IT a anatomie firmy

(Řízení a analytika údržby)

(pracovní dokument)



Vojtěch Jiránek

VŠE Praha, 2026



Mapa dokumentu podle kapitol textu (s odkazy)

[1] Vymezení řízení údržby	
[2] Strategie a přístupy v řízení údržby	**[3] Aktuální trendy v oblasti údržby**
[4] Datová analytika v řízení údržby	**[5] Prediktivní údržba**
[6] Vstupní analýza řešení	
[7] Návrh a tvorba řešení – deskriptivní úlohy	**[8] Návrh a tvorba řešení – prediktivní úlohy**

Účelem dokumentu je prezentovat kvalitní a efektivní datově analytické řešení pro různé typy firem, se zaměřením na řízení a analytické úlohy v oblasti údržby, které bude zahrnovat:

- jak ***deskriptivní analytiku***, která bude společnosti poskytovat kvalitní a přehledné reporty o aktuálním stavu,
- tak ***prediktivní analytiku***, která by poskytovala přínosy na základě předpovědí do budoucna, které by společnosti poskytovaly výhodu i v rámci podmínek, které teprve nastanou.

Obsah

1. Vymezení řízení údržby ... 7

1.1 Vymezení údržby ... 7

1.2 Závěry ... 8

1.3 Typy údržby ... 8

1.3.1 Údržba po poruše (Reaktivní údržba) ... 9
1.3.2 Preventivní údržba – s předem stanovenými intervaly ... 9
1.3.3 Preventivní údržba – podle technického stavu ... 9
1.3.4 Prediktivní údržba ... 9
1.3.5 Proaktivní údržba ... 10
1.3.6 Závěry ... 11

1.4 Organizace údržby ... 11

1.4.1 Centralizovaná údržba ... 11
1.4.2 Decentralizovaná údržba ... 11
1.4.3 Kombinovaná údržba ... 11
1.4.4 Externí zajištění údržby ... 11
1.4.5 Outsourcing údržby ... 11
1.4.6 Závěry ... 12

1.5 Typy poškození strojů v rámci údržby ... 12

1.5.1 Erozivní opotřebení ... 12
1.5.2 Únavové opotřebení ... 12
1.5.3 Koroze kovů ... 13
1.5.4 Otlačení ... 13
1.5.5 Deformace ... 13
1.5.6 Trhliny a lomy ... 13
1.5.7 Závěry ... 13

2. Strategie a možné přístupy v rámci údržby ... 14

2.1 Condition based maintenance (CBM). Údržba podle skutečného technického stavu ... 14

2.2 Reliability centered maintenance (RCM). Údržba zaměřená na bezporuchovost ... 14

2.3 Total Productive Maintenance (TPM). Totální / komplexní produktivní údržba ... 15

2.4 Závěry ... 17

3. Aktuální trendy v rámci tématu údržby ... 18

3.1 Průmysl 4.0 a Maintenance 4.0 ... 18

3.2 Průmysl 5.0 a Human-centric Maintenance ... 18

3.3 E-Maintenance ... 18

3.4 Digital Twin ... 19

3.5 Čištění olejů ... 19

3.6 Závěry ... 20

4. Datová analytika a řízení údržby ... 21

4.1 Hlavní KPI a metriky ... 21

4.1.1 MTBF – Mean Time Between Failures ... 21
4.1.2 MTTF – Mean Time To Failure ... 21
4.1.3 MTTR – Mean Time To Repair ... 21
4.1.4 OEE – Overall Equipment Effectiveness ... 21
4.1.5 Equipment Downtime ... 22

4.1.6 Asset Utilization ... 22
4.1.7 RAV – Remaining Asset Value ... 22
4.1.8 SMCP – Standard Maintenance Cost per Unit ... 22
4.2 Data důležitá pro údržbu ... 22
4.3 Závěry ... 23
5. Údržba a prediktivní analytika – Prediktivní údržba ... 24
5.1 Prediktivní analytika a údržba ... 24
5.2 Proces a hlavní kroky prediktivní údržby ... 24
5.2.1 Získávání dat (Data Acquisition) ... 25
5.2.2 Přenos a diagnostika dat (Diagnostics) ... 25
5.2.3 Prognózování (Prognostics) ... 25
5.2.4 Health management ... 26
5.3 Vybrané metody prediktivní údržby ... 26
5.3.1 Data Mining ... 26
5.3.2 Machine learning ... 26
5.3.3 Deep learning ... 27
5.4 Často používané algoritmy ... 27
5.4.1 Neuronové sítě ... 27
5.4.2 Random Forest ... 28
5.4.3 k-nearest neighbor ... 28
5.5 Obecně používané analýzy ... 28
5.5.1 Vibrační analýza ... 28
5.5.2 Termografická analýza ... 29
5.5.3 Olejová analýza ... 29
5.5.4 Akustická analýza ... 29
5.6 Rizika a bariéry zavedení datové analytiky ... 29
5.6.1 Bezpečnost dat ... 29
5.6.2 Kvalita a dostupnost dat ... 30
5.6.3 Správa a konektivita zařízení IoT ... 30
5.6.4 Výběr algoritmu ... 30
5.6.5 Odpor lidí a integrace do strategie ... 30
5.7 Závěry ... 30
6. Vstupní analýza řešení ... 31
6.1 Cíle a přínosy datové analytiky ... 31
6.2 Katalog uživatelů ... 31
6.2.1 Role zaměstnanec managementu společnosti ... 31
6.2.2 Role manažer údržby ... 32
6.2.3 Role zaměstnanec oddělení údržby ... 32
6.3 Závěry ... 32
6.4 Analýza faktorů řešení ... 32
6.5 Analýza datových zdrojů ... 32
6.5.1 Výrobní data ... 32
6.5.2 Poruchy ... 33
6.5.3 Údržba ... 33
6.5.4 Zdroje pro dimenzionální tabulky ... 34
6.6 Závěry ... 34
7. Návrh a tvorba řešení – deskriptivní úlohy ... 35
7.1 Definice KPI a metrik ... 35

