ANN architecture called the *Perceptron*. In TLU, the inputs and outputs are numbers (instead of binary values) and each input connection is associated with weight. The TLU computes a weighted sum of its inputs, then applies a *step function* to that sum and outputs the result h_w, The schematic of TLU is shown in Figure 40. Step function is usually *heaviside function*
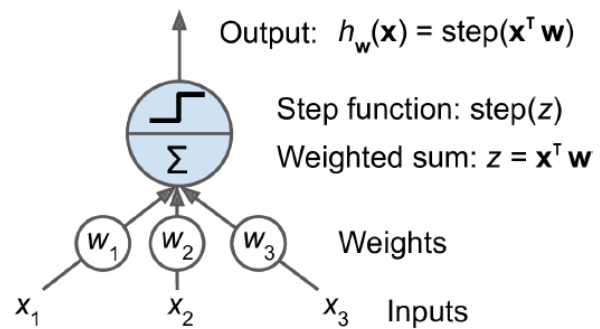

Fig. 40 - Threshold logic unit: an artificial neuron which computes a weighted sum [35]

which produces 0 if z < 0 and 1 if z ≥ 0. This single TLU can be used for simple linear binary classification – based on different inputs, TLU computes linear combination of them and if the result exceeds a threshold, it outputs the positive class, otherwise the negative.

Perceptron is then composed as a single layer of TLUs with each TLU connected to all the inputs. When all the neurons in a layer are connected to every neuron in the previous layer, the layer is called *dense layer* or *fully connected layer*. Perceptron also contains *input neurons*, which the inputs of Perceptron are fed in. Input neurons output whatever they are fed. There is also *bias neuron* that represents the extra bias feature and outputs 1 every time. Example of the perceptron architecture is shown in Figure 41. This architecture has two inputs, one bias feature and 3 outputs.
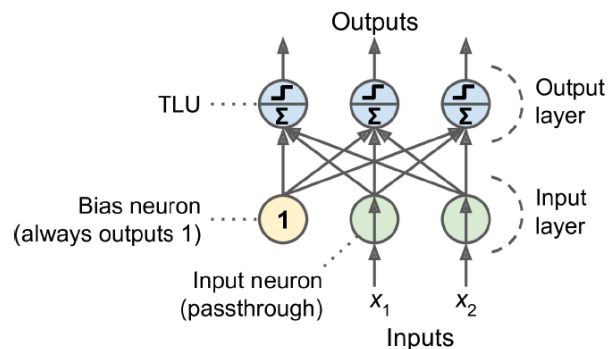

Fig. 41 - Single layer perceptron architecture with two inputs, one bias neuron and 3 outputs [35]

It can classify instances simultaneously into three different binary classes, making it a *multioutput classifier*. Outputs of dense layer of neurons is efficiently calculated using formula in equation (28) where X is the matrix of input features, W is weight matrix containing all the weight except for bias neuron, b is the bias vector containing all weights between bias neuron and all artificial neurons, ϕ is the *activation function* such as sigmoid, tanh or RELU. In the case of artificial neurons being TLUs, it is called step function as mentioned previously. The activation function will be discussed more promptly later. [35][52]

$$h_{W,b} = \phi(XW + b) \tag{28}$$

Perceptron is then trained by employing the *Hebb's rule* which suggests that when biological neuron triggers another neuron often, the connection between the two grows stronger. In the case of perceptron, every output neuron that produced a wrong prediction, connection weights for all the inputs that would provide the correct prediction are reinforced. Simple perceptron architectures are also unable to learn complex patterns because the decision boundary of each input neuron is linear or solve trivial problems such as XOR function. More on the topic is explained in various literature, but some of the limitations of simple single layer perceptrons can be eliminated by stacking multiple perceptron layers. This ANN architecture is called *Multilayer perceptron (MLP).* [35][52]

MLP is composed of multiple layers. One *input layer* (passthrough – outputs input), one or more *hidden layers* (TLUs – input 0 or 1 based on exceeding threshold) and one *output layer* (TLUs). This architecture can be seen in Figure 42. If the ANN contains a deep stack of hidden layers, it is called *deep neural network (DNN)*. The layers close to input layer are usually called *lower layers* and layers close to output layer are *upper layers*. Every layer except the output layer is fully connected to the next layer and includes a bias neuron. In this architecture, the signal flows
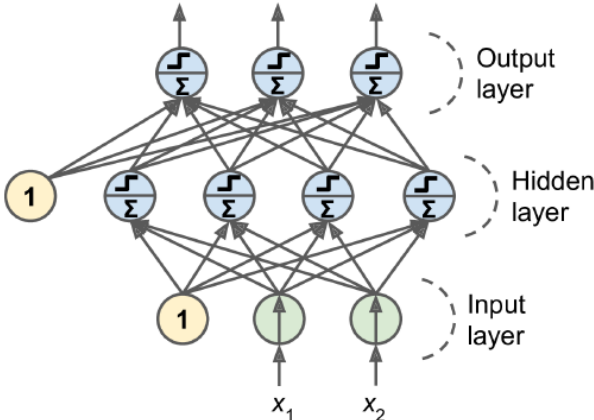

Fig. 42 - Multilayer perceptron with two inputs, one hidden layer of four neurons and output layer of three neurons [35]

only in one direction, from inputs to outputs, making it the example of *feedforward neural network (FNN)*.

Until 1986, training of MLPs was without success, until Rumelhart, Hinton and Williams introduced *backpropagation* training algorithm that is still used today. Backpropagation works like Gradient Descent – in one forward and one backward pass through the network, the backpropagation algorithm can compute the gradient of network's error with regard of every single model parameter i.e. find how each connection weight and each bias term can be tweaked to reduce the error. Then the regular GD step is performed and the whole process of backpropagation is repeated until the network converges to the solution. [35][52]

Gradient descent approach brought a key change to MLP architecture – the heaviside step function was replaced by activation function. The Heaviside function only has flat segments and the resulting derivative for gradient would be zero. On the other hand, activation functions such as *sigmoid*, *hyperbolic tangent (tanh)* or *Rectified Linear Unit (RELU)* function have better

properties regarding the differentiation and are therefore more suitable for this backpropagation (using gradient descent). [35][52]

Different activation functions present different valuable properties. Sigmoid (logistic) function has a well-defined derivative everywhere that allows GD to make progress in every step. Tanh function is also differentiable everywhere and its ranging from -1 to 1 tends to make output of each layer more or less centered around zero, which speeds up the convergence. RELU function is not differentiable at zero, but in practice works very well and is fast to compute. Because of not having maximum output value, RELU reduces some issues during GD. RELU is usually the default activation function nowadays. Alternative later used in the thesis is the Leaky RELU function – which instead of flat slope for negative values as RELU, has slightly decreasing slope for negative values. All of these functions and their derivatives can be seen in Figure 43. Without the activation function, i.e. without non-linearity between the layers, we wouldn't be able to solve very complex problems. The reason for that is, because the chaining of two linear functions f(x) and g(x) such that f(g(x)) gives you also the linear function. If this were the case, even deep stack of layers would behave as one single layer and the intention of having multiple layer network would be lost. As mentioned before, the activation function introduces non-linearity, by deciding if the neuron should be activated (when input to it exceeds the threshold of activation function) or remain inactive and if so, to what extent it should activate. [35][52]
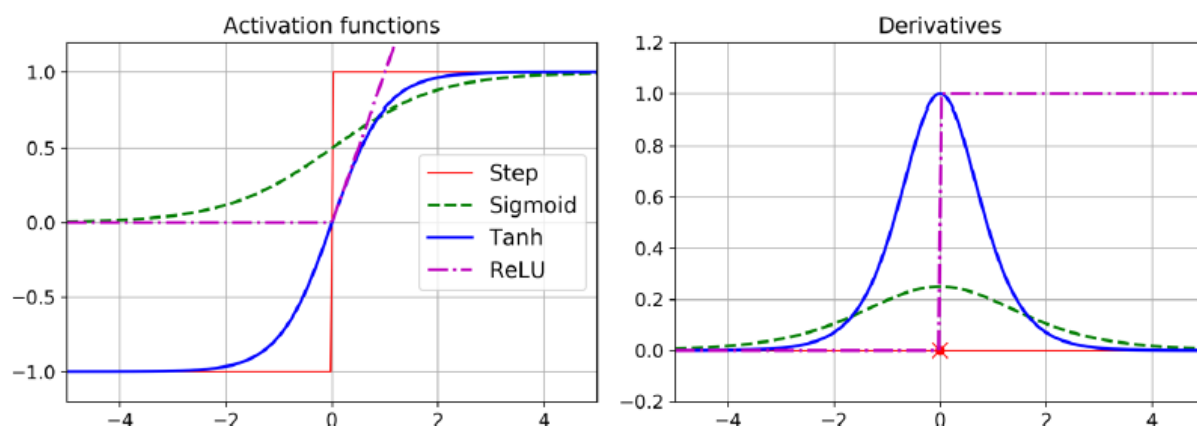

Fig. 43 - Activation functions sigmoid, tanh and RELU and heaviside step function with their derivates [35]

Now that we know what MLP is, Figure 44. shows the typical regression MLP architecture.

Hyperparameter	Typical value
# input neurons	One per input feature (e.g., 28 x 28 = 784 for MNIST)
# hidden layers	Depends on the problem, but typically 1 to 5
# neurons per hidden layer	Depends on the problem, but typically 10 to 100
# output neurons	1 per prediction dimension
Hidden activation	ReLU (or SELU, see Chapter 11)
Output activation	None, or ReLU/softplus (if positive outputs) or logistic/tanh (if bounded outputs)
Loss function	MSE or MAE/Huber (if outliers)

Fig. 44 - Typical regression MLP architecture [35]

*MNIST stands for Modified National Institute of Standards and Technology database of handwritten digits that are 28x28 pixels and the dataset is used as a worldwide machine learning benchmark.

As Figure 44 shows, the Neural Networks have many hyperparameters, which offers good flexibility but also is one of their main drawbacks – many hyperparameters to tweak. Another major drawbacks are the "black box" nature of NNs, tendency to overfit, computational intensity and as with all ML models – dependence on the data quality and quantity. On the other hand, they offer many advantages such as the ability to model nonlinear relationships, learn features itself, scalability, universal approximation, adaptability and parallel processing.

Back to hyperparameters tweaking, there are multiple options to find the hyperparameters for best performing NN. One option is to try many combinations of hyperparameters to see which works best on validation set or use K-fold cross-validation. This can be done via GridSearchCV or RandomizedSearchCV to explore the hyperparameter space. GridSearchCV is good for small hyperparameter space as it tries the parameters and values that you specify. RandomizedSearchCV is better for large hyperparameter spaces, since it evaluates a given number of random combinations by selecting a random value for each hyperparameter at every iteration. RandomizedSearchCV is better for NNs. Randomized search is good for problems with small feature number and small dataset, otherwise it is very time consuming. Good practice is to run randomized search on wide range of parameters and then run another one on the region that performed the best in hyperparameter space. Some dedicated Python libraries, such as *Hyperopt*, do the zooming process and focus the hyperparameter search on the good performance regions in hyperparameter space. [35][52]

Most important hyperparameters are number of hidden layers, number of neurons per hidden layer, learning rate, optimizer, batch size, activation function, number of iterations and regularization techniques. The following hyperparameters will be concisely explained in the following paragraph. [35][52]

The number of hidden layers depends on the task complexity. Many problems get reasonable results with one or two hidden layers in ANN but for more complex problems, the better. If the problem is complex, ramp up the number of hidden layers until the model starts overfitting. The number of neurons in the input and output layers is determined by the type of input and output. One neuron per one input/output. In our case, 3-7 neurons in input layer for non-constant parameters of experiment or calculated features and one neuron in output layer that is predicted wear volume. The number of neurons in hidden layers needs to be set. Nowadays, most architectures are made with the same number of neurons in the hidden layer and sometimes the first hidden layer is bigger. It is better to increase the number of hidden layers instead of neurons in the hidden layer. In practice, it is also better to choose a more complex model with more hidden layers and neurons than you need and then apply early stopping or other regularization techniques to prevent overfitting. The learning rate as mentioned earlier is the step size in optimizers (e.g. GD), which influences the rate at which the model updates its parameters and arguably the most important hyperparameter. In general, the optimal learning rate is about half of the maximum learning rate i.e. the learning rate above which the algorithm diverges. A good approach is to train model for few hundred iterations with learning rate ranging from 10e-5 to 10, plot loss-learning rate and setting the optimal learning rate of the model to the value before the loss in plot starts to climb up. Optimizer is an algorithm that updates the model's parameters based on the gradients of loss function, with the learning rate value being a step size. Simple optimizers are the Stochastic and Mini-batch gradient descents, but better performing optimizers such as ADAM, RMSProp or AdaGrad can be used. Batch size specifies the number of training samples processed in each iteration or mini-batch during training. Suggested strategy by the author is to try a large batch size, using learning rate warmup and

if the training is unstable or the final performance is disappointing, use small batch size instead. Activation functions were discussed earlier. In general RELU function will be good default for all hidden layers and the default function for output layer depends on the task. Number of iterations known as *epochs* denotes the number of times the entire training dataset is passed forward and backward through the network during training. Epochs usually doesn't need to be tweaked, use the early stopping instead. [35][52]

Neural Networks, especially Deep NNs usually have lots of parameters which gives them an incredible amount of freedom. As mentioned before, it's great for fitting complex datasets but prone to overfitting. To counter this, *regularization* is applied. Most notable regularization techniques are Early stopping, L1 and L2 Regularization, Dropout and others. Early stopping regularizes the iterative learning algorithm by stopping the training as soon as the validation error reaches a minimum i.e. the validation error would want to climb higher which would suggest overfitting. L1 and L2 are techniques to penalize large weights in NNs cost function (MSE, RMSE etc.). L1 adds a penalty term proportional to the absolute value of the weights (L1 norm) to the cost function and L2, sometimes referred to as *weight decay*, adds a penalty term proportional to the square of the weights (L2 norm). Dropout is the technique of shutting down random neurons during certain training step (setting them to zero) which prevents the network from relying too much on individual neurons and therefore promoting better generalization. Dropout is controlled by parameter *dropout rate* which is the probability for each individual neuron being dropped, usually set between 10-50 %. [35][52]

1.5.10. Machine learning research in tribology

Since Holm, Archard and others began experiments and theoretical exploration regarding predicting the wear in 40's and 50's of 20th century, lot of technology evolved. In recent years, few researchers have delved into the combination of predicting wear and machine learning. This part of the thesis is a short summary of their discoveries, which will be later used as a starting point in methodology chapter to develop the most suited ML model for our use case. Most notable articles related to our use case are presented below in chronological order.

In 1997, Jones et al. performed combinatorics experimental designs and employed neural network model, both to reduce the number of experiments and getting data without performing certain experiments during investigation of different material/mechanical systems for friction and wear observation. The article is focused on showing feasibility of neural networks to predict life data and defining which input variables will influence the tribological behavior of the material/mechanical system. Jones et al. investigated 15 different neural network architectures on three different test rigs: rub shoe rig, pin-on-disk rig and four-ball rig. The article shows that the discussed models have been capable of predicting the wear rates regardless of the lubricants used, that the models are able to interpolate and extrapolate approximate wear rate values for conditions not run experimentally. The overall best architecture proved to be the Input layer dampened recurrent network. Linear scaling functions and hyperbolic tangent or logistic activation functions were beneficial. [53]

In 2000, Velten et al. extended the work of Jones et al. by measuring a total of 10 input variables/experimental conditions on a total of 72 wear volume measurements. They also reduced dimensionality by principal component analysis, improving predictive capabilities of NN by using Bayesian regularization instead of Early stopping and therefore also identifying the optimal size of NN layers. Velten et al. used a large number of randomly chosen test datasets and obtained worse average results than Jones et al. but still reasonable according to them. Out of 72 measurements, 10 were used as test set and the rest as train set. It is shown, that for this amount of measurements, the satisfactory prediction of wear volume data by feed-forward NNs can be expected only for a subset of any new data set, not for all data, and proposed that this probably holds true also for other NN architectures as well. [54]

In 2015, Radhika et al. performed dry sliding experiments on aluminium/alumina/graphite hybrid metal matrix composite on pin-on-disc tribometer and varied parameters such as load, velocity, distance, and Al composition. Using this empirical data, they obtained regression equation using NNs and optimized it using Genetic Algorithm (GA). A total number of 81 experiments was conducted and based on it, the optimal NN architecture of 3-10-1 was created. Data is normalized and split into 70/20/10 train/valid/test datasets and network trained by trial and error until satisfactory regression was obtained. Overall regression was $R^2 = 0.9046$ with curve shown in Figure 45. This NN is then fed into GA as the fitness function, iterates 51 generations and the algorithm produces single optimum best

Fig. 45 - Overall regression plot of Radhika et al. NN+GA approach [55]

solution of lowest wear rate for following parameters: Load = 33.8 N, Velocity = 2.05 m/s, Distance = 1500 m and Al composition = 5,49 wt. % for lowest wear rate = 4.41 mm3/m 10^-3. [55]

In 2020, many authors conducted research in the field of modeling wear with machine learning. Most notable of them will be mentioned in following paragraphs. Thankachan et al. predicted and analyzed dry sliding wear rates on novel copper-based surface composites. They deposited Boron nitride particles on the surface through friction stir processing from 5 to 15 vol% BN and proved that BN particles greatly reduce wear rate. During low load conditions, analysis of worn out surface revealed mild adhesive wear and during high load conditions, abrasive wear. Then, the NN with FFBP model with topology 4-7-1 was modeled and predicted good agreement with

experimental outcomes. The varying parameters were BN volume percentage [vol%] = A, load [N] = B, sliding velocity [m/s] = C and sliding distance [m] = D. Authors used MINITAB software to develop a general regression equation for wear rate [·10-5 mm3/Nm] based on the specified parameters as shown in equation (29). $R^2 = 0.89$ for given model. The NN model was trained on 24 values and tested on 3 values. The value of nodes present in hidden layer iterated from 2 to 15 and 7 hidden nodes was found to produce lowest value of MAE. This observation leads to 4-7-1 architecture of the NN, 4 being the number of input features, 7 being the number of hidden layers giving the lowest MAE, hence best prediction, and 1 being output
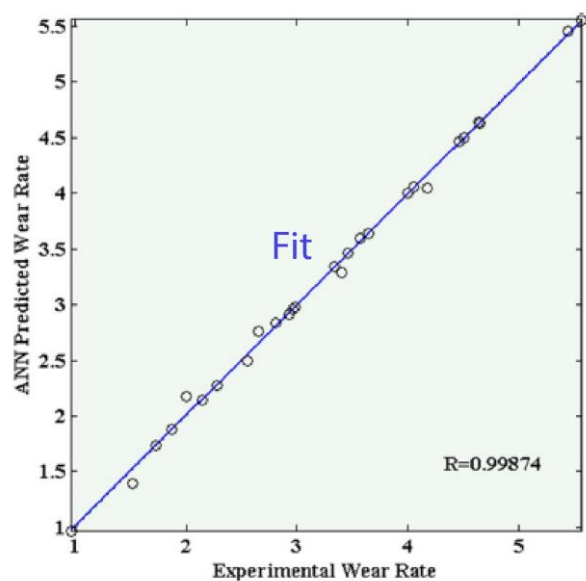

Fig. 46 - Thankachan et al regression using NN 4-7-1 with FFBP on wear rate for boron nitride deposited on surface [56]

layer. This NN for this use case of investigation Wear rate based on BN particles vol% on surface, load, velocity and sliding distance, produced R2 = 0.9987, as shown in Figure 46, which is almost ideal and model is therefore very good. [56]

$$Wear\ rate = -0.209 - 0.109 \cdot A + 0.1 \cdot B + 1.13 \cdot C + 0.00087 \cdot D\ by\ [56] \quad (29)$$

Many authors in the field of tribology use for their calculations Specific Wear Rate, which is the average value of wear volume divided by load and sliding distance. Hence wear rate is not the same during the whole experiment, Argatov et al. decided to use the so-called True Wear Coefficient which is regarded as a function of the sliding distance. Hence, they use integral and differential form of Archard equation for wear rate to develop NN for the predictions. Most relevant plot to our use case is shown in Figure 47. In part (a), it shows wear volume as a function of sliding time. Real measured data in red, sample of Polished WC-CoCr, compared with well-fitting ANN with 1-3-1 architecture and ill-fitting basic exponential model. A similar plot is shown in (b), where wear volume rate is plotted and also ANN variant with Finite-difference derivative and L·SWR predictions are added. L·SWR stands for Load multiplied with Specific wear rate = W/Ls. In (b), all predictions are quite close, apart from basic calculation of L·SWR, which produces larger values of wear volume rate. Argatov et al.

Fig. 47 - Argatov et al. comparing wear volume (a) and wear volume rate (b) as a function of time based on different measurements and predictions. [57]

proposed that the developed ANN model can be utilized in studies of wearing-in period which they defined as the initial time interval during which the TWC's value is stabilized. [57]

Faris et al. preferred insight into the underlying system and easier evaluation before predictive power, so they opted for Genetic algorithm (GA) instead of NNs. They used GAs for predicting quantities of adhesive wear for low and medium carbon steel. The resulting model is a function of the load, sliding speed and sliding time based upon pin-on-disc configuration of tribo test. Genetic programming was set with following parameters: Population size = 1000, Maximum generations = 200, Mutation rate = 15%, Tournament selector as selection mechanism, Elites = 1 and Operators = {+, -, ·, /, sin, cos, tan}. The train-test split is 70/30. The equations, coefficients and performance in form of VAF (authors used EVS formula explained before but named as VAF – Variance Accounted For and multiplied x100 to get result in percentage) and RMSE are shown in Figures 48 and 49. The superiority of GP approach over Linear regression is shown in article as well. [49]

$$Wear_{low} = (t \cdot s \cdot (c_0 \cdot s + \sin((c_1 \cdot t + c_2 \cdot s))) \cdot (c_3 \cdot s + c_4 \cdot t + \tan(c_5 \cdot s) \cdot c_6 + \tan(c_7) \cdot c_8) \cdot c_9 + c_{10}$$

$c_0 =$	0.27029	$c_6 =$	−1.0
$c_1 =$	0.95369	$c_7 =$	1.4368
$c_2 =$	−0.41068	$c_8 =$	1.2596
$c_3 =$	0.75134	$c_9 =$	$1.8049E-10$
$c_4 =$	0.95369	$c_{10} =$	0.00053583
$c_5 =$	1.0216		

Best performance:	Testing performance:
VAF = 99.5 %	VAF = 90.6 %
RMSE = 3.23e-4	RMSE = 1.5e-3
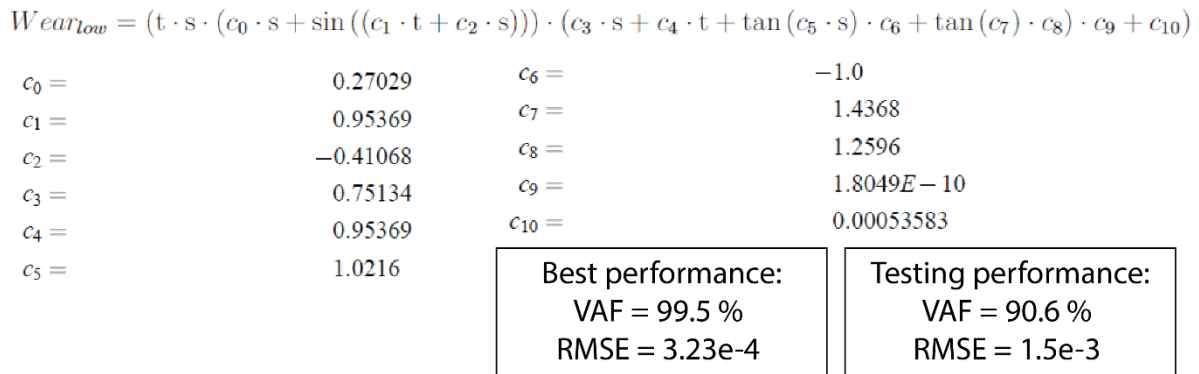
Fig. 48 - Genetic algorithm by Faris et al. for low carbon steel [49]

$$Wear_{medium} = \left(\frac{s \cdot t \cdot c_0}{\sin((c_1 \cdot 1 + c_2)) \cdot \cos(\cos((c_3 \cdot s + c_4 \cdot 1))) \cdot c_5} + c_6 \right)$$

$c_0 =$	$-8.9815E-08$
$c_1 =$	−1.1003
$c_2 =$	8.0088
$c_3 =$	1.827
$c_4 =$	−1.1003
$c_5 =$	0.2108
$c_6 =$	−0.0005508

Best performance:	Testing performance:
VAF = 94 %	VAF = 94.3 %
RMSE = 7.34e-4	RMSE = 6.06e-4
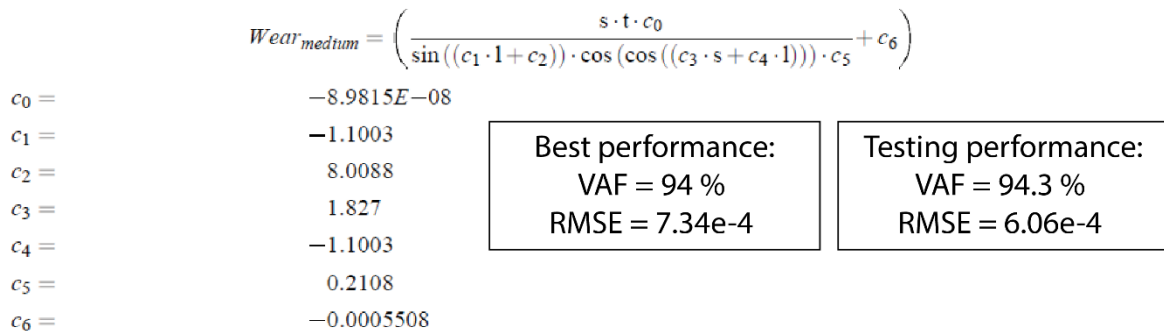
Fig. 49 - Genetic algorithm by Faris et al. for medium carbon steel [49]

Altay et al. performed a total of 99 measurements on 11 specimens of AISI1020 steel with coatings applied by plasma transfer arc welding and later applied LR, SVR and Gaussian process regression (GPR) to predict wear quantities. The following features were chosen: powder composition for plasma arc deposition, average micro-hardness of the coating layers, applied load and sliding distance. They used interaction function LR version of the LR algorithm, cubic function in kernel in SVR algorithm and rational quadratic function in GPR algorithm. Cross validation using k-fold with k = 5 was used. Metrics to compare were MAE, RMSE and R2. Linear regression performed the value R2 = 0.93. SVM and GPR were slightly better with R2 = 0.96. Although the R2 values for SVM and GPR were the same, the GPR was chosen as a better algorithm because of its potential to create an even better performing model. [58]

In 2022, Algur et al. used ML algorithms to predict wear performance of modified ZA-27 alloy in dry sliding. They varied normal load, sliding speed and sliding distance in the total number of 75 experiments performed in room temperature. Then, they used RF, GPR, K-nearest neighbors (KNN), SVR and LR algorithms to predict the wear loss. Load, speed and sliding distance are the features and wear loss as the target value. The train-test split

Table 3 - Evaluation of ML models by Algur

Model	R2 [-] train set	R2 [-] test set
RF	0.987	0.943
GPR	0.903	0.866
KNN	0.848	0.761
SVR	0.809	0.771
LR	0.753	0.745

is 70/30. In GPR, rational quadratic function is used. In SVR, cubic kernel function is used. Bagging technique for RF regression and interaction function is applied in LR algorithm. Cross-validation using k-fold with k = 10 was used. Metrics compared are MAE, RMSE and R2. R2 values are shown in Table 3. Results show that all the constructed models R2 values are close for train and test set, which signals their good fit and good ability to generalize. Best performing model for the use case of Algur et al. regarding R2 was RF – Random Forest with R2 = 0.987 [59]