7.2 Dimenze ... 39
7.3 Znázornění vazeb mezi dimenzemi a KPI ... 42
7.4 Dimenzionální model ... 42
7.5 Reporty deskriptivní části ... 44
7.5.1 Report Spolehlivost ... 44
7.5.2 Report Výkonnost ... 45
7.5.3 Report Efektivita oddělení údržby ... 46
7.5.4 Report Ekonomika údržby ... 47
7.5.5 Report Stav aktiv ... 48
7.6 Závěry návrhu a tvorby deskriptivního řešení ... 48
8. Návrh a tvorba řešení – prediktivní úlohy ... 49
8.1 Definice cílů prediktivní analytiky – predikční úlohy ... 49
8.1.1 Predikce – Porucha do 72 hodin ... 49
8.1.2 Predikce – Typ poruchy ... 49
8.1.3 Predikce – Remaining Useful Life (RUL) ... 49
8.2 Analýza dat z pohledu prediktivní analytiky (EDA) ... 50
8.2.1 Základní charakteristiky datasetu ... 50
8.2.2 Cílové proměnné ... 51
8.2.3 Závěry analýzy dat z pohledu prediktivní analytiky ... 54
8.3 Feature Engineering ... 54
8.3.1 Klouzavé a rozdílové proměnné ... 54
8.3.2 Historické poruchové proměnné ... 55
8.3.3 Závěr Feature Engineering ... 55
8.4 Definice cílových proměnných (Target variables) ... 56
8.4.1 Target – Porucha do 72 hodin ... 56
8.4.2 Target – Typ poruchy ... 56
8.4.3 Target – RUL ... 56
8.4.4 Závěr definice cílových proměnných ... 57
8.5 Data preprocessing ... 57
8.5.1 Seřazení datasetu ... 57
8.5.2 Tvorba dataframe pro každou úlohu ... 57
8.5.3 Odstranění NaN hodnot hours_since_last_failure ... 58
8.5.4 Výběr vhodných features ... 58
8.5.5 Odstranění zbylých NaN hodnoty ... 58
8.5.6 Rozdělení dat na trénovací a testovací ... 58
8.5.7 Škálování hodnot ... 59
8.5.8 Závěr Data preprocessing ... 59
8.6 Tvorba a trénování prediktivních modelů ... 59
8.6.1 Modely – Porucha do 72 hodin ... 60
8.6.2 Modely – Typ poruchy ... 60
8.6.3 Modely – RUL ... 61
8.7 Vyhodnocení a validace modelů ... 62
8.7.1 Způsob vyhodnocení modelů ... 62
8.7.2 Vyhodnocení – Porucha do 72 hodin ... 63
8.7.3 Vyhodnocení – Typ poruchy ... 64
8.7.4 Vyhodnocení – RUL ... 66
8.7.5 Závěr vyhodnocení a validace modelů ... 68
8.8 Interpretace a vizualizace výsledků ... 68
8.8.1 Interpretace a výsledky – Porucha do 72 hodin ... 68
8.8.2 Interpretace a výsledky – Typ poruchy ... 70
8.8.3 Interpretace a výsledky – RUL ... 71

8.9 Závěr návrhu a tvorby prediktivního řešení 73

Zdroje 74

1. Vymezení řízení údržby

Účelem je vymezení údržby a popis obecných principů a přístupů, nastínění aktuálních trendů v rámci údržby, analýzy vztahu údržby a deskriptivní a prediktivní analytiky a určitá rizika a bariéry, týkající se zavedení a provozování údržby.

1.1 Vymezení údržby

Definice údržby může existovat v mnoha různých formách a podáních. Jednu z definic přináší norma ***ČSN EN 13306 Terminologie údržby*** a ta ji definuje jako *„kombinace všech technických, administrativních a manažerských opatření během životního cyklu objektu, zaměřených na jeho udržení ve stavu nebo jeho navrácení do stavu, v němž může vykonávat požadovanou funkci“* (Česká agentura pro standardizaci, 2018).

Z dobře fungující údržby pak vyplývají ***mnohé efekty***, mezi které můžeme zařadit:

- snižování negativních následků poruch plynoucích především z neplánovaných prostojů,
- zvyšování efektivnosti hlavního výrobního procesu,
- dlouhodobé pozitivní ovlivnění ekonomické stránky fungování společnosti (Rada, 2022).

Mezi ***cíle údržby*** můžeme zařadit následující (Kasalová, 2016; Pechová, 2025):

- předcházení poruchám a jejich minimalizace pomocí preventivních oprav a výměn,
- zajištění efektivního a nepřetržitého provozu stroje po dobu jeho nastavené životnosti,
- snížení dopadu poruch a prodloužení životnosti zařízení a majetku,
- zajištění bezpečnosti provozu a zaměstnanců a prevence nehod či zranění,
- snižování výrobních nákladů (Kasalová, 2016; Pechová, 2025).

Po staletí prováděli opravy na strojích buď sami výrobci anebo i pouze samotní uživatelé. Změna přišla až z **průmyslovou revolucí**, kdy se začali objevovat profese údržbáře, zodpovědného za opravu zařízení a následně i specializace údržbářů (Rada, 2022). Následně přišel postupný vývoj, jelikož jak se vyvíjel průmysl, stejně tak se musela posouvat i údržba, aby vyhovovala požadavkům aktuální doby.

V dnešní době můžeme mluvit o ***čtvrté generaci***. Ta je umožněná masivním rozšířením internetu do všech oblastí lidského života. Prvky využívanými v průmyslu jsou IoT, Big Data, Cloud, AI, 3D tiskárny a mnoho dalšího. Tyto moderní technologie dali možnost vzniku prediktivní údržby, která je označována jako Údržba 4.0 (dle průmyslu 4.0). (Rada, 2022).

V moderních společnostech se tak velmi často ustupuje od klasického řízení údržby a přechází se ke spojení řízení majetku a jeho údržby. To dobře nastiňuje přechod k trendu, že údržba začíná být jasně vnímána jako součást jakéhokoliv vlastněného majetku a přidává se k managementu majetku. V rámci managementu údržby existují tři podoblasti, které jsou součástí i hlavního managementu. Je to ***strategický management***, který stejně tak jako ten celopodnikový plánuje a sleduje strategické cíle podniku v rámci údržby a vytváří strategii budoucího rozvoje. Zahrnuje analýzy současného stavu, formulaci cílů a volbu, tvorbu, a především realizaci strategie. Další je ***taktický management***, jenž se zabývá postupy a prostředky, jak co nejefektivněji dosáhnout cílů strategie. V rámci své činnosti konkretizuje strategické záměry, snaží se zvyšovat produktivitu, nastavuje organizaci a určuje kritéria měření výkonu. Poslední je ***operativní management***, který je ve spodu hierarchie a má na starosti nastavení každodenního fungování údržby. To zahrnuje rutinní činnosti a řešení problémů, incidentů a poruch. Současně sleduje plnění nastavených cílů, dodržování pravidel a využívání dostupných zdrojů a financí (Rada, 2022).

Samotný manažer údržby by měl s podporou vrcholového managementu zajistit, že údržba bude poskytována jako součást celkového plánu provozu společnosti a provozní spolehlivosti zařízení tak, aby byly splněny ekonomické provozní podmínky. Politika údržby se stanovuje v souladu s politikou celé společnosti a měla by obsahovat alespoň, jak se bude řídit ***plánování, provedení, posuzování,***

analýza, zlepšování a zajištění údržby. Při plánování údržby musí být vzaty v potaz následující aspekty (Legát et al., 2018):

- cíle provozní spolehlivosti, cíle bezpečnosti a nařízení či předpisy,
- vybraný systém/typ údržby, který chce být implementován,
- metodika optimalizace údržby,
- zdroje, kterými společnost disponuje,
- spoluodpovědnost vrcholového managementu (jak velkou roli bude mít v rámci údržby) (Legát et al., 2018).