In 2023, Rajput et al. applied ML approach on data from various journals and theses. They gathered a total of 300 data points with 23 independent input parameters and material loss in mm3 as output parameter. Input variables consist of pin sample wt% of constituents, structural ingredients in vol% such as ferrite, bainite and other structures of steel. Other parameters were pin hardness, temperature and time of heat treatment and finally working temperature, sliding speed, load and sliding distance. Authors decided to exclude parameters such as CoF, humidity, uniform and total elongation, temperature

Fig. 50 - Comparison of actual and predicted material volume loss by ANN by Rajput et al [61]

increase by friction because scarcity of data in these variables. Neuromat tool [60] with train-test split 80/20 was used. Data is initially normalized and then ANN model is built. Activation function h is hyperbolic tangent. Authors use various metrics to find optimal number of hidden units (HU) in NN. Based on this optimization, the best model (lowest error) was incorporated with four hidden units and four seeds. Figure 50 shows the performance of this model on unseen data. For most of the datapoints, the acceptable correlation between experimental and predicted
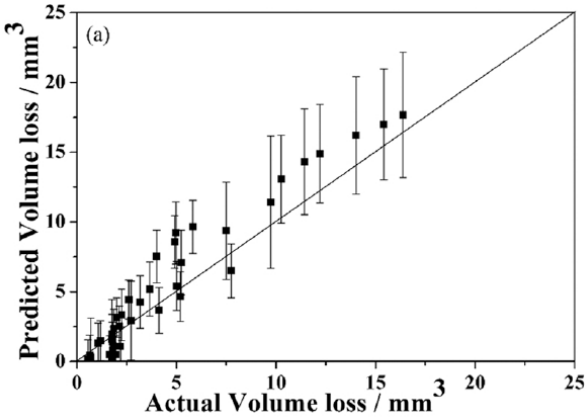
values is visualized. Larger error bars indicate that the model is not entirely-self-confident to make predictions and more experimental work needs to be done. The difference calculated between these values ranges from 0.3 to 4.2 mm3, which authors deem acceptable. Rajput et al. arrived at satisfactory result regarding the correlation between predicted and experimentally obtained results. [61]

This year, January 2024, Zhu et al. performed an experiment using ball-on-disk setup with lubrication. They varied parameter as sliding speed, sliding distance, normal load, temperature and oil film thickness and obtained 81 datapoints with wear depth as dependent parameter. The train-test split is 75/25. They employed four different ML algorithms: RF, KNN, SVM and Extreme Gradient Boosting (XGB). Optimization for RF, KNN and SVM models is done via least squares method to minimize errors between data and regression line estimation. XGB method optimizes the objective

Fig. 51 - Best performing model (XGB) by Zhu et al. [62]

function itself. Evaluation metrics were MAE, MSE, RMSE and R2. Results were validated using k-fold validation with k = 5. Results are shown in Table 4. Conclusion is that SVM and KNN models are not very accurate, but RF and XGB give satisfactory results. XGB produced best generated R2 = 0.88 and RF performed almost identical with R2 = 0.84. Actual vs predicted wear depth for the best performing model – XGB - is shown in Figure 51. [62]

Table 4 - Evaluation of ML models by Zhu

ML model	MAE	MSE	RMSE	R2
KNN	4.54	38.41	5.69	0.68
RF	3.25	18.3	4.14	0.84
SVM	7.29	87.2	8.81	0.29
XGB	2.95	14.48	3.62	0.88

Zhu et al. also performed feature importance analysis on its input values. Result is shown in Figure 52. Most important feature is sliding distance and then the thickness of lubricating film. Authors also conducted Pearson correlation analysis in the form of heat map and Taylor diagram in order to evaluate the best model in the proposed model. As said before, the XGB model was chosen as the best. [62]
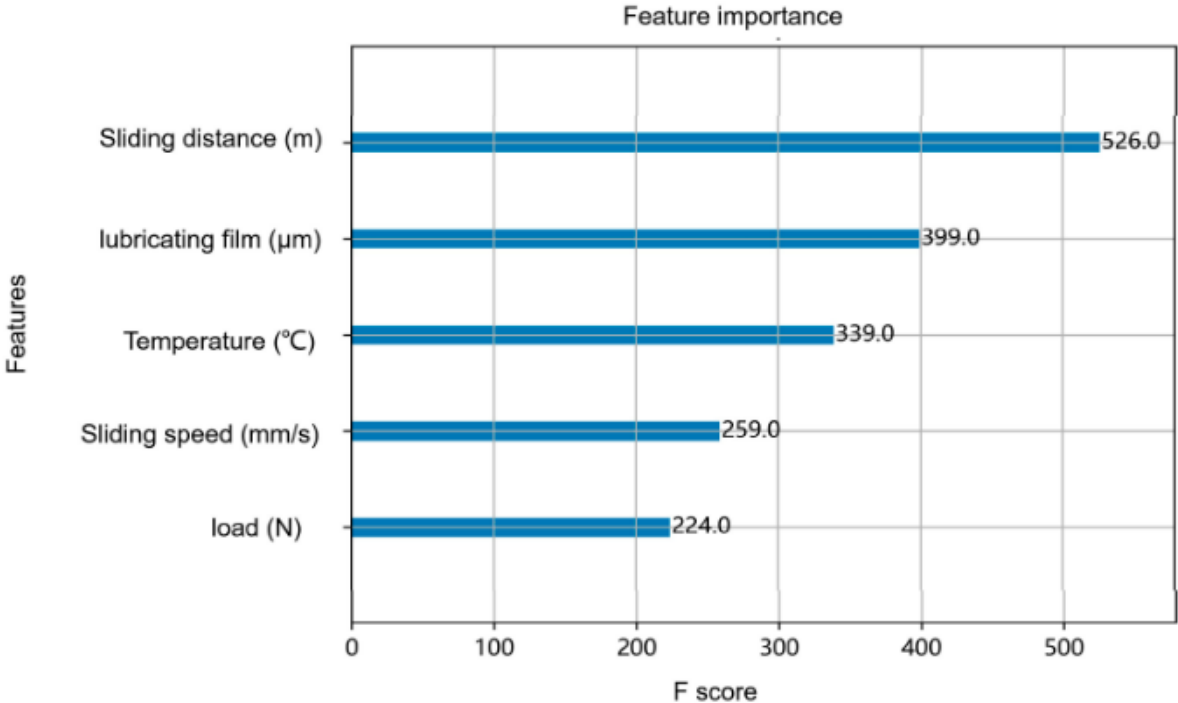
Fig. 52 - Feature importance analysis by Zhu et al. [62]

2. Specimens, Instruments and Software

Experimental data were obtained from company Garrett Motion Czech Republic s.r.o. and author is very thankful to them for the opportunity of making master thesis there. Garrett Motion is a worldwide company primarily specialized in development and engineering of turbochargers for automotive industry from small personal cars to large vehicles and industrial equipment.

The main goal of this thesis is to predict wear behavior in turbocharger kinematic mechanism, concretely rotational joints regarding the Arm and Bushing internal kinematic (green) and External kinematic (red) as shown in Figure 53. In these joints, small rotational motion is present and over time, the surfaces in contact are worn out.
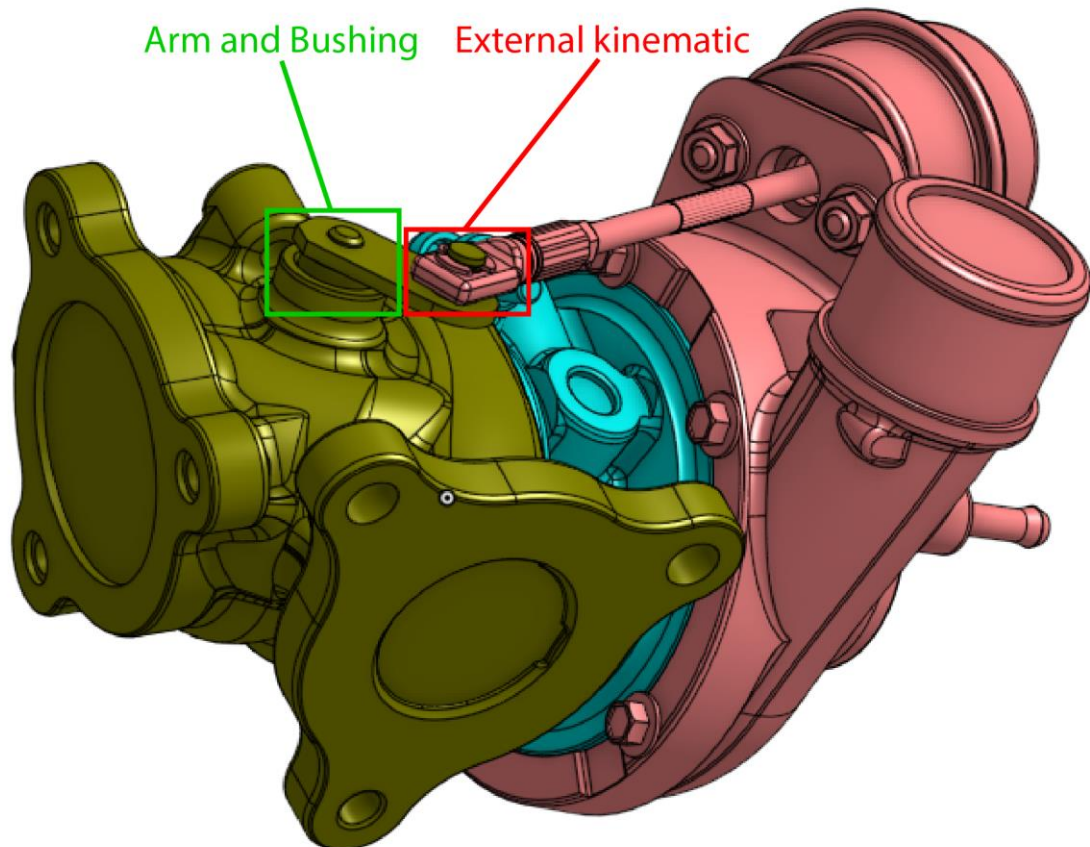
Fig. 53 - Turbocharger and shown kinematic parts, which wear behavior we are interested in. Arm and bushing in green and External kinematic in red.

2.1. Tribometer DN55 - Testing method principle, Specifications

To simulate wear behavior for a given case explained above, experiments are conducted on DN55 tribometer. The principle of it is shown in Figure 54. Specimens are the two plates and one pin in the middle. Plates are always of the same material. Material couple is therefore formed of two materials – one for pin and other for plates. Pin and plates are put into the arms of machine to conduct the experiment. The whole chamber of furnace is heated to a certain temperature up to the maximum temperature noted in the machine specification. After the chamber is heated to the temperature of the experiment, force is applied from the plates on the sides putting it into contact with the pin in the middle from both sides. Mechanism of force production is automatically set to keep the applied force the same during the whole experiment. In this moment, experiment can begin. The middle arm starts to produce cyclic motion up and down in the given range. This experiment is conducted for a certain time and as a result, we get worn pin and plates to some degree. So, to summarize, pin and plates in contact are subject to cyclic sliding motion for certain sliding distance with certain normal applied force at certain temperature.
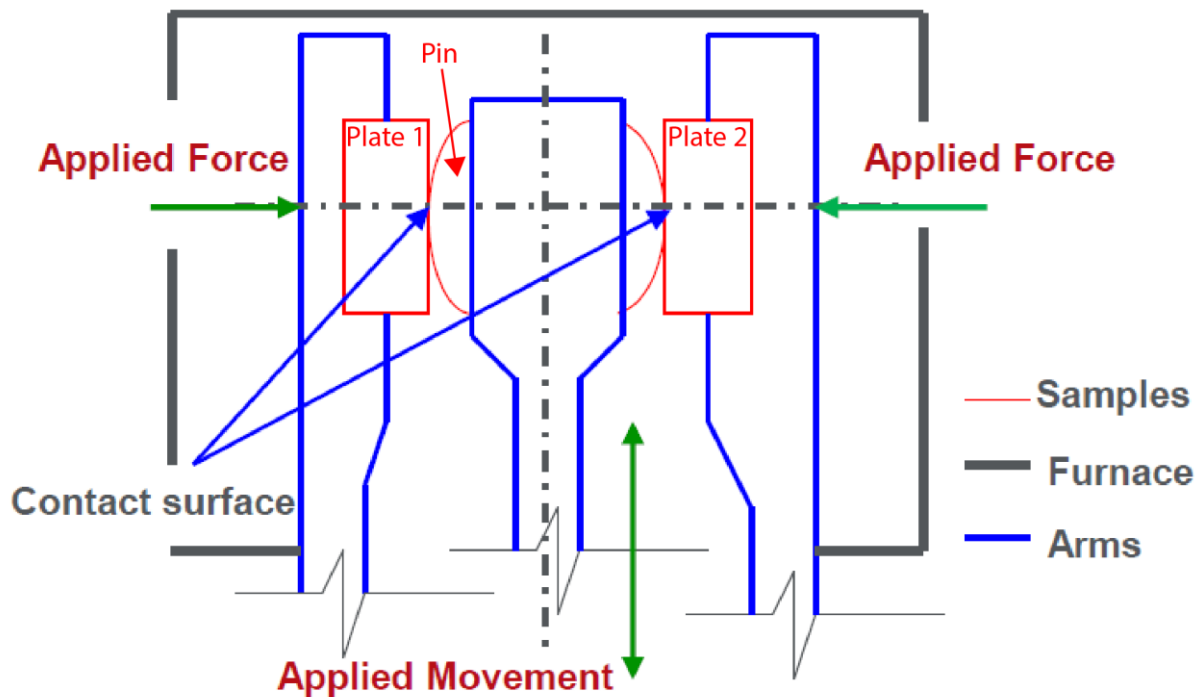
Fig. 54 - Testing principle of DN55 Tribometer

DN55 Tribometer can be seen in Figure 55. Parameters the DN55 can provide for experiments are noted in Table 5. Also, we must note the ability to create test programs to most suit the user needs and measure CoF-time charts. After the specimens, pin and plates, are subject to experiment, the surface in line contact looks like the surface shown in Figure 56.

Table 5 - Parameters of DN55

Parameter	Range
Temperature [°C]	Up to 1000
Load [N]	18-200
Stroke range [mm]	0.01 - 20
Frequency [Hz]	1 – 300
Type of contact [-]	Line, Point

Fig. 55 - DN55 Tribometer

Fig. 56 - Specimens, pin (a) and plates (b) after experiment on DN55 Tribometer

2.2. Laser profilometer Zygo NewView 300

After the pin and plates are subject to wear experiment on DN55 Tribometer, they are brought to laser profilometer. In Garrett Motion company, the profilometer used is Zygo NewView 8300 as shown in Figure 57. It is based on the principle of 3D coherence scanning interferometry. The scanner being precision piezo drive with closed loop capacitance gauge control and crash protection. Available objectives for magnification are from 1x up to 100x with turret

Fig. 57 - Laser profilometer Zygo NewView 8300 [63]

head for 4 objectives. Field of view is 0.04 to 16mm and Z-drive is 100mm range with 0.1 μm. The vertical scan range is 150 μm with precision piezo drive or 20 mm with extended scan. The repeatability of surface topography is 0.2 nm and repeatability of root mean square is 0.01 nm. More specifications can be found in [63].

The profilometer can measure 3D profiles such as shown in Figure 58. This is the worn specimen after the DN55 tribometer experiment. Also, the 2D profile in different places across the 3D profile can be obtained such as shown in Figure 59.
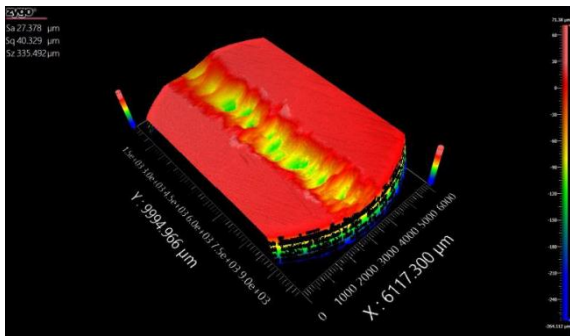
Fig. 58 - 3D profile of a specimen after tribometer test
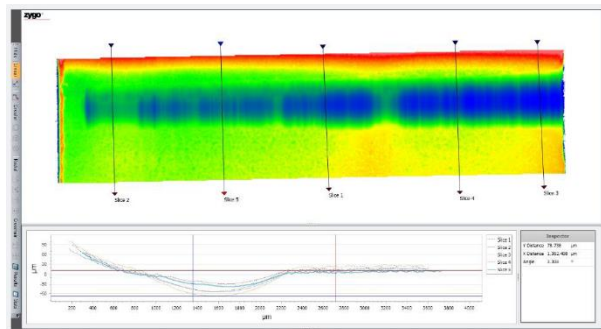
Fig. 59 - 2D profile the profilometer can measure

Based on this, the profilometer measures various surface profile properties, but our object of interest is the calculation of the wear volume, meaning total displaced volume of material. Total displaced volume is calculated as Negative volume + Positive volume such as shown in Figure 60. The zero line is made on the profilometer after the experiment. Reference plane is set in the area which was not subject to wear experiment and serves as zero line for the calculation of displaced volume in the area subject to wear experiment.

Fig. 60 - Total displaced volume in schematic representation and real surface profile

All wear volume values mentioned in this thesis are the product of this measurement of the laser profilometer.

2.3. Laptop parameters, Software, IDE

Methodology of creating the model was conducted on Acer Aspire E15 (E5-572G-74LM) laptop with specifications found in [64]. Chosen programming language is Python with various libraries. Most notably Pandas and Numpy for data management, Scikit-learn, TensorFlow and Keras for machine learning, Hyperopt for hyperparameter space search and Matplotlib and Seaborn for visualization. The online platform Google Colab [65] was used as a computing platform and interactive environment to perform ML modelling with better resources.

3. Methodology – Goryacheva Equation

The Methodology chapter is about exploring the potential methods of wear prediction and modelling possibilities based on available data. Various methods will be used with the intention to find relationships between data, correlations and develop a methodology for predicting the wear coefficients K, α and β based on material characteristics and experiment parameters. The first of these methods used will be testing if the Goryacheva wear equation from the literature hold some significance in our use case.

Experiments were conducted with the parameters specified in Table 6. For best prediction, only some of the experimental values were chosen. Main criteria were for materials to be of metallic character (no ceramics, no composites) and to be homogenous or with heat treatment finish (no coatings). Three sets of data were obtained. Firstly, measurements at 300°C

Table 6 - Experiment parameters for our use case

Parameter	Value set
Temperature [°C]	300, 700, 850
Load [N]	40
Type of cycle (move)	Sine
Amplitude [mm]	1
Frequency [Hz]	10
Time [s]	7200
Type of contact [-]	Line

for external kinematics, then measurements at 300, 700 and 850°C for the internal kinematics (arm and bushing) and coefficients of friction for 300, 700 and 850°C. External kinematics in the operation is subject to much lower temperatures (around 300°C) than internal kinematics which is subject up to temperatures of 700 and 850°C. This temperature is the set temperature of the oven. Temperature contribution because of friction is not counted because of its complexity and unavailability to properly measure it. It is therefore necessary to mention this simplification also.

3.1. Calculation of parameters for Goryacheva equation

$$K = \frac{V}{p^\alpha v^\beta t} \tag{30}$$

In our chosen equation (30), some parameters are directly given, such as wear volume V and time t. Others are not directly given but they can be calculated from the experiment parameters, such as pressure p or speed v. Lastly, the „free" parameters, namely K, α, β will be used in prediction models to best predict wear behavior of given material based on material characteristics such as hardness and experimental parameters as given in equation. Regarding the indirect parameters, we must do calculations. Total distance L is calculated using equation (31) where L is the total distance, N is the number of cycles in sine function and L_c is length of one cycle, therefore 4 amplitudes x_m. Then, using the total distance, we calculate speed in equation (32). This is simplification of reality because the speed changes in oscillatory motion. Its maximum in the equilibrium in the middle and minimum (zero) in the amplitude points.

$$L = N \cdot L_c = f \cdot t \cdot 4x_m = 10 \, Hz \cdot 7200 \, s \cdot 4 \cdot 0.001 \, m = 288 \, m \tag{31}$$

$$v = \frac{L}{t} = \frac{288 \, m}{7200 \, s} = 0.04 \, ms^{-1} \tag{32}$$

The second indirect parameter that needs to be calculated is the pressure p as shown in equation (33). The contact in our experiment is "line" in simplification, although the area is slowly getting bigger. We will use simplified value of constant pressure. N is the normal load used in experiment that is being held constant throughout the experiment. y is the distance of the line contact, perpendicular to the oscillatory motion in the x direction of the sliding, namely 10mm.

$$p = \frac{N}{y} = \frac{40\ N}{0{,}01\ m} = 4000\ Nm^{-1} \tag{33}$$

After calculation of the speed and pressure, the parameters used for prediction of wear are specified in Table 7. It consists of material characteristic hardness. Experiment parameters are temperature, pressure, speed and time. Wear volume is the result of the experiment and is measured after the experiment using the laser profilometer and parameters K, α and β are free for predicting the wear behavior using different methods such as statistics, data analysis and machine learning. Possible uses of these methods will be part of this chapter.

Table 7 - Parameters used for prediction

Parameters used for prediction	Values
Temperature [°C]	300, 700, 850
Coeff of friction [-]	Various
Hardness [HV]	Various
Wear volume [m^3]	Various
Pressure [Nm^-1]	4000
Speed [ms^-1]	0.04
Time [s]	7200
K, α, β	Free parameters for prediction

3.2. Exploratory data analysis

To begin the process of data analysis, the data are plotted in scatter plot to give us first insight into wear behavior of External kinematics and Arm and Bushing mechanism, both being plotted together, because the experiment is the same. This scatter plot can be seen in Figure 61. Wear volume is given in original units of measurement on laser profilometer, hence um^3.
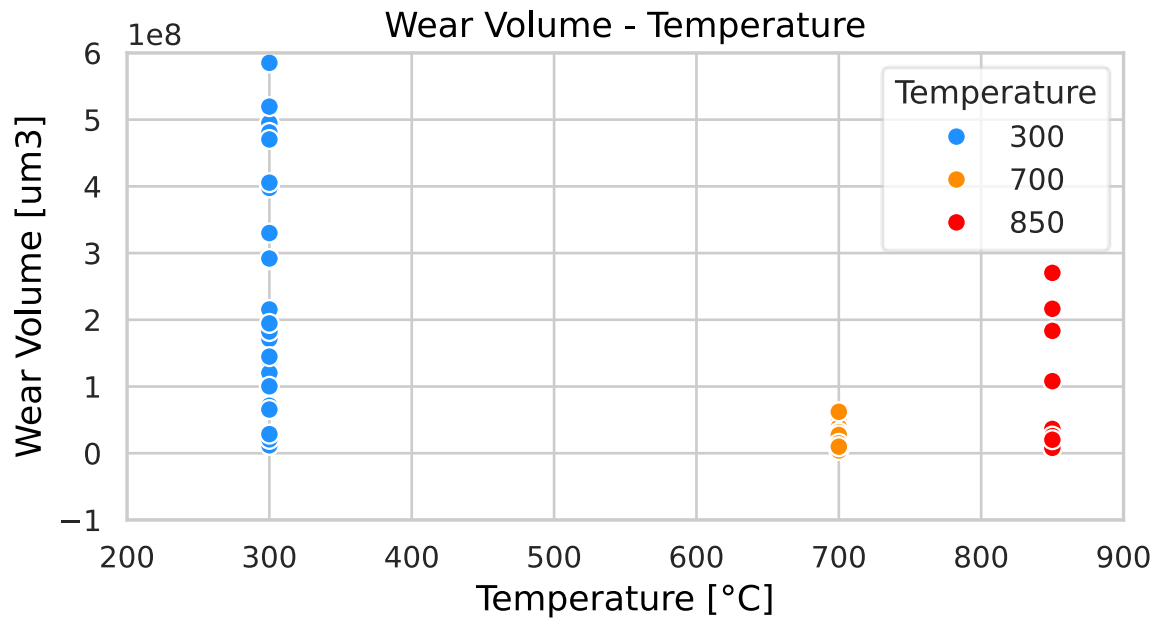
Fig. 61 - Wear volume - Temperature dependency of the measured data

As we can see, the wear volume has the largest dispersion at temperature 300°C, ranging from 0 to 6e8 um3 . Then at 700°C, the dispersion of wear volume values is quite small, only from 0 to 1e8. Then at 850°C, the dispersion of wear volume values gets bigger again, ranging from 0 to 3e8. This can be due to many reasons and no direct tendency and behavior of data is observed yet. Later, the data will be fitted by Goryacheva equation and Machine Learning techniques to predict the behavior and figure out which models are best.

Similar first insight can be done with the measured values of Coefficient of friction as shown in Figure 62. In this case, with the increase in temperature, the CoF values are becoming smaller. When we look at these data, we can assume that linear regression fitting might be a good option, but this dataset is too small to make some predictions upon it.
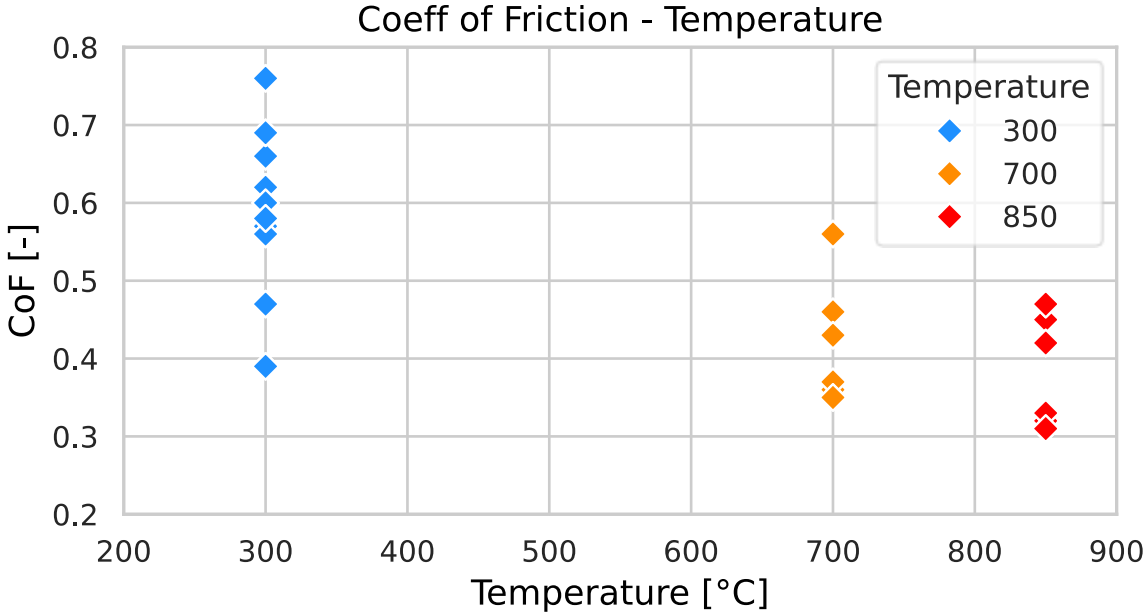
Fig. 62 - Coefficient of Friction - Temperature dependency of the measured data