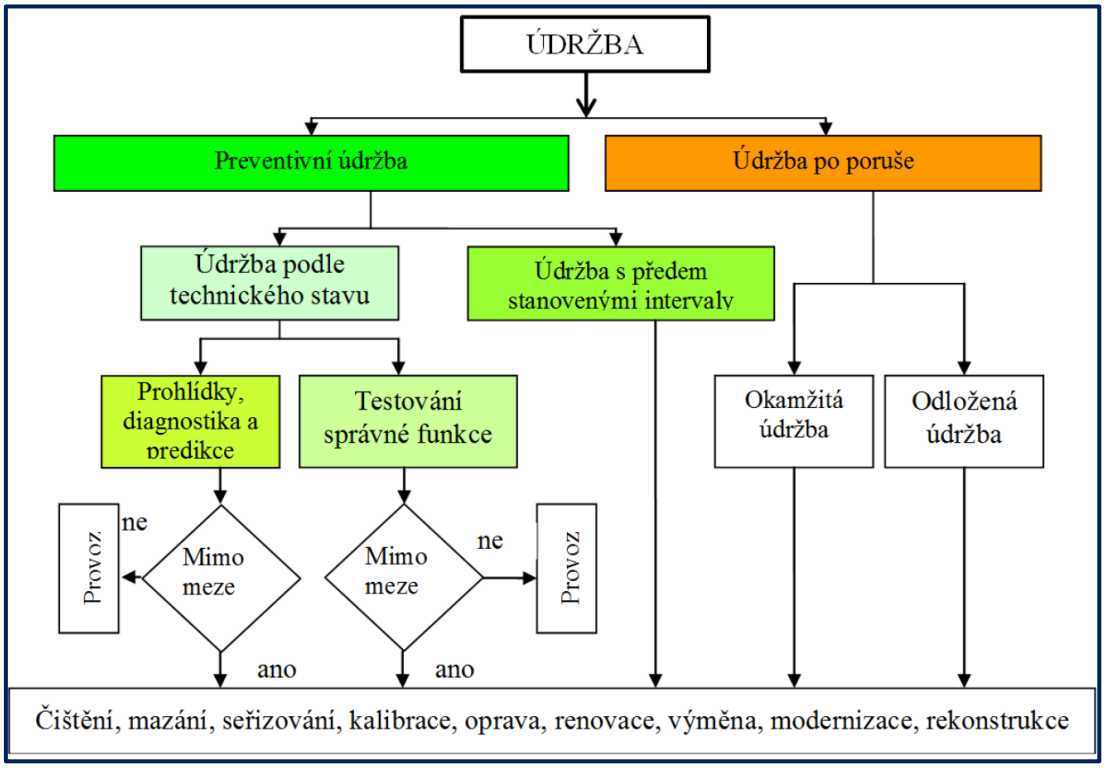
1.2 Závěry



- Vymezení údržby tvoří nezbytný ***teoretický základ*** pro pochopení tématu údržby, na který navazuje další řešení a umožňuje ***zasadit údržbu do kontextu společnosti***.
- Řešení má založeno na použití těch ***nejaktuálnějších postupů a nástrojů***.
- Firemní prostředí může významně ovlivnit ***dostupnost dat, volbu ukazatelů a KPI*** a další.

1.3 Typy údržby

Typy údržby se liší se především podle toho, ***jakým způsobem je údržba prováděna***. Hlavní rozdíly jsou v tom, zda již došlo k poruše či nikoliv. Rozhodnutí o typu prováděné údržby musí být součástí ***strategie managementu údržby***. Na základě vybraného typu ***se pak mohou lišit*** vniklé prostoje, jejich množství, finanční náklady, finanční ztráty a využití jiných zdrojů, jako je lidská síla a další. Základní rozdělení údržby dokumentuje Obrázek 1-1.



Obrázek 1-1: Dělení údržby (Legát et al., 2018)

Hlavní rozdělení je na Údržbu po poruše a Preventivní údržbu. Ta po poruše může být rozdělena buď na okamžitou nebo odloženou údržbu. Preventivní se dělí na Údržbu podle technického stavu a Údržbu s předem stanovenými intervaly. Na ně pak navazují i údržby prediktivní a proaktivní.

1.3.1 Údržba po poruše (Reaktivní údržba)

Dodnes se používá díky ***jednoduchosti*** a díky tomu, že na jakémkoliv stroji ***může vzniknout nečekaná závada*** i s tou nejlepší přípravou. Základní princip je, že se vyčkává, dokud na zařízení nevznikne porucha a ***reaguje se až na poruchový stav***. Následně je cílem, co nejrychleji uvést stroj do stavu, kdy může opět plnit svou funkci. ***Negativní stránky*** jsou množství neplánovaných odstávek a prostojů a nadbytečné náklady na náhradní díly (Kasalová, 2016; Legát et al., 2018; Rada, 2022).

Dělí se na:

- údržbu ***okamžitou***, kdy dojde k opravě ihned, což bývá žádoucí ve většině případů,
- údržbu **odloženou**, což může vzniknout například v případě zajišťování náhradního dílu z externího zdroje (Kasalová, 2016; Rada, 2022).

1.3.2 Preventivní údržba – s předem stanovenými intervaly

Periodická údržba, kde se provádějí kontroly, činnosti či jiné úkony v souladu s předem stanovenými intervaly, ***v určitém počtu cyklů*** nebo v rámci daných nastavených jednotek používání (danou věc lze využít jen po omezený počet použití/množství apod.). Preventivní údržba probíhá i ***bez předešlé kontroly skutečného stavu*** dané části/stroje/součástky. Může se stát, že někdy nastane porucha i po provedení tohoto úkonu a následuje údržba po poruše.

V praxi ***nastavení intervalů*** není zcela snadné a musí se vycházet z aspektů, jako je doporučení výrobce, odhadu, zkušenosti se zařízeními, statistické vyhodnocení apod. Některé kontroly či výměny mohou být dány i zákonem nastavenými pravidly pro danou věc. Předpokládá se u ní ***důsledné plánování potřebné práce*** a často vede ke snížení nákladů na opravy a ponížení možnosti vzniku závady a tím pádem nechtěných prostojů a finanční ztráty z nich plynoucí. Stinnou stránkou je, že při provádění nastavených úkonů či výměně částí nevykazujících poruchu může dojít k jinému poškození daného stroje/součástky. (Kasalová, 2016; Legát et al., 2018; Rada, 2022).

1.3.3 Preventivní údržba – podle technického stavu

V pravidelných intervalech se provádí ***monitoring a kontrola zařízení/ součástky*** / výkonnosti/parametrů a na základě diagnostiky stavu dané věci se realizuje údržba. Sledování může být prováděno i nepřetržitě nebo na vyžádání. Po diagnostice se ***rozhodování provádí dvoustavově***: vyhovuje nebo nevyhovuje. ***Metody***, které se zde používají pro sledování stavu jsou například vibrodiagnostika, termodiagnostika, hluková diagnostika anebo lze použít i lidských smyslů jako je sluch, hmat, čich, zrak a k tomu zkušenosti s danými stroji.