Using scatter plots, another great insight into wear behavior is produced when we plot Wear Volume – Coefficient of Friction together. This is shown in Figure 63. This scatter plot combines the observations from the first two plots, shown in Figures 61 and 62, that at 300°C, the wear volume and CoF dispersion is the greatest. Then at 700°C, the wear volumes are small and dispersion of CoF is in range 0.35-0.57. Then at 850°C, most of the measured values have the smallest measured wear volumes and coefficients of friction.
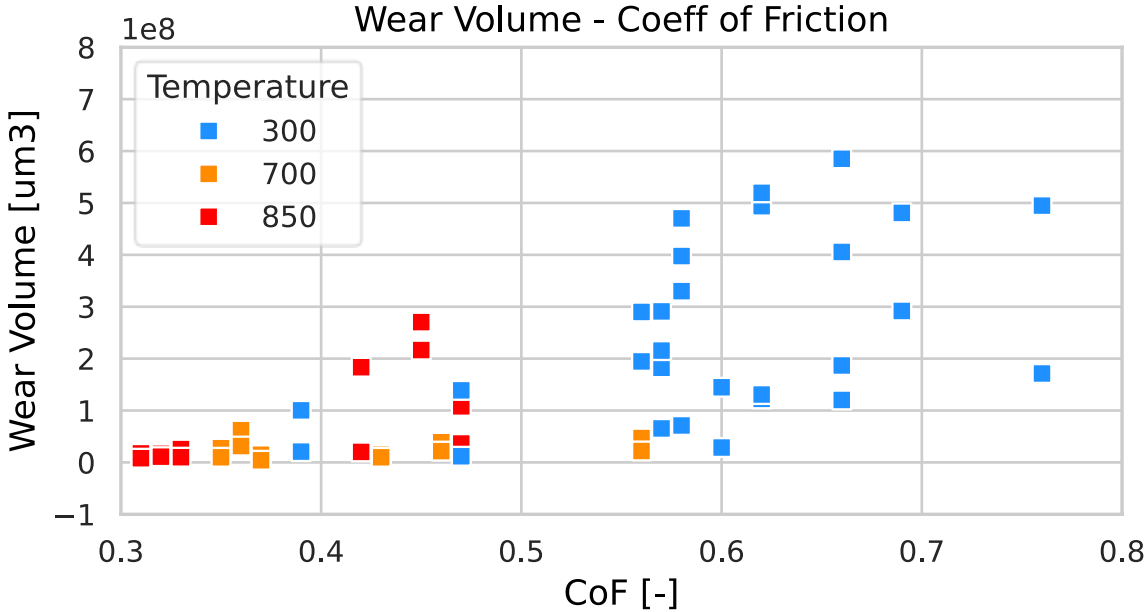
Fig. 63 - Wear volume - Coefficient of friction dependency of the measured data

In total, we have 14 labeled couples as specified in Table 8. Wear Volume values were plotted for each material couple in Figure 64.

Fig. 64 - Wear Volume values sorted by Material Couple

Table 8 - Assignment of material couples

Material Couple	Label
M1/M2	1
M1/M3	2
M1/M4	3
M1/M5	4
M1/M6	5
M1/M7	6
M3/M3	7
M5/M5	8
M8/M7	9
M9/M9	10
M10/M7	11
M10/M10	12
M11/M4	13
M12/M12	14

These scatter plots give us great first insight into the data structures and possible underlying patterns.

3.3. Use of Goryacheva equation for measured data

As shown again in equation (34), the Goryacheva equation is empirical prediction of wear volume based on experimental parameters and wear coefficients K, α, β.

$$K = \frac{V}{p^\alpha v^\beta t} \tag{34}$$

During the measurements in Garrett Motion, the pressure, speed and time of the experiment were constant in order to create a comparative method for wear volume.

$$K = \frac{V}{4000^\alpha 0.04^\beta 7200} \tag{35}$$

We therefore have 1 equation with 3 unknown variables, as shown in equation (35). Also, we can see, that with the α and β being an arbitrary number, they still lead to constant number as constant to the power of constant is still constant. K and V will in this case therefore have a linear relationship of some sort.

We unfortunately doesn't know the values of α and β for some prediction of K and in case with pressure and speed being constant, it doesn't matter. Let's initially assume that α = 1 and β = 1 for this scenario and calculate K from mean value of wear volume. If α = 1 and β = 1, then the relationship between K and V for our given experiment parameters is shown in equation (36).

$$K = \frac{V}{1\,152\,000} = 8.68056 \cdot 10^{-7} \cdot V \tag{36}$$

When we put in mean value of wear volume, which is 1.5738e-10 m^3, we obtain K = 1.3662e-16 [-]. This value of K is quite small regarding the general range of 1e-3 to 1e-7 mentioned in literature. With those parameters of ours, we can assume that coefficients α and β therefore need to be smaller than 1 to obtain some expected values of K. This is done via 3D plot of alfa-beta-K relationship. Alfa and Beta shown in Figure 65.
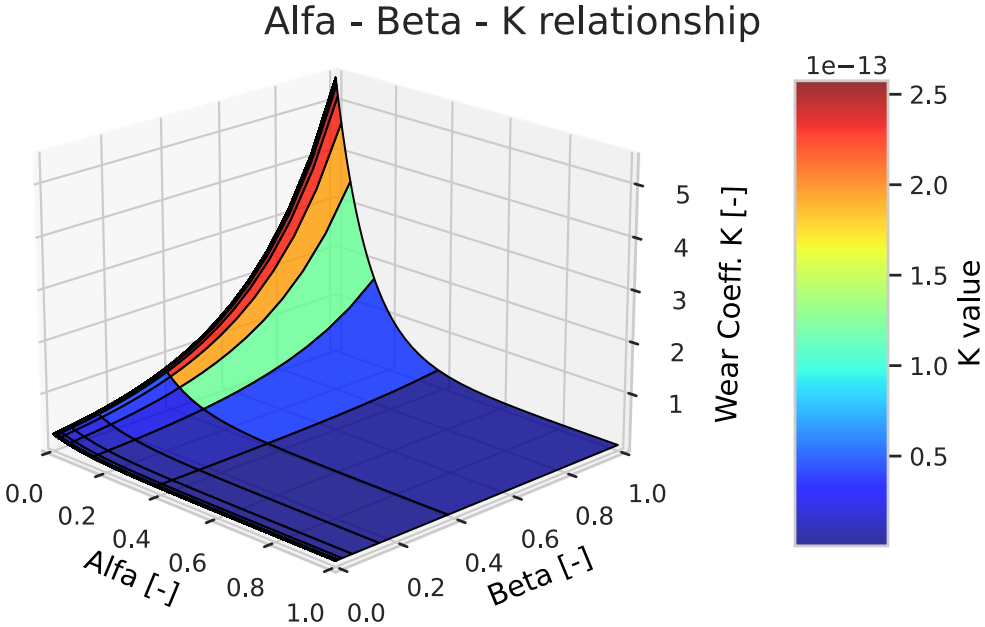

Fig. 65 - Alfa - Beta and K coefficients relationship in Goryacheva equation for our use case

Based on this plot, author decided to verify if K converges based on lowering α and β values or not. This convergence is shown in Figure 66, where is calculated mean K based on mean Wear volume of our samples and its changes with changing α and β coefficients separately.
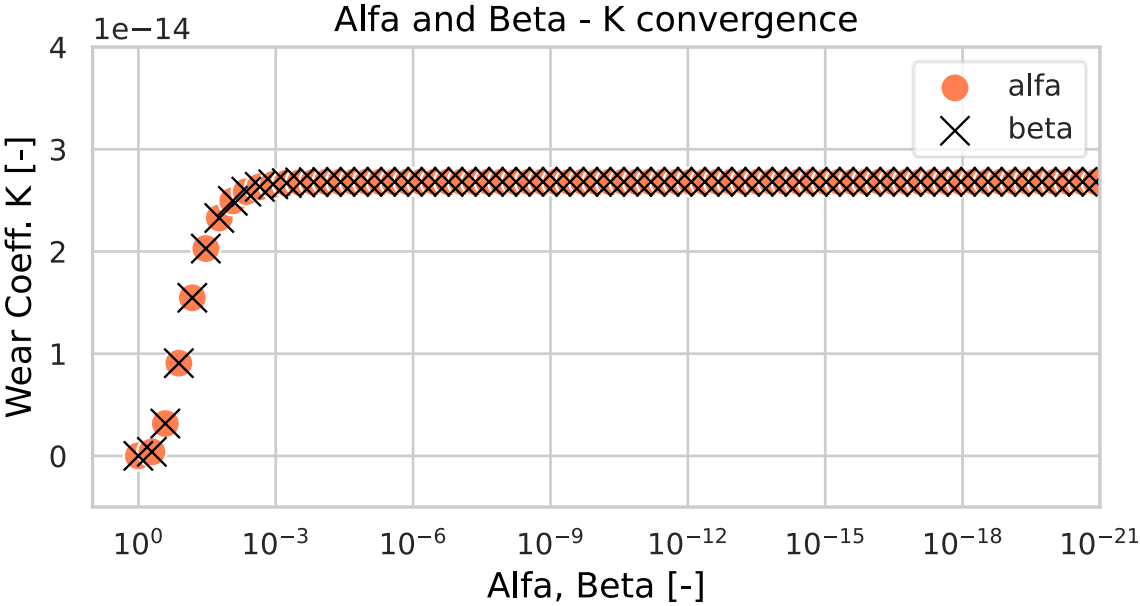
Fig. 66 - Verifying K convergence based on Alfa and Beta values

We can see in Figure 66, that from 10^{-3} the K values converge to the value of 2.6818e-14 [-]. Based on this verification, we plot all our wear volume values with parameters α = 1 and β = 1and α = 1e-20 and β = 1e-20 to compare them. This is shown in Figure 67, where Y axis is in log-scale for better understanding.
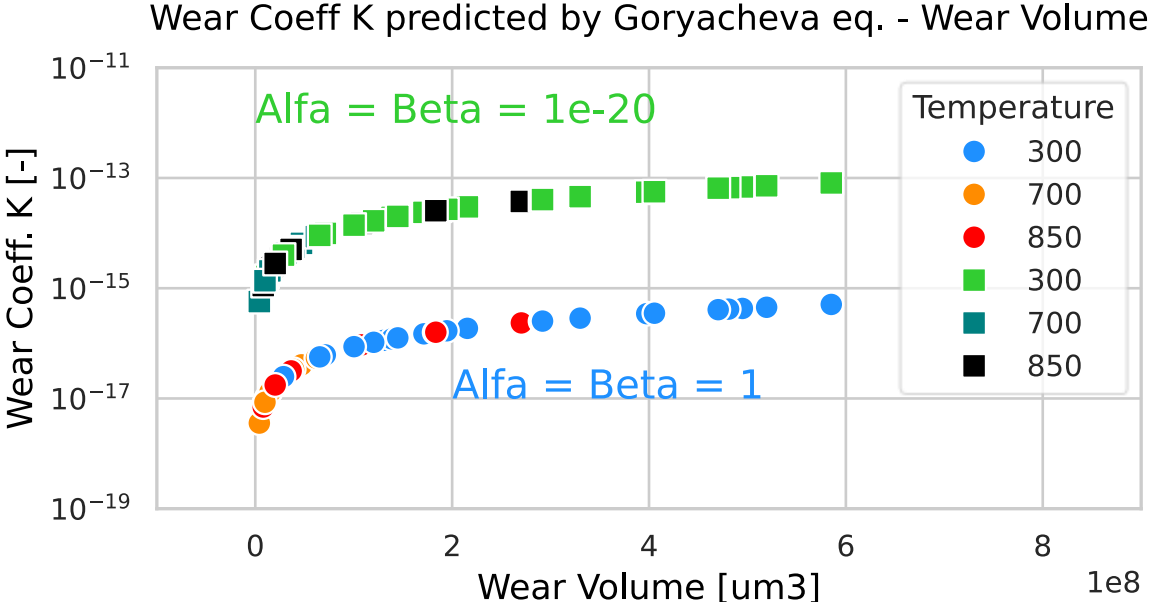
Fig. 67 - Comparison of wear coefficient K values with alfa, beta = 1 and alfa, beta = 1e-20

We can see that K values are higher with lower α and β values, but even in the lowest scenario of convergence, which we confirmed earlier, K values are still in order of 10^{-13} which is still 6 orders lower than expected lowest values of coefficient K mentioned in the literature.

Based on this, the author proposes that for parameters used in this use case and measurement methodology, the Goryacheva equation is far from useful and other methods will be deployed. To conclude more research in the topic of Goryacheva equation, the author suggests that more measurements with varying experimental parameters are conducted, instead of measuring methodology on comparative basis.

4. Methodology – Machine Learning
4.1. Data pre-processing, Feature engineering

With machine learning (ML), one of the most important parts of the project is to have good and representative data of our problem to make good predictions. The data obtained from Garrett Motion company were measured on DN55 tribometer as specified earlier and methodology was built comparatively – apply same experiment parameters on different material couples to see which material couple has lowest wear volume, therefore is best for the kinematics mechanisms outside the turbocharger. This is unfortunate for the ML approaches, hence many of the parameters are constants and therefore offer no predictive ability for ML approaches. Nevertheless, we will proceed with available data that are not constant for all the experiments and find the best machine learning approaches for this use case.

Machine Learning dataset
Feature Matrix X — Target Vector y

couple [-]	spec [-]	temp [°C]	hard_ratio [-]	cof [-]	volume [um3]
1	plate	300	0,8997	0,58	397673161
1	pin	300	0,8997	0,58	71155353
1	plate	700	0,8997	0,56	46930815
1	pin	700	0,8997	0,56	22005355
1	plate	850	0,8997	0,32	16521947
1	pin	850	0,8997	0,32	11040056
2	plate	300	0,3915	0,62	493756419
2	pin	300	0,3915	0,62	122426157

Fig. 68 - Dataset for our project - feature matrix X and target vector y

The data has been put into **feature matrix X** and corresponding target vector y as shown in Figure 68. This matrix consists of feature vectors with following parameters:

- **Couple** – In chapter 3.2 Exploratory data analysis explained, as each of the 14 different material couples was assigned number from 1 to 14
- **Specimen** – Either pin or plate, which is later for machine learning methods converted that pin = 0 and plate = 1.
- **Temperature** – Temperature at which the experiment was conducted: 300, 700 or 850 °C

- **Hardness Ratio** – Hardness of the surface is an important parameter for prediction of wear behavior and since our prediction will be done for material couples, each material couple has characteristic ratio of hardnesses between its surfaces. The obtained hardness data were measured using HV1 – Vickers scale. Hardness ratio is calculated as HV_specimen1 / HV_specimen2 Hardness ratios for material couples are shown in Figure 69.
- **CoF** – Average coefficient of friction calculated from the values measured in time during experiment are presented.

Material couple		Hardness ratio [HV1/HV1]
M1/M2	1	0,8997
M1/M3	2	0,3915
M1/M4	3	0,7596
M1/M5	4	0,3861
M1/M6	5	1,1488
M1/M7	6	0,5101
M3/M3	7	1,0000
M5/M5	8	0,9972
M8/M7	9	1,8404
M9/M9	10	1,0000
M10/M7	11	1,4312
M10/M10	12	1,0000
M11/M4	13	2,7404
M12/M12	14	1,0000

Fig. 69 - Material couples with corresponding hardness ratio's

Each feature vector x in feature matrix X corresponds to target (value) in target vector y. In our case, this target value is wear volume, measured on the laser profilometer. All of the wear volume values therefore form a target vector y which will be the predicted outcome of model. To summarize, feature vector having certain value or option in variables: couple, specimen, temperature, hardness ratio and CoF corresponds to certain value of wear volume as target. Dataset shown is sorted by default, but for the training of the ML algorithms, the dataset is shuffled to remove inherent ordering or any bias. We will use this set of features to train models and then predict the target values and compare the metrics to see which of the models performs best for our use case.

4.2. Model selection, design and fine-tuning

The process of finding the best ML model begins with the estimation of the few best performing models and then fine-tuning them to get the best possible results. The following ML models were used for initial performance estimation: Linear Regression, Polynomial Regression, Linear SVR, SVR, Multilayer Perceptron NN (Sequential), Decision Trees, Random Forest, Extreme Gradient Boosting (XGB), Gaussian Process Regression (GPR) and Genetic Algorithm (GA). All these modules were utilized using Python libraries, namely Scikit-learn, TensorFlow, Keras, GPlearn and XGboost. Criteria for selection of best models were R2 and MSE. If you are interested in more information about these models, see the attached file rychly_ml_wide.ipynb, only models that were chosen for closer examination will be discussed more in depth.

4 best performing models were chosen for closer examination. Namely, Support Vector Regression (SVR), Random Forest (RF), Genetic Algorithm using Symbolic Regressor (GA) and Neural Networks using Sequential Multilayer Perceptron (NN). All these models are explained in chapter 1.5 in the machine learning section so this chapter can be dedicated to finetuning the design of those models.

4.2.1. Support Vector Regression

Using Hyperopt Python library, the hyperparameter space for SVR model was searched and Table 9 shows the best hyperparameter values. The Polynomial features degree is 1, therefore original values. More about parameters can be found in sklearn.svm.SVR documentation or in chapter 1.5.

Table 9 - SVR fine-tuned architecture

Hyperparameters for SVR	Value
C - Regularization param. [-]	1.3
Epsilon – width of street [-]	0.1
Kernel [-]	rbf
Gamma – kernel coeff [-]	auto
Polynomial Features degree [-]	1

4.2.2. Random Forest

The Hyperparameter space of RF model was searched via Hyperopt and the results for best performing model are in Table 10. Model best performed for Polynomial features degree of 3. More about parameters can be found in sklearn.ensemble.RandomForestRegressor documentation or in chapter 1.5.

Table 10 - RF fine-tuned architecture

Hyperparameters for RF	Value
N estimators [-]	50
Max depth [-]	10
Min samples split [-]	2
Min samples leaf [-]	1
Min weight fraction leaf [-]	0.0186
Max leaf nodes [-]	None
Min impurity decrease [-]	9.47e-6
Polynomial Features degree [-]	3

4.2.3. Genetic Algorithm using Symbolic Regressor

Genetic algorithm was programmed using both recommended values and values proposed by Faris [49] and optimal values found using Hyperopt grid search. These parameters are presented in Table 11.

Table 11 - GA fine-tuned architecture

Hyperparameters for GA	Value	Hyperparameters for GA	Value
Function set [-]	+,-,/,·	P crossover [-]	0.515
Population size [-]	200	P subtree mutation [-]	0.178
Generations [-]	150	P hoist mutation [-]	0.110
Tournament size [-]	30	P point mutation [-]	0.099
Stopping criteria [-]	0.001	P point replace [-]	0.168
Max samples [-]	0.5	Polynomial Features degree [-]	1
Parsimony coefficient [-]	0.001	-	-

The best performing Genetic algorithm used the equation shown in equation (37). Where X_1 is the number of material couple, X_2 is specimen type, either pin or plate, X_3 is temperature and X_5 is coefficient of friction. Unused X_4 is hardness ratio.

$$GA\ equation = ((X_2 + X_5) - 0.266) \cdot \left(\frac{X_5}{3X_3 + 2X_5 + X_1}\right) \qquad (37)$$

4.2.4. Sequential Multilayer Perceptron NN

Sequential Multilayer Perceptron NN was created and its parameters also with Polynomial Features degree were optimized using Hyperopt. The best performing architecture is shown in Table 12. The number of hidden layers is 1 with 3 hidden units, which corresponds to the number of input features. Batch size is 1 and the activation function for network is Leaky RELU – Rectified Linear Unit. The Polynomial Features degree was 1, therefore original features performed the best.

Table 12 - NN fine-tuned architecture

Hyperparameters for NN	Value
Number of hidden layers [-]	1
Number of hidden units [-]	3
Batch size [-]	1
Activation function [-]	Leaky RELU
Learning rate [-]	0.00018022
Polynomial Features deg. [-]	1

4.3. Evaluation, Comparison, Feature importance

The metrics of performance for the 4 best performing models are shown in Table 13. R2 train mean was calculated using KFold cross validation with K=5. For NN, the validation split was 0.2 of the training set. R2 train mean is therefore mean performance on the training set, indicating how model generalizes on different training sets. R2 test also means the generalization error is the R2 calculated from performing on the test set i.e., unseen data. MSE and RMSE values are calculated for final model parameters used for prediction on test set, not cross validation.

Table 13 - Comparison of 4 best performing ML models

Evaluation metrics	SVR	RF	GA	NN
R2 train mean [-]	0.351	0.170	0.212	0.194
R2 test [-]	0.429	0.621	0.37	0.477
MSE [um3^2]	1.856e16	1.232e16	2.047e16	1.939e16
RMSE [um3]	1.362e8	1.109e8	1.430e8	1.393e8

From the total of 4 fine-tuned models, the Support Vector Regression model is chosen as the best model. Random Forrest and Genetic Algorithm models are weak regarding R2 metrics and regarding the score difference between R2 mean train and R2 test which signals overfitting. NN model is also weak regarding the metrics but performs slightly better than the RF and GA models. NN performance is sometimes like SVR model, but the calculation of the model is in wide range of values for both R2 train and R2 test. R2 train and R2 test were both calculated training multiple NNs and taking the average values of both. This could be because of the small quantity and quality of the dataset or little variation in the dataset.

SVR model is the best performing model among these 4 fine-tuned models that were chosen from wide pool of possible models, this pool of models can be seen in attachment rychly_ml_wide.ipynb. Choice of the best model was based on the values of R2 = 0.351 for training set and R2 = 0.429 for test set, which was among the best of trained models for our use case, but still indicates notions of overfitting and the performance can be much better and R2 closer to 1. This overfitting is, as with NN and other models, most likely due to the small quantity and quality of training data and little variation in the measured data. MSE for SVR model is $1.856 \cdot 10^{16}$ $(um^3)^2$ and RMSE is $1.362 \cdot 10^8$ um^3. So, the average root mean squared error of the model's prediction of the wear volume values is $1.362 \cdot 10^8$ um^3.

Figure 70 shows the performance of the 4 fine-tuned models visually. True vs Predicted values of Wear volume are shown. The ideal shape of the curve is linear i.e. each predicted value corresponds to true value.
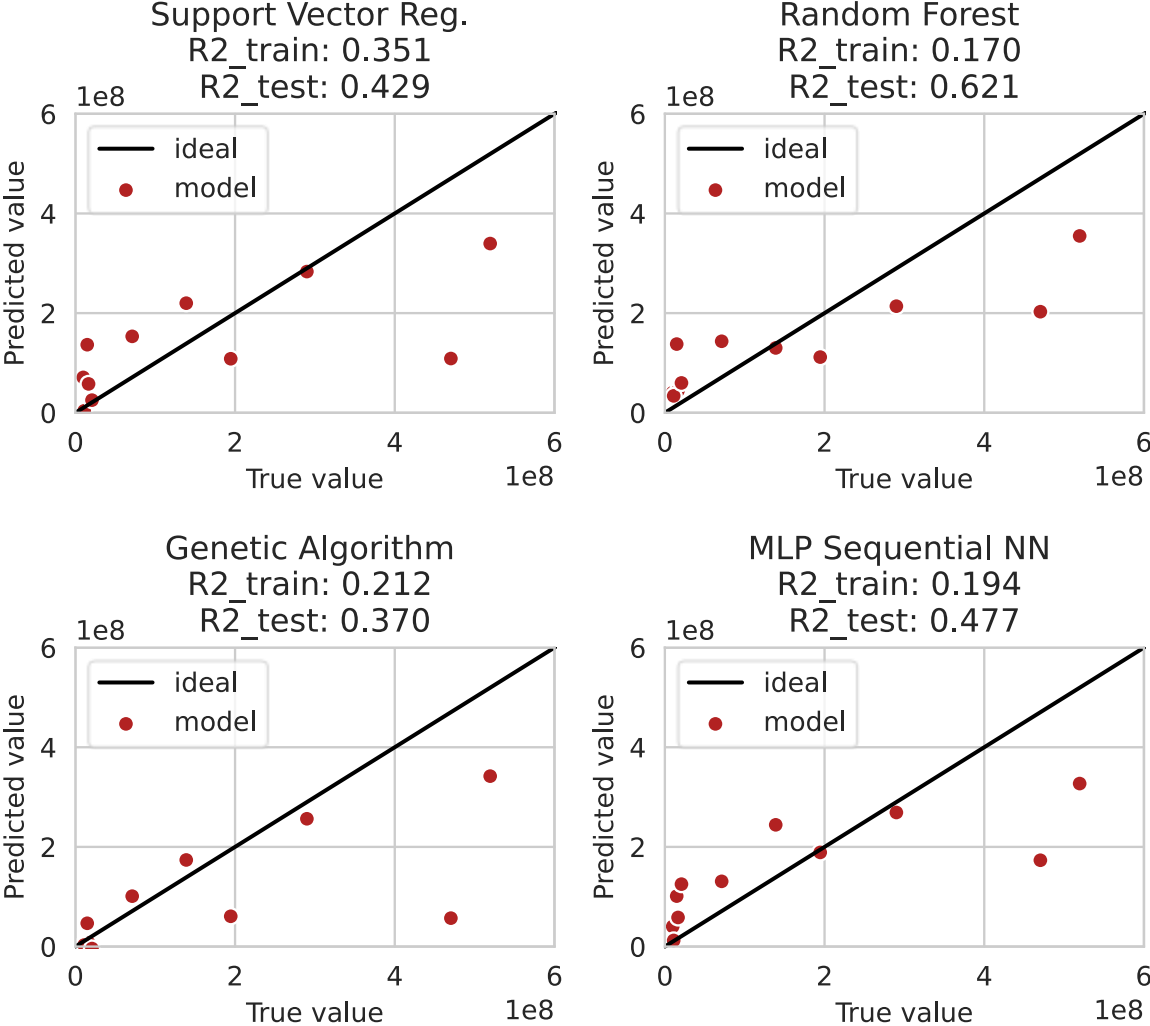
Fig. 70 - Comparison of 4 best performing models for our use case - SVR, RF, GA and MLP NN

For the best chosen model, SVR, feature importance analysis is performed. Feature importances are calculated based on Support Vectors (boundaries in which the data is) and Dual Coefficients, which represent the weight of each support vector in the decision function. Importance of certain feature is then inferred by examining how much each support vector contributes to the prediction, weighted by the vector's corresponding dual coefficients. This is calculated as a dot product of Support vectors and Support vectors coefficients. The Polynomial Features degree for the best performing model is 1, therefore original set of features from feature matrix X. Visually shown in Figure 71. The most important feature in the SVR model is coefficient of friction, then specimen, either pin or plate, next is temperature and lowest importances hold the couple number and hardness ratio. The author suggests measuring more data with more variety of experimental parameters which would broaden the field of potential features. Feature engineering with varied values of time, load, temperature and many others would, in the author's humble opinion, present more prediction value and better model.
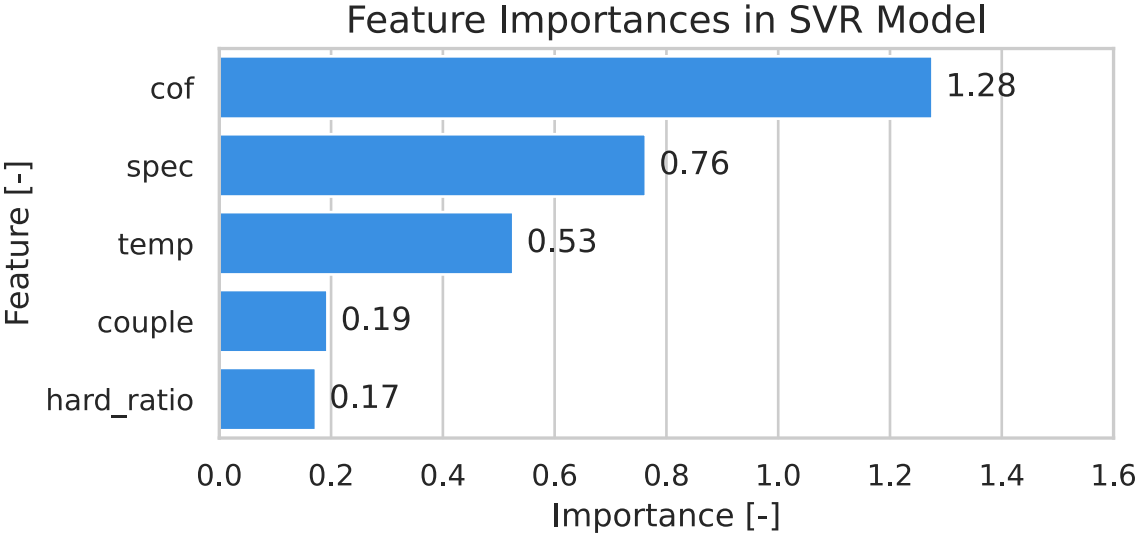
Fig. 71 - Feature importances in SVR model for our use case

Using a small dataset for predictions makes the model heavily influenced by the test-train split. This influence was examined by comparing the best performing SVR model architecture with different data split values. The split was characterized by the "ratio" parameter determining where is the boundary between train and test. Therefore ratio = 0.75 means 75/25 train-test split. Ratios varied from 0.6 to 0.95 with step 0.05. The best performing models were with 60/40 split and 80/20 split. The metrics of those 2 models are shown in Table 14. Model performance is quite similar – 60/40 split is slightly better regarding the overfitting and 80/20 gives smaller, better, RMSE and generalizes better to testing data. The author suggests trying both data splits for optimal performance of your own model and proposes 80/20 split as the best result due to slightly better R2 scores. 60/40 and 80/20 are common splits in literature and this experimental finding confirms their usefulness.