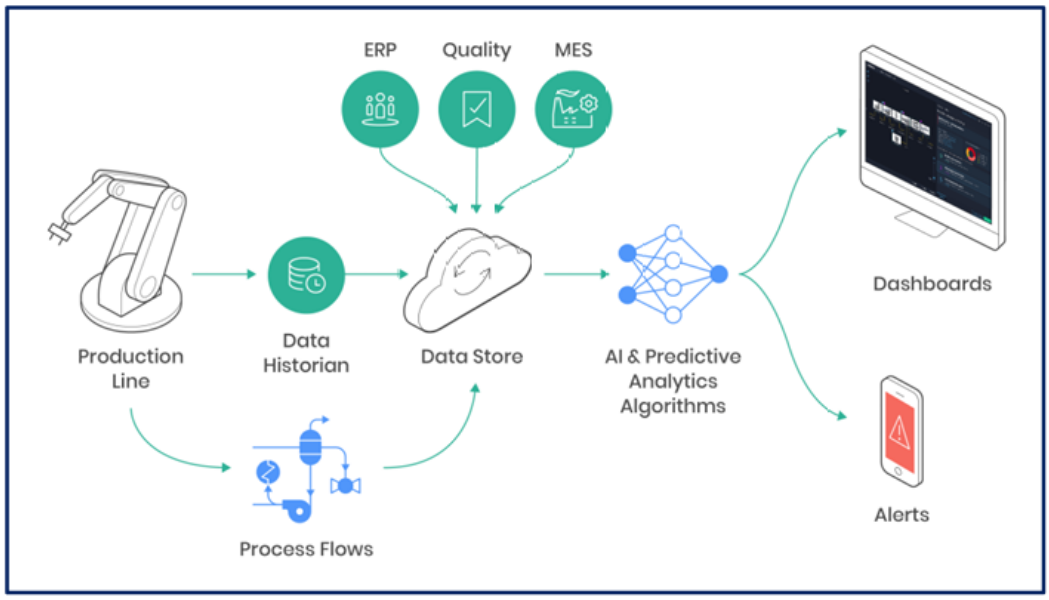
Na základě smyslů lze sledovat znaky jako přehřátí, hluk, kouř, barvu, fyzické poškození a jiné. Mimo člověka existuje obrovské množství ***snímačů či senzorů***, které mohou sledovat velké množství dat a informací a vypozorovat tak změny i v pro člověka nezaznamenatelných jevech. Tyto senzory se používají v místech, kde se stav může měnit náhle a senzor je pak schopný změnu zaznamenat okamžitě. V místech, kde může být změna pozorována až dlouhodobě se naopak kontrola může provádět až na požádání.

Nevýhodou je, že tento typ sledování, senzory a diagnostické přístroje jsou velmi nákladné, a proto se musí velmi pozorně určit, kde je potřeba provádět tuto kontrolu a jaká data je potřeba sledovat. ***Výhody*** a přínosy jsou, že se realizuje tam, kde je opravdu potřebná. Velmi efektivně může zlepšovat poznatky o strojích, snižovat výskyt poruchových stavů, zvyšovat bezpečnost, zásah se provádí, až když je to opravdu potřeba a má minimální vliv na životní prostředí (Kasalová, 2016; Legát et al., 2018; Rada, 2022).

1.3.4 Prediktivní údržba

Prediktivní údržba se používá, stejně jako preventivní údržba ***podle technického stavu***, metod ***diagnostiky***. Narozdíl od té preventivní ale získané informace, sesbíraná data a výsledky z diagnostik používá k tomu, aby se snažila predikovat / ***předpovídat možné budoucí či potenciální poruchy***, změny různých parametrů, optimální čas pro provedení údržby a jiné stavy do budoucnosti ještě, než nastanou. Stanoví pro ně řešení či opatření, které dané poruše předejde nebo alespoň sníží jeho vliv. Tedy předpověď budoucí poruchy na základě analýz (Obrázek 1-2).

Mezi ***metody***, které se pro ni využívají, patří například neuronové sítě, Machine-Learning, multiparametrické analýzy nebo dnes již aktuální umělá inteligence. Její součástí je i ***odhalení příznaků***, které by mohly způsobit poruchu, výpočet pravděpodobnosti vzniku poruchy nebo předpověď následků a závažnosti, pokud porucha vznikne. Jednotlivé stroje či ***zařízení jsou rozděleny do kategorií kritičnosti*** a na základě toho jsou používané nástroje schopny i určit kritičnost vlivu poruchy, jaké poruchy by měly být řešeny přednostně a počítat kritičnost konkrétní možné závady. Takové přístupy jsou samozřejmě velmi nákladné a často také náročné na implementaci. Na druhou stranu jejich vliv a přínos může být extrémně vysoký. Proto je důležité držet se pravidla zavádět prediktivní analytiku všude tam, kde je to technicky možné a finančně výhodné (Kasalová, 2016; Legát et al., 2018; Rada, 2022).



Obrázek 1-2: Schéma prediktivní údržby (MBI tým, 2023)

1.3.5 Proaktivní údržba

Proaktivní údržba je typ, který je ***založený na údržbě po poruše, preventivní údržbě a prediktivní údržbě.*** Navíc se při provádění údržby také sleduje, zda daná porucha nemá nějaké další ***konkrétní příčiny***. To znamená, že při poruše nebo potenciální poruše se provádí i technická, technologická nebo konstrukční opatření, která mají za úkol odstranit příčinu poškození daného dílu/součástky/stroje. Jako konkrétní praktický příklad může být, že než údržbář vymění součástku, například ložisko, tak se přesvědčí, zda příčinou poruchy a jeho krátké životnosti není důvod jako špatné mazání, vibrace či špatná pozice. Následně nejdříve odstraní tyto příčiny a pak vymění ložisko, což prodlouží jeho budoucí životnost. ***Proaktivní údržbář*** se tak snaží zabránit pravidelně se opakující poruchám, které mají jasné příčiny, mezi které může patřit špatně provedená údržba, špatná instalace, špatné nastavení nebo špatná obsluha (Kasalová, 2016; Legát et al., 2018).

Prvky, které proaktivní údržba přináší navíc jsou:

- důraz na rozbor příčin poruch a jejich aktivní odstraňování,
- technický stav je předvídán do budoucna a komplexněji se zpracovává diagnostika,
- posílení úlohy týmu a zahrnutí všech, kteří se na správné funkci stroje nějak podílejí,
- vyšší důraz na dlouhodobou stabilitu funkčnosti zařízení (Kasalová, 2016).

S ***plánováním budoucí údržby*** na nějakém zařízení by se mělo začít ještě před příchodem daného zařízení. Měly by se zvážit aspekty jako možnost údržby konkrétního stroje, cenu a dostupnost náhradních dílů, zainteresované osoby, možnost komunikace s dodavatelem či instrukce a spolehlivost zařízení. Následně se ***sestaví tým***, který bude za zařízení zodpovědný a ten by měl dát dohromady veškerou dokumentaci, potřebné díly, parametry a další nutné věci potřebné pro provádění proaktivní údržby na daném zařízení. Na základě nich ***se pak sestaví plán údržby, intervaly preventivní údržby, diagnostické metody, přístroje,*** které budou využity a další (Kasalová, 2016).

1.3.6 Závěry



- Vývoj ukazuje, že ***jednotlivé typy se nevylučují***, a naopak se často prolínají a kombinují ve společnosti.
- Hlavní pozornost je dále věnována ***prediktivní údržbě***. Ta nejvíce využívá datovou analytiku.
- Řešení musí ale zohlednit i ***prvky proaktivní a preventivní údržby***, jelikož jsou úzce propojeny.

1.4 Organizace údržby

Existuje ***pět základních typů***, které musí každá firma zvážit při zavádění údržby a je jasné, že za určitých podmínek mohou využívat i více typů organizace najednou. Každá z nich má samozřejmě svá pozitiva a negativa.

1.4.1 Centralizovaná údržba

Forma organizace spočívající v existenci ***jednoto centrálního orgánu***, který má na starosti provádění veškeré údržby a všech činností. Údržbáři jsou dle specializace ***rozděleni do týmu a přiřazeni specificky ke svým strojům*** nebo činnostem, které musí spravovat. ***Výhodou*** je velká profesní úroveň zaměstnanců nad danými stroji/zařízeními/oborem a také vybavení speciálním nářadím, které potřebují pro specifické úkony. ***Nevýhodou*** zde je komplikovaná komunikace mezi údržbáři a obsluhou strojů a také neznalost okolností a pracovních podmínek, ve kterých zařízení pracují (Kasalová, 2016; Rada, 2022).

1.4.2 Decentralizovaná údržba

Forma, kdy je specializovaná údržba ***rozdělena do skupin***. To znamená, že ke ***každému zařízení je přiřazena skupina***, která obsahuje nejen zaměstnance výroby, ale i údržbu. Často se zde využívá i údržba přímo zaměstnanci výroby. K ***výhodám*** této formy patří velmi dobrá znalost pracovních podmínek zařízení a dobrá komunikace se spolupracovníky v rámci jak údržby, tak právě např. výroby a obsluhy stroje. ***Nevýhodami*** však je nejednotné odborné vedení jednotlivých týmů, horší využití materiálu a náhradních dílů a nedostatečné vybavení údržby (Kasalová, 2016; Rada, 2022).

1.4.3 Kombinovaná údržba

Jde o ***kombinací formy centralizované a decentralizované***. To znamená, že ve společnosti ***existují oba typy skupin***, kdy ***centralizované*** jsou ty, které potřebují vysokou úroveň profesní specializace a odbornost (to mohou být například specializované práce, nákup náhradních dílů, IT služby atd.). ***Decentralizované*** skupiny naopak spolupracují s týmem výroby anebo se právě přímo stávají jeho součástí (Kasalová, 2016; Rada, 2022).

1.4.4 Externí zajištění údržby

Znamená ***nákup služby*** u externí společnosti, a to u naprosté většiny společností. Často se jedná o nákup a objednání údržby na speciální úkony. Důvody mohou být například ***specializované práce***, které se provádějí jen jednou ***za dlouhé časové období***, práce, které musí být prováděny jen oprávněnou licencovanou firmou nebo úkon, vyžadující použití drahých, nedostupných a speciálních zařízení. ***Výběr externí firmy*** bývá často snadný, jelikož v blízkém okolí jich bývá malé množství anebo se rozhoduje podle ceny. Odpadá zde nutnost zařizování a skladování speciálních součástek, jelikož je firma zajistí a nutnost kupovat speciální drahá zařízení, které bude mít firma k dispozici (Rada, 2022).

1.4.5 Outsourcing údržby

Spočívá ***v dlouhodobém zajišťování*** buď celé, ale většinou části údržby ***externím dodavatelem***. Společnost se musí rozhodnout, jakou bude mít strategii a zda chce část údržby outsourcovat.

Následně se musí rozhodnout, do jaké míry chce nahradit vlastní údržbu externí a ***stanovit si smluvní požadavky a výše plateb***.

Důvody pro outsourcing bývají většinou navýšení kvality a odbornosti údržby a snížení finančních nákladů. S outsourcingem odpadá neustálá nutnost nákupu dílů, nákupu vybavení, nákupu zařízení pro údržbu a pouze se platí za využívané služby. Samozřejmě s sebou ale outsourcing nese i jistá rizika a nevýhody a v případě nespokojenosti může dojít k ukončení spolupráce a vráceni zodpovědnosti do rukou vlastní údržby (Kasalová, 2016; Rada, 2022).

Výhody a důvody pro outsourcing:

- snížení nákladů
- zvýšení kvality
- nedostatek vlastních zkušeností
- snížení stavu zaměstnanců
- provede speciální licencované práce
- dostupnost speciálních zařízení
- nedostatek vlastních zkušeností (Kasalová, 2016; Rada, 2022)

Nevýhody a hlavní důvody proti outsourcingu:

- složitá kontrola údržby
- neúměrné nebo příliš vysoké náklady ku odvedené práci
- obavy z nekvalitně provedené práce
- závislost na externí firmě
- složitost zadání požadavků externí firmě
- nezkušenost nebo nevybavenost dodavatele
- obava z úniku interních dat, informací nebo know-how (Kasalová, 2016; Rada, 2022)

1.4.6 Závěry



- V praxi lze používat ***pouze jednu vybranou***, ale ve většině případů se setkáme ***s kombinací***, kdy může být například centralizovaná s částí úkonů outsourcovaných.
- Největší vliv bude mít ***struktura společnosti***, kdy se od výchozího stavu bude odvíjet, jakým způsobem by mělo být datově analytické řešení nastaveno a zaváděno, jaký je tok informací či jací jsou aktéři.

1.5 Typy poškození strojů v rámci údržby

Existují tyto typy poškození/opotřebení, které mohou vzniknout na strojích a součástkách v průmyslových podnicích. Je důležité vědět, s čím se v rámci takových strojů údržba potýká a co například výrobní linky může poškodit.

1.5.1 Erozivní opotřebení

Jedná se o ***proces, kdy kapalina nebo plyn unáší nějaké částečky*** tvrdšího materiálu anebo částečky obsahuje samotná proudící látka. Tyto částečky narážejí do materiálu, který je kolem nich, a poškozují ho. Dochází k rýhování a seřezávání materiálu a typický výsledek je ***poškozený zvrásněný a nerovnoměrný materiál*** s prohlubeninami. Tyto materiály pak musí být vyměněny (Pošta et al., 2002; Rada, 2022).

1.5.2 Únavové opotřebení

Typ opotřebení, které vzniká na součástkách a materiálech, které procházejí ***opakovaným vysokým kontaktním tlakem*** mezi dvěma tělesy v cyklickém pohybu. Klasickým příkladem je vzájemný pohyb ozubených kol. Opakovaným kontaktem a působením tlaku začnou v kolech vznikat praskliny na povrchu zubů, které se nazývají ***únavové praskliny***. Praskliny se opakujícím se kontaktem postupně zvětšují, až začne docházet k vylamování kusů materiálu. K šíření poškození pomáhá i mazivo, které

vniká do prasklin, tam je uzavíráno, zvyšuje tlak a zrychluje šíření prasklin (Pošta et al., 2002; Rada, 2022).

1.5.3 Koroze kovů

Proces, kterým se označuje ***změna povrchu kovu a ztráta jeho vlastností*** na základě elektrochemických a chemických vlivů okolního prostředí, kdy kov oxiduje a vzniká koroze. Koroze často probíhá ***současně s jinými typy opotřebení***, jejich vliv se kumuluje a rychleji snižují funkčnost kovu a zvyšují opotřebení. Pro kovy jsou vždy uvedené ***korozní rychlosti***, které označují rychlost intenzity koroze a lze podle nich určit, zda bude součást schopna pracovat v korozním prostředí. Koroze ***se dělí*** na plošné koroze (po celém povrchu zařízení), místní, a atmosférická, která vzniká, když je materiál ve velmi vysoké vlhkosti. ***Chránit kov před korozí*** pak můžeme například úpravou prostředí, kde se materiál nachází, elektrochemickou ochranou, úpravou povrchu nebo konstrukčním řešením (Pošta et al., 2002; Rada, 2022).

1.5.4 Otlačení

Ve chvíli, kdy na materiál působí ***vnější tlaky a síly*** způsobené například nějakým jiným objektem, dochází k trvalé změně povrchu či tvaru materiálu, tedy otlačení. V případě, že jsou na první pohled znatelné, lze to brát už jako poškození. Pokud je materiál dostatečně pevný, tak materiál neubývá, ale dochází k jeho přelévání na jiné místo, tedy změně tvaru. ***V závislosti na změně tvaru*** dané součástky pak může dojít k výskytu opotřebení na jiných místech nebo opotřebení i u jiných součástek v jeho blízkosti (Pošta et al., 2002; Rada, 2022).

1.5.5 Deformace

Typ opotřebení, kdy dojde ***k viditelné změně geometrického tvaru nějaké součástky***. To může být způsobeno opět působením vnějších vlivů, kdy může dojít například k ohybu materiálu. Zde ale záleží na typu materiálu, jelikož u křehčích materiálů, jako je sklo, je i mírná deformace moc na to, aby materiál vydržel, dojde ke zlomu, a to už může být i zničení součástky. Může dojít i k ***deformaci vnitřní***, kdy se působením tlaku deformuje součástka uvnitř, ale zvenku to na první pohled znatelné není. I to ale může součástku poškodit. K deformaci může dojít i působením vysoké teploty, kdy, pokud je materiál náchylný na teplo vystaven pro něho příliš vysoké teplotě, může dojít ke změně tvaru. (Pošta et al., 2002; Rada, 2022).

1.5.6 Trhliny a lomy

Trhlina je typ poškození, kdy dojde u součástky přímo k narušení jejího povrchu ***ve tvaru viditelného řezu***. Pokud se řez vyskytne v celé šířce dané součástky, tak by to bylo nazýváno lomem. Způsob jejich vzniku je podobný, jako tomu bylo u otlačení a deformace, tedy působení vnějšího tlaku a sil, které překročí pevnost nebo unavenost materiálu a způsobí, že materiál nějakým způsobem praskne. Výskyt trhlin ze své podstaty ***způsobí narušení pevnosti a výdrže materiálu*** a po nějaké době většinou končí lomem. Velmi zde ale záleží na křehkosti a houževnatosti materiálu, tedy jeho schopnosti absorbovat vnější energii. (Pošta et al., 2002; Rada, 2022).

1.5.7 Závěry



- Typy poškození strojů a cesty, ***jakým způsobem se zařízení opotřebovávají***, jsou zajímavé z technického hlediska pro všechny, kdo se na opravách strojů podílejí.
- Umožňuje to lepší pochopení příčin poruch a ***hledání lepších preventivních opatření***.
- Pro datově analytické řešení ale není klíčový samotný fyzikální mechanismus poškození, ale spíše ***důsledek, tedy porucha či zhoršení funkčnosti*** zařízení.
- ***Analytické aplikace*** pracují hlavně s provozními daty, indikátory a historií poruch, které tyto mechanismy odrážejí.
- ***Popis poškození*** slouží více jako podpůrný kontext pro pochopení vzniku poruch než jako cíl analytického zkoumání a predikcí.

2. Strategie a možné přístupy v rámci údržby



Účelem je vymezení několika ***strategií a přístupů***, jak je možné celkově vést a zavádět údržbu v podniku. Uvedené konstrukty nelze brát zcela jako typy údržby, ale jsou to spíše přístupy, které kombinují různé vlastnosti, postupy a specifika.

Těmi nejznámějšími jsou CBM, RCM a TPM.

2.1 Condition based maintenance (CBM). Údržba podle skutečného technického stavu

CBM je ***strategie preventivní údržby***, která se zaměřuje na ***monitorování a sledování skutečného stavu*** zařízení či stroje a současně jeho potenciálního poklesu funkčnosti či poškození. Využívají se k tomu kontroly jeho vizuálního stavu, senzory a další techniky a zařízení, díky kterým sledujeme trend stavu zařízení buď v reálném čase nebo v pravidelných intervalech.

Pokud objevíme ***degradaci stavu zařízení*** (pokles jeho efektivity, pravidelnější poruchy, opotřebení), tak se na základě toho optimalizuje údržbový plán, jako může být častější a intenzivnější údržba, abychom předešli náhlým poruchám a časovým prostojům. To pak znamená, že ***zařízení se vyřadí z provozu až ve chvíli, kdy je údržba nevyhnutelná***, a ne po uplynutí nějakého nastaveného časového období (IBM.com, 2023; Puškár, 2024).

CMB ***lze spojit s prediktivní údržbou***, která díky předpovědím ještě zefektivní tento přístup. Úprava údržbového plánu je pak prováděna nejen na podobě aktuálního stavu zařízení, ale právě i ***na předpovědích***, které nám mohou pomoci udělat tuto úpravu plánu dopředu s patřičnou přípravou. Na rozdíl od preventivní údržby v pravidelných intervalech, která dělá údržbu na základě daného rozpisu, ***CBM umožňuje identifikovat potřeby stroje*** a nastavit údržbu efektivněji a finančně výhodněji (IBM.com, 2023; Puškár, 2024).

V rámci CBM lze používat ***různé praktiky***, mezi které patří například:

- infračervená termografie – využívá termální obraz pro detekci přehřívání nebo jiných výkyvů spojených s teplotou
- monitorování vibrací – také vibrační analýza, využívá senzory snímající vibrace, pro detekci a měření jejich frekvence a identifikaci výchylek,
- analýza tlaku – využívaná především v zařízeních, kde musí být daný nějaký tlak a senzory vyhodnocují změny či pokles tohoto tlaku,
- analýza olejů – vyhodnocuje kvalitu a podmínky olejů a jiných kapalin, sledujíce například jejich viskozitu či kyselost pro detekci například kontaminace (IBM.com, 2023).

Životní cyklus CBM je velmi přímočarý a skládá se z částí plánování, implementace, monitorování, analýzy a konečné zlepšení údržby. ***Výhody CBM*** jsou takové, že zabraňuje selhání vybavení a prostojům, zvyšuje životnost zařízení, zvyšuje bezpečnost, zvyšuje efektivitu údržby a snižuje náklady. ***Nevýhodami*** ale je, že má vysoké náklady na implementaci, sběr dat může být složitý, integrace musí být komplexní a důkladná a nese s sebou nutnost vysoké bezpečnosti dat (IBM.com, 2023).

2.2 Reliability centered maintenance (RCM). Údržba zaměřená na bezporuchovost

Jejím hlavním cílem je ***zvyšovat spolehlivost*** zařízení tak, že se konstantně snaží ***identifikovat kořenové příčiny poruch***, být na ně připravená a tím snižovat náklady. To znamená, že se snaží určit každý jeden prvek zařízení, který by mohl selhat. Při správném pochopení možností, jak by mohl selhat, ***lze optimalizovat údržbový plán*** a nastavit úkony tak, aby se jim průběžně snažili předcházet anebo na ně byly připravené.

Na rozdíl od CBM tedy ***nereaguje na vývoj*** stavu zařízení, ale již ***na všechny možné příčiny***, jak by potenciální porucha mohla vzniknout. Zásahy preventivní údržby a managementu se určují i podle provozních, bezpečnostních a ekonomických následků poruch. Jejím ***výsledkem je posudek*** k provedení údržbářského úkonu, změny zařízení či jiná potenciální možnost pro zlepšení (Kasalová, 2016; Puškár, 2024).

RCM má své ***základní definované principy***, kterými je potřeba se řídit:

- Primárním cílem RCM je zachovat funkce systému.
- Identifikovat potenciální způsoby poruchy stroje či součástky, které mohou způsobit nefunkčnost zařízení.
- Klasifikovat způsoby poruch dle stromu logického rozhodování.
- Hledat užitečné a efektivní preventivní činnosti (Kasalová, 2016).

Tyto čtyři základní ***principy pak byly rozpracovány do navazujících otázek,*** které by si měl každý uživatel RCM položit:

1. Jakou funkci a výkon má zařízení splňovat v chtěných provozních podmínkách?
2. Co může způsobit selhání této funkce zařízení?
3. Co může zapříčinit funkční poruchy?
4. Co se stane při vzniku této poruchy?
5. Jak porucha ovlivní zařízení?
6. Co můžeme udělat preventivně, abychom předešli této poruše?
7. Co můžeme udělat, pokud není k dispozici žádné preventivní řešení (Kasalová, 2016)?

Stejně jako CBM má pak RCM svůj daný ***životní cyklus***. Začíná se ***zahájením a plánováním***, kdy se například stanovují cíle a obsah analýz či se identifikují zkušenosti, specializace a odpovědnosti. Výsledkem je plán analýzy. Pokračuje se právě ***analýzou poruch funkce***, kde dochází například k určení možných poruch, jejich důsledků a kritičnosti. Následuje ***volba úkonu*** na základě předchozí analýzy, kdy se určí i potřebné intervaly úkonů. Pak je ***implementace a výsledkem je plán údržby***. Na to již opět navazuje neustálé zlepšováni a zvyšování efektivity vytvořeného systému (Kasalová, 2016).

Výhody této strategie jsou, že zlepšuje efektivitu údržby, poskytuje návod k jejímu zvládnutí, snižuje náklady na údržbu, vytváří dobře zdokumentované procesy údržby a možnost jejich neustálého zlepšování a dobré pochopení cíle a účelu údržby všemi zainteresovanými. ***Nevýhodou*** pak je opět velmi složitá a nákladná tvorba, složitost analýz a implementace a současně při změně zařízení musí dojít i ke změnám analýz a úkonů (Kasalová, 2016; Puškár, 2024).

2.3 Total Productive Maintenance (TPM). Totální / komplexní produktivní údržba

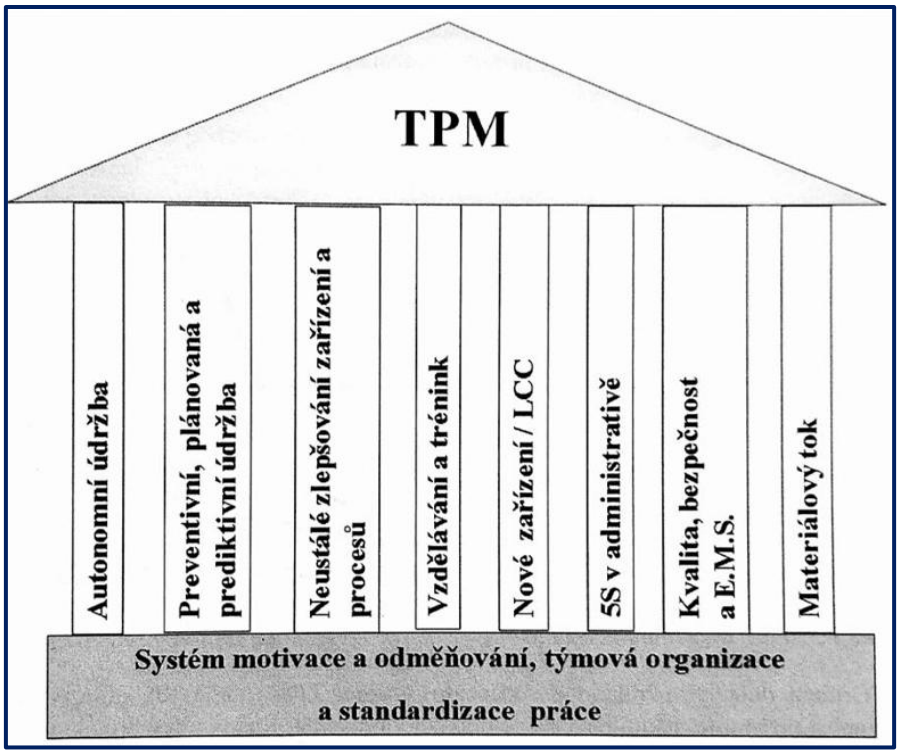
TPM je ***strategie*** údržby ***založená na proaktivním typu*** údržby. Často je označována jako typ filozofie pro efektivní údržbu. Důraz se zde klade na ***odstranění dělení pracovníků*** na obsluhu stroje a údržbu a ***zapojení všech*** zaměstnanců do procesu údržby. TPM velmi silně zasahuje až do samotné firemní kultury. Předpokládá se, že není nutné vše přenechat údržbě a ten, co stroj obsluhuje, může jako první zaznamenat drobné změny či nedostatky na stroji, zapojit se do procesu, který normálně provádí údržbáři a efektivně tak odhalením maličkostí či abnormalit předejít budoucím potenciálním poruchám. Toto se nazývá tzv. ***autonomní údržba*** a je efektivní, protože obsluha má největší znalosti o daném zařízení a dodává to zodpovědnosti a důležitosti obsluhy vůči své roli.

TPM pak vyžaduje a zahrnuje velmi úzkou ***spolupráci všech*** tedy operátorů, údržby, ale i managementu, vedení a dalších. Věří, že touto úzkou spoluprací se pak dosáhne nejefektivnějšího předcházení poruch a navazujících prostojů a co nejrychlejší reakce na vznikající poruchy. Mimo samotných poruch, tak také ***vypomáhá předcházet chybám na výrobcích***, neshodám a přešlapům výroby, zvýšení bezpečnosti a předcházet možným úrazům, což jsou další benefity TPM. Zavádí se tedy ***nejen jako údržba, ale celopodnikově, i v rámci prodeje, financí, vývoje***, nákupu či jiné administrativy, právě aby se zapojili zcela všichni a především všichni, co mohou nějak přispět (Kasalová, 2016; Puškár, 2024; Rada, 2022).

TPM byla vytvořena Japoncem jménem Seiche Nakajima a ten také definoval ***8 základních pilířů*** správně fungujícího TPM:

- autonomní údržba,
- preventivní, plánovaná a prediktivní údržba,
- neustálé zlepšování zařízení a procesů,
- vzdělání a trénink,
- nové zařízení, LCC (Life Cycle Cost),

- 5S v administrativě (metoda pro organizaci pracoviště založená na 5S),
- kvalita a bezpečnost,
- materiálový tok (Kasalová, 2016).



Obrázek 2-1: Pilíře TPM (Rada, 2022)

Zavádění TPM se pak dělí do **čtyř základních fází** a celkově se jedná o **12 kroků**, které je potřeba splnit (Kasalová, 2016).

1. ***Fáze: Příprava***
   1) Oznámení vrcholovým vedení o zavedení TPM.
   2) Start vzdělávání a kampaně o TPM.
   3) Vytvoření organizační jednotky pro TPM.
   4) Vytvoření vize a politiky TPM.
   5) Formulace hlavního plánu.
2. ***Fáze: předběžná implementace***
   6) Zahájení celého TPM.
3. ***Fáze: implementace TPM***
   7) Zlepšování výkonnosti zařízení.
   8) Návrh autonomní údržby.
   9) Návrh plánu údržby pro oddělení údržby.
   10) Zahájení tréninků operací údržby.
   11) Vytvoření časového programu managementu zařízení.
4. ***Fáze: stabilizace***
   12) Zlepšování TPM a jeho rozšiřování (Kasalová, 2016).

Výsledky zavedení a výhody pak jsou ***zlepšování kvality, snížení nákladů, zvýšení morálky a zlepšení bezpečnosti*** na pracovišti. Dále pak snížení poruchovosti, zkrácení doby údržby a snížení nákladů na údržbu, zkrácení výrobních časů, zlepšení procesů a mnoho dalších. V konečném důsledku je pak výsledkem i ***zvýšení konkurenceschopnosti*** celé společnosti. Udává se, že zlepšení pak můžeme sledovat již během šesti měsíců. ***Nevýhody*** je opět velmi náročná implementace. Díky své

složitosti a zásahu do celé struktury společnosti je to nevyhnutelné. Stejně tak může být problémem odpor zaměstnanců k novým postupům a dále nutnost kvalitního a pravidelného proškolování zaměstnanců, které je nevyhnutelné a zásadní (Kasalová, 2016; Puškár, 2024; Rada, 2022).

2.4 Závěry



- Uvedené strategie údržby nejsou všechny, ***existuje řada dalších***.
- ***CBM*** je založená na skutečném stavu.
- ***RCM*** je založená na základě rizika.
- ***TPM*** zasahuje do celé organizace společnosti.
- ***Analytické řešení*** pak slouží jako podpora pro rozvoj a rozhodování v rámci dané strategie.

3. Aktuální trendy v rámci tématu údržby

Účelem je specifikace trendů, které v současnosti panují v oblasti údržby. Zahrnují několik konceptů, které aktuálně převládají. Je důležité pak určit, jakým směrem se novodobá údržba ubírá a co lze čekat do budoucna.

Jedním z těchto trendů je aktuálně i právě ***Predictive maintenance*** (***prediktivní údržba***) a je to jeden z těch největších trendů.

3.1 Průmysl 4.0 a Maintenance 4.0

Koncept Průmysl 4.0 se dostal a adaptoval i do oboru údržby, kde vytváří velký posun. Vznikají ***obrovské propojené systémy komponent, které sbírají data*** z linek a jiných zařízení, jako jsou senzory a čidla. Ty pak posílají data v reálném čase s informacemi o jejich stavu, spolehlivosti a efektivitě do systémů, které je zpracují a poskytují okamžitě velmi přesná data a výsledky. Ty pak pomáhají dělat rozhodnutí pro co nejlepší údržbu a další oblasti jako je správa majetku. Data se ***průběžně ukládají*** a pracovník je může mít kdykoliv dostupná. To nese i ***negativa***, jelikož když jsou data přenášena přes internet, tak je riziko, že se k nim dostane i někdo cizí a zneužije je. (Rada, 2022; Štaffenová et al., 2021).

3.2 Průmysl 5.0 a Human-centric Maintenance

Průmysl 5.0 se stále snaží procesy automatizovat, ale vzniká v něm opět nové ***silné zaměření na člověka (human-centric)***, který v té čtvrté byl spíše mimo. Člověk nesmí nikdy být z procesu odstraněn, měl by být vždy alespoň do určité míry přítomen a stroj by měl především podporovat, zjednodušovat a zefektivňovat jeho práci. ***Stroj nemůže udělat svou práci sám bez kontroly*** a schopnost uvažování a zasáhnutí člověka musí být zásadní. Oba dva tak mají zásadní roli a dle 5.0 oba musí koexistovat, nikoliv jeden vyřadit druhého (Cummins et al., 2024).

Na jeho základě tak vznikají i nové trendy a koncepty v rámci údržby. Mezi tyto nové vznikající pojmy pak patří například ***XAI (eXplainable AI) nebo iML (interpretable Machine Learning)***. Tyto dva koncepty jsou v současné době velmi intenzivně zkoumány z různých směrů a jeden z těchto směrů je právě i efektivní využití v prediktivní údržbě. Využití těchto dvou pojmů a spojením s prediktivní údržbou nám pak vzniká pojem ***XPM = eXplainable Predictive Maintenance***. Z praktického hlediska to pak znamená, že například AI nedá pouze tvrdé výsledky a tvrzení, že se stroj rozbije, ale poskytne i kontext, důvody, rozhodovací logiku a další důkazy, proč by se tak mělo stát, a to poskytuje naprosto transparentní informace pochopitelné pro kohokoliv. (Cummins et al., 2024).

3.3 E-Maintenance

E-Maintenance je systém, který lze jednoduše definovat jako ***údržbu, kde jsou stroje monitorovány a řízeny přes internet***. Je to síť, která ***integruje a spojuje jednotlivé aplikace údržby*** a zařízení tak, aby sbíraly společně informace a doručily je tam, kde jsou potřeba. Tento technologicky vyspělý, efektivní a rychlý systém je schopný při správném použití dovést společnost, aby dosáhla téměř nulových prostojů. Tento sběr dat a jejich okamžité sdílení je pak schopný pomoci společnosti k vysoké úrovni transparentnosti a maximálně efektivní okamžité informovanosti. Systémy se pak ***dají integrovat s dalšími odlišnými systémy*** společnosti a sesbíraná data mohou být propojena s daty jiných oddělení či dokonce jiných závodů, čímž získáme okamžitě výsledky, které se dříve velmi složitě počítaly a vyhodnocovaly. To velmi markantně zefektivňuje a zjednodušuje práci i tisíců zaměstnanců a jiných zainteresovaných stran (Iung et al., 2007).

Tento koncept dnes dostává ***obrovskou pozornost***, díky svému vlivu na operování společnosti a procesy. Velmi zvyšuje bezpečnost, zvyšuje spolehlivost, snižuje náklady obecně a také snižuje náklady na údržbu, jelikož se odhaduje, že ***náklady na údržbu mohly běžně činit 15–40 %*** z ceny vyrobeného produktu a e-Maintenance toto ***tlačí dolů***. V konečné důsledku pak ovlivňuje i samotnou konkurenceschopnost společnosti. Proto je tento systém čím dál rozšířenější. Nese s sebou ale i ***negativa***, jelikož jeho zavedení může být velmi složité