Table 14 - Comparison of data splits for SVR model

Split	MSE	RMSE	Mean train score	Test score
60/40	2.0416e16	1.4288e8	0.312	0.374
80/20	1.8558e16	1.3622e8	0.351	0.429

As a conclusion of this methodology chapter, SVR model with following architecture was chosen as the best performing model of all models for this use case with the recommendation to gather more data with greater variety of experimental parameters to maximize its potential. Architecture is shown in Table 15. It is however necessary to again bring up the fact that wear is a very complex phenomena and its prediction is influenced by many physical, chemical, and mechanical phenomena and predictions might not be straightforward. This model gives the reader and authors of following researches good starting point for their own predictions using machine learning in modelling of wear.

Table 15 - Best architecture of SVR for our use case

Best model for our use case - SVR	Value
C - Regularization param. [-]	1.3
Epsilon – width of street [-]	0.1
Kernel [-]	rbf
Gamma – kernel coeff [-]	auto
Polynomial Features degree [-]	1
Data split [-]	80/20

5. Conclusion, Further work

The main mission of this thesis was to predict wear behavior on concrete use case and save resources necessary for the experiments using both empirical equations from literature and machine learning approaches. To grasp this topic, throughout research of the relevant tribology topics was conducted. Chapter 1 is dedicated to introduction to tribology, friction and wear and different wear mechanisms. Later, the adhesive wear is chosen as the main mechanism involved in our use case and literature search is primarily performed to find empirical relations for predicting this type of wear. Throughout literature search for empirical predictions of adhesive is presented in chapter 1.4 and the Goryacheva Equation is chosen as the best equation for conducting methodology. Shown in equation (38).

$$V = Kp^{\alpha}v^{\beta}t \tag{38}$$

Empirical predictions of wear using the Goryacheva equation are complemented by the machine learning approaches based on experimental data obtained from Garrett Motion company. Various models were tried and sub-chapter 1.5 explains 4 models that were fine-tuned for the best possible model. These models are Support Vector Regression (SVR), Random Forest (RF), Genetic Algorithm using Symbolic Regressor (GA) and Multilayer Perceptron Sequential Neural Network (MLP NN). Chapter 1.5 also contains the research conducted in the field of tribology using the machine learning approaches of various authors. Their findings were used as a starting point and help design models in this thesis. Chapter 2 introduces the problematic and its testing method, the instruments of measurement and ways of obtaining the data later used for empirical predictions and models.

The methodology chapter 3 focuses on exploratory data analysis and predictions using the Goryacheva equations. Equation is calculated using the experimental parameters and the variation of K, α and β wear coefficients is conducted. 3D plot and α, β convergence analysis is presented. Result is that Goryacheva equation even with various values of α and β is weak tool in predicting the wear behavior.

In chapter 4 methodology of machine learning, the whole process of creating ML model is presented. Feature Matrix contains couple number, specimen, temperature, hardness ratio and coefficient of friction therefore making it 5 features for predicting one target vector variable – wear volume. Various models are employed and best models are selected for fine tuning: Support Vector Regression (SVR), Random Forest (RF), Genetic Algorithm using Symbolic Regressor (GA) and Multilayer Perceptron Sequential Neural Network (MLP NN). Best performing architectures are presented for each model, the models are compared using metrics R2 train, R2 test, MSE and RMSE. SVR – Support Vector Regression model is chosen as the best for given use case and both architecture and resulting metrics are shown in Table 16.

Table 16 - Best performing model SVR and its evaluation metrics

Architecture		Evaluation	
Best model for our use case - SVR	**Value**	**Metric**	**Value**
C - Regularization param. [-]	1.3	R2 train mean [-]	0.351
Epsilon – width of street [-]	0.1	R2 test [-]	0.429
Kernel [-]	rbf	MSE [um3^2]	1.856e16
Gamma – kernel coeff [-]	auto	RMSE [um3]	1.362e8
Polynomial Features degree [-]	1		
Data split [-]	80/20		

For the best SVR model, feature importance analysis is performed with CoF being the most important feature for prediction. Feature importance analysis is shown in Figure 72.
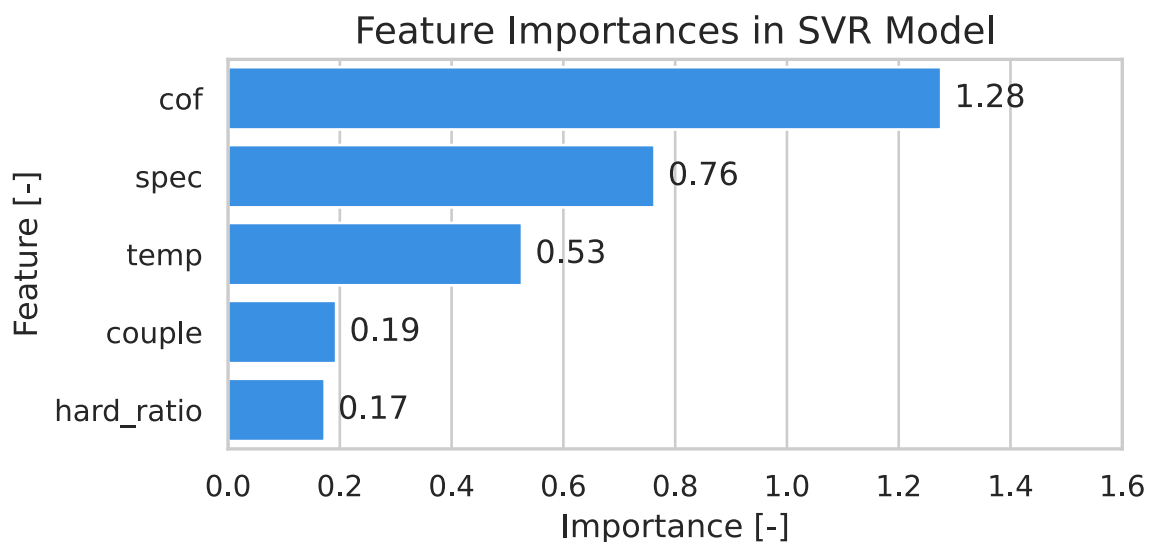
Fig. 72 - Feature importances in best performing SVR model for our use case

To maximize the research potential of this thesis, analysis of optimal data split value is performed because in small datasets, the data split heavily influences the final model. 80/20 train/test split is chosen as the best with 60/40 close behind.

To produce better results based on the research in this thesis, the author suggests gathering more training data with various experimental parameters set. This will greatly improve machine learning models via feature engineering (adding more features) and allow the models to learn on bigger samples. Data is the key to machine learning and the author suggests focusing on this part of the model creation – obtaining of the data and feature engineering. Another possible approach to try is to add material properties as single features, such as concentrations of elements in % or phases as performed in Rajput [61]. Another approach to try is to compare wear coefficient K values from this thesis with literature findings of certain materials and optimize Goryacheva equation or delve deeper into the relationships of variables inside the wear coefficient K itself.

6. Literature

[1]   GORYACHEVA, I.G. 2013. Contact Mechanics in Tribology [online]. 1. Dordrecht: Springer. Available at: https://link.springer.com/book/10.1007/978-94-015-9048-8#book-header

[2]   HUTCHINGS, Ian and Philip SHIPWAY. 2017. Tribology [online]. 2. Elsevier. Available at: https://www.sciencedirect.com/book/9780081009109/tribology

[3]   RABINOWICZ, Ernest and David TABOR. 1951. Metallic transfer between sliding metals: an autoradiographic study. The Royal Society Publishing [online]. The Royal Society Publishing, (Volume 2081095). Available at: https://doi.org/10.1098/rspa.1951.0174

[4]   COLLINS, Danielle. 2019. Motion Control Tips [online]. Available at: https://www.motioncontroltips.com/why-is-static-friction-greater-than-kinetic-friction/

[5]   BURWELL, J.T. and C.D. STRANG. 1952. Metallic wear. Proc. R. Soc. [PDF]. London, (A212), 470-477. Available at: https://royalsocietypublishing.org/doi/10.1098/rspa.1952.0239

[6]   NEALE, M.J. 1995. The tribology handbook [online]. 2. Butterworth-Heinemann. Available at: https://doi.org/10.1016/B978-075061198-5/50001-1

[7]   BOWDEN, Frank Philip and David TABOR. 2001. The friction and lubrication of solids. Oxford university press.

[8]   GOEDECKE, Andreas. 2014. Chapter 1 - Introduction. Transient Effects in Friction - Fractal Asperity Creep [online]. Wien: Springer, p. 1-3. Available at: https://link.springer.com/book/10.1007/978-3-7091-1506-0

[9]   Archard equation. 2023. Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 17.01.2023. Available at: https://en.wikipedia.org/w/index.php?title=Archard_equation&oldid=1134127939

[10]   REYE, Karl Theodor. 1860. "Zur Theorie der Zapfenreibung" [On the theory of pivot friction]. Der Civilingenieur - Zeitschrift für das Ingenieurwesen [online]. München, (6), 235-255. Available at: https://www.digitale-sammlungen.de/de/view/bsb10048713?page=132 (The German version was scanned via OCR and translated to English using ChatGPT. Possible inaccuracies may exist due to this translation process.)

[11]   HOLM, Ragnar. 1967 (2013). Electric Contacts Theory and Application [online]. 4. Berlin, Heidelberg: Springer. Available at: https://doi.org/10.1007/978-3-662-06688-1

[12]   BURWELL, J. T. and C. D. STRANG. 1952. On the Empirical Law of Adhesive Wear. Journal of Applied Physics [online]. 23(1), 18-28. Available at: https://pubs.aip.org/jap/article/23/1/18/160173/On-the-Empirical-Law-of-Adhesive-Wear

[13]   ARCHARD, J.F. 1953. Contact and Rubbing of Flat Surfaces. J. Appl.Phys [online]. 24(8), 981-988. Available at: https://doi.org/10.1063/1.1721448

[14]   WEN, Shizhu and Ping HUANG. 2017. Principles of Tribology [online]. 2. Tsinghua: Tsinghua University Press. Available at: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119214908

[15]   RABINOWICZ, E. 1981. The Wear Coefficient—Magnitude, Scatter, Uses. Journal of Lubrication Technology [online]. 103(2), 188-193. Available at: https://asmedigitalcollection.asme.org/tribology/article/103/2/188/423574/The-Wear-CoefficientMagnitude-Scatter-Uses

[16]   RABINOWICZ, Ernest. 1958. The effect of size on the looseness of wear fragments. Wear [online]. 2(1), 4-8. Available at: https://linkinghub.elsevier.com/retrieve/pii/0043164858903351

[17]   AGHABABAEI, Ramin, Derek H. WARNER and Jean-Francois MOLINARI. 2016. Critical length scale controls adhesive wear mechanisms. Nature Communications [online]. 7(1). Available at: https://www.nature.com/articles/ncomms11816

[18]   ARCHARD, J. F. 1961. Single Contacts and Multiple Encounters. Journal of Applied Physics [online]. 32(8), 1420-1425. Available at: https://pubs.aip.org/jap/article/32/8/1420/163208/Single-Contacts-and-Multiple-Encounters

[19]   GEE, M. G.; NEALE, M. J. General approach and procedures for unlubricated sliding wear tests. 2002.

[20]   POPOV, VL and R POHRT. 2018. Adhesive wear and particle emission: Numerical approach based on asperity-free formulation of Rabinowicz criterion. Friction [online]. 2018(6(3), 260-273. Available at: https://doi.org/10.1007/s40544-018-0236-4

[21]   HOLM, Ragnar. 2013. Elektrische Kontakte / Electric Contacts Handbook [online]. 3. Berlin, Heidelberg: Springer. Available at: https://link.springer.com/book/10.1007/978-3-662-25893-4

[22]   NEALE, M.J. and M GEE. 2001. A Guide to Wear Problems and Testing for Industry [online].William Andrew.Available at: https://www.sciencedirect.com/book/9780815514718/a-guide-to-wear-problems-and-testing-for-industry#book-info

[23]   BHUSHAN, Bharat. 2013. Principles and applications of tribology [online]. 2. Hoboken, USA: Wiley. Available at: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118403020

[24]   BURWELL, John T. 1957. Survey of possible wear mechanisms. Wear [online]. 1957(Volume 1, Issue 2), 119-141. Available at: https://www.sciencedirect.com/science/article/pii/0043164857900054

[25]   TERWEY, J. Torben, Mohamed Ali FOURATI, Florian PAPE and Gerhard POLL. 2020. Energy-Based Modelling of Adhesive Wear in the Mixed Lubrication Regime. Lubricants [online]. 8(2), 1. Available at: https://www.mdpi.com/2075-4442/8/2/16

[26]   RABINOWICZ, Ernest. 1984. The least wear. Wear [online]. 100(1-3), 533-541. Available at: https://linkinghub.elsevier.com/retrieve/pii/0043164884900310

[27]   GÅÅRD, A, N HALLBÄCK, P KRAKHMALEV and J BERGSTRÖM. 2010. Temperature effects on adhesive wear in dry sliding contacts. Wear [online]. (268, 7-8), 968-975. Available at: https://www.sciencedirect.com/science/article/pii/S0043164809006292

[28]   RHEE, S.K. 1970. Wear equation for polymers sliding against metal surfaces. Wear [online]. 16(6), 431-445. Available at: https://linkinghub.elsevier.com/retrieve/pii/0043164870901705

[29]   YANG, L. J. 2004. Prediction of Steady-State Wear Coefficients in Adhesive Wear. Tribology Transactions [online]. 47(3), 335-340. Available at: http://www.tandfonline.com/doi/abs/10.1080/05698190490455366

[30]   BHATTACHARYYA, H. P. 1956. On the Law of Adhesive Wear. Journal of Applied Physics [online]. 27(6), 661-662. Available at: https://pubs.aip.org/jap/article/27/6/661/161632/On-the-Law-of-Adhesive-Wear

[31]   BURWELL, J. T. 1956. Law of Adhesive Wear. Journal of Applied Physics [online]. 27(12), 1561-1561. Available at: https://pubs.aip.org/jap/article/27/12/1561/161484/Law-of-Adhesive-Wear

[32]   VARENBERG, Michael. 2022. Adjusting for Running-in: Extension of the Archard Wear Equation. Tribology Letters [online]. 70(2). Available at: https://link.springer.com/10.1007/s11249-022-01602-6

[33]   MISHINA, Hiroshi and Alan HASE. 2013. Wear equation for adhesive wear established through elementary process of wear. Wear [online]. 308(1-2), 186-192. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0043164813004092

[34]   QUEENER, C.A., T.C. SMITH and W.L. MITCHELL. 1965. Transient wear of machine parts. Wear [online]. 8(5), 391-400. Available at: https://linkinghub.elsevier.com/retrieve/pii/0043164865901705

[35]   GÉRON, Aurélien. 2019. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems [online]. Second edition. Beijing: O'Reilly. Available at: https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/

[36]   MIRJALILI, Sayedali, Faris HOSSAM and Ibrahim ALJARAH. 2020. Evolutionary Machine Learning Techniques - Algorithms and Applications [online]. Singapore: Springer. Available at: https://doi.org/10.1007/978-981-32-9990-0_1

[37]   BANKO, Michele and Eric BRILL. 2001. Scaling to very very large corpora for natural language disambiguation. Proceedings of the 39th Annual Meeting on Association for Computational Linguistics - ACL '01 [online]. Morristown, NJ, USA: Association for Computational Linguistics, , 26-33. Available at: http://portal.acm.org/citation.cfm?doid=1073012.1073017

[38]   HALEVY, Alon, Peter NORVIG and Fernando PEREIRA. 2009. The Unreasonable Effectiveness of Data. IEEE Intelligent Systems [online]. 24(2), 8-12. Available at: http://ieeexplore.ieee.org/document/4804817/

[39]   BROWNLEE, Jason. 2020. How to Calculate Feature Importance With Python. BROWNLEE, Jason. Machine Learning Mastery [online]. Available at: https://machinelearningmastery.com/calculate-feature-importance-with-python/

[40]   GITHUB. 2024. Train_Test_Split [online]. Available at: https://github.com/mGalarnyk/Python_Tutorials/blob/master/Sklearn/Train_Test_Split/TrainTestSplitScikitLearn.ipynb

[41]   GITHUB. 2018. Root Mean Squared Error Versus Mean Absolute Error [online]. Available at: https://jmlb.github.io/flashcards/2018/07/01/mae_vs_rmse/

[42]   Mean absolute error. 2001-. Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation. Available at: https://en.wikipedia.org/wiki/Mean_absolute_error

[43]   Mean squared error. 2001-. Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation. Available at: https://en.wikipedia.org/wiki/Mean_squared_error

[44]   Root mean square deviation. 2001-. Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation. Available at: https://en.wikipedia.org/wiki/Root_mean_square_deviation

[45]   Coefficient of determination. 2001-. Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation. Available at: https://en.wikipedia.org/wiki/Coefficient_of_determination

[46]   EVS - Explained Variance Score. 2021. THIEU. Permetrics Python library [online]. Available at: https://permetrics.readthedocs.io/en/latest/pages/regression/EVS.html

[47]   SCIKIT-LEARN. 2024. SVR. SCIKIT-LEARN. Scikit-learn [online]. Available at: https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html

[48]   SCIKIT-LEARN. 2024. RandomForestRegressor. SCIKIT-LEARN. Scikit-learn [online]. Available at: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html

[49]   FARIS, Rana, Bara'a ALMASRI, Hossam FARIS, Faris M. AL-OQLA and Doraid DALALAH. 2020. Evolving Genetic Programming Models for Predicting Quantities of Adhesive Wear in Low and Medium Carbon Steel. Evolutionary Machine Learning Techniques [online]. Singapore: Springer Singapore, , 113-127. Algorithms for Intelligent Systems. Available at: http://link.springer.com/10.1007/978-981-32-9990-0_7

[50]   MATHWORKS. 2024. How the Genetic Algorithm Works. MATHWORKS. MathWorks [online]. Available at: https://www.mathworks.com/help/gads/how-the-genetic-algorithm-works.html

[51]   GPLEARN. Symbolic Regressor. GPLEARN. GPLearn documentation [online]. Available at: https://gplearn.readthedocs.io/en/stable/reference.html

[52]   CHOLLET, François and . 12 April 2020AD. The Sequential model. KERAS. Keras [online]. Available at: https://keras.io/guides/sequential_model/

[53]   JONES, Steven P., Ralph JANSEN and Robert L. FUSARO. 1997. Preliminary Investigation of Neural Network Techniques to Predict Tribological Properties. Tribology Transactions [online]. 40(2), 312-320. Available at: http://www.tandfonline.com/doi/abs/10.1080/10402009708983660

[54]   VELTEN, K., R. REINICKE and K. FRIEDRICH. 2000. Wear volume prediction with artificial neural networks. Tribology International [online]. 33(10), 731-736. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0301679X00001158

[55]   RADHIKA, N., K.T. VIJAYKARTHIK and P. SHIVARAM. 2015. Adhesive wear behaviour of aluminium hybrid metal matrix composites using genetic algorithm. Journal of Engineering Science and Technology [online]. School of Engineering, Taylor's University, 10(3), 258-268. Available at: https://www.researchgate.net/publication/281943697_Adhesive_wear_behaviour_of_alumini um_hybrid_metal_matrix_composites_using_genetic_algorithm

[56]   THANKACHAN, Titus, K. SOORYA PRAKASH, V. KAVIMANI and S. R. SILAMBARASAN. 2021. Machine Learning and Statistical Approach to Predict and Analyze Wear Rates in Copper Surface Composites. Metals and Materials International [online]. 27(2), 220-234. Available at: https://link.springer.com/10.1007/s12540-020-00809-3

[57]   ARGATOV, Ivan I and Young S CHAI. 2021. Artificial neural network modeling of sliding wear. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology [online]. 235(4), 748-757. Available at: http://journals.sagepub.com/doi/10.1177/1350650120925582

[58]   ALTAY, Osman, Turan GURGENC, Mustafa ULAS and Cihan ÖZEL. 2020. Prediction of wear loss quantities of ferro-alloy coating using different machine learning algorithms. Friction [online]. 8(1), 107-114. Available at: http://link.springer.com/10.1007/s40544-018-0249-z

[59]   ALGUR, Veerabhadrappa, Poornima HULIPALLED, V. LOKESHA, Madeva NAGARAL and V. AURADI. 2022. Machine Learning Algorithms to Predict Wear Behavior of Modified ZA-27 Alloy Under Varying Operating Parameters. Journal of Bio- and Tribo-Corrosion [online]. 8(1). Available at: https://link.springer.com/10.1007/s40735-021-00610-8

[60]   RESEARCH, INNOVATION AND DISSEMINATION CENTER FOR NEUROMATHEMATICS. 2013. NES Neuromat [online]. Available at: https://neuromat.numec.prp.usp.br/nes/

[61]   RAJPUT, Ajeet Singh and Sourav DAS. 2023. A machine learning approach to predict the wear behaviour of steels. Tribology International [online]. (vol. 185). Available at: https://linkinghub.elsevier.com/retrieve/pii/S0301679X23002876

[62]   ZHU, Chenrui, Lei JIN, Weidong LI, Sheng HAN and Jincan YAN. 2024. The Prediction of Wear Depth Based on Machine Learning Algorithms. Lubricants [online]. 12(2). Available at: https://doi.org/10.3390/lubricants12020034

[63]   Zygo NewView™ 8300 Specifications [online]. 2014. LAUREL BROOK ROAD • MIDDLEFIELD, CT 06455: ZYGO CORPORATION. Available at: https://www.optixs.cz/data/ke-stazeni/NewView%208300%20Datov%C3%BD%20list-d575.pdf

[64]   CZC.CZ. 2015. Acer Aspire E15 (E5-572G-74LM) [online]. Available at: https://www.czc.cz/acer-aspire-e15-e5-572g-74lm-cerna/165369/produkt

[65]   GOOGLE. 2024. Google Colaboratory [online]. Available at: https://colab.research.google.com/

7. List of used symbols and abbreviations

Abbreviation	Description	Abbreviation	Description
Δw	Weight Loss	ML	Machine Learning
A	Area	MLP	Multilayer Perceptron
a	Radius of Asperity	Mo	Molybdenum
A_n	Nominal Contact Area	MSE	Mean Squared Error
A_t	True Contact Area	N	Normal Force
a, b	Statistical Parameters in Mishina Equation	n	Number of Wear Elements Generated at Junction
a,b,c	Coefficients in Rhee Equation	n	Number of values
Ag	Silver	Nb	Niobium
AI	Artificial intelligence	NN	Neural Network
Al	Aluminium	OOB	Out of bag error in random forest model
Alfa	Wear Coefficient in Goryacheva Equation	p	Pressure
ANN	Artificial neural network	p	Total number of explanatory variables in model
B, C	Wear Coefficients in Yang Equation	p_m	Flow Pressure
Beta	Wear Coefficient in Goryacheva Equation	PCA	Principal component analysis
Beta, n	Coefficients in Queneer Equation	π	Ludolf Number
BN	Boron nitride	ppm	Particles per million
C	Regularization Parameter for SVR Model	R2	Coefficient of determination
CoF / μ	Coefficient of Friction	RELU	Rectified linear unit function
Couple	Material Couple	RF	Random Forest
Cu	Copper	RMSE	Root Mean Squared Error
DN55	Tribometer	s	Average Shear Strength
DNN	Deep neural network	SHAP	SHapley Additive exPlanations
DT	Decision tree	σ_y	Yield Stress
E	Young Modulus	σ_{yf}	Yield Strength of Fragment Material
e	Euler Number	Spec	Specimen (Pin or Plate)
E-P	Extreme Pressure Lubricant	SR	Symbolic regressor
ε	Width of the Street in SVR Model	SSR	Sum of squared residuals
EVS	Explained variance score	SVM	Support vector machine
F	Friction Force	SVR	Support Vector Regression
f	Frequency	t	Time

FFBP	Fast forward back propagation	tanh	Hyperbolic tangent
FNN	Feedforward neural network	Temp	Temperature
GA	Genetic Algorithm	Theta	Fraction of Contacts that Give Rise to Wear Particles
GA	Genetic algorithm	TLU/ LTU	Threshold logic unit/ Linear threshold unit
Gamma	Kernel Coefficient in SVR Model	TSS	Total sum of squares
GBM	Gradient boosting machines	V	Wear Volume
GD	Gradient descent	v	Speed
GP	Genetic programming	V_p	Volume of Spheric Particle
GPR	Gaussian process regression	VAF	Variance accounted for
H	Hardness	var	Variance
h	Wear Depth	W	Work of Adhesion
Hard_ratio	Hardness Ratio	W_d	Depth Wear Rate
HV	Hardness Vickers	W_w	Wear Rate
IDE	Interactive development environment	WC	Tungsten carbide
IF	Isolation forest	X	Feature matrix
K	Wear Coefficient	x_m	Amplitude on DN55
K_s	Steady State Wear Coefficient	X_test	Feature matrix for testing predictions
KNN	K-nearest neighbors	X_train	Feature matrix for training
L	Sliding Distance	XGB	Extreme gradient boosting
L_c	Length of one cycle on DN55	XOR	Exclusive OR logic gate
L1, L2	Regularizations in machine learning	y	Distance of the line contact
lambda	Chemisorption Rate	y	Target vector in ML
LogR	Logistic Regression	ŷ	Predicted value
LR	Linear regression	y_pred	Predicted vector
M_X	Material number in material couple notation	y_test	True values for testing set
MAE	Mean Average Error	y_train	Target vector for training
μ_d	Coefficient of Dynamic Friction	y_i	Actual value
μ_s	Coefficient of Static Friction	-	-

8. List of Figures

9. List of Attachments

dataset_ml.xlsx – excel file with dataset used for machine learning

rychly_goryacheva.ipynb – interactive Python notebook with code for chapter 3 including exploratory data analysis, 3D plot K-α-β and convergence of α, β

rychly_ml_wide.ipynb – interactive Python notebook for chapter 4 including wide search for best ML models

rychly_ml_finetuning.ipynb – interactive Python notebook for chapter 4 including finetuning of 4 best performing model: SVR, RF, GA and NNs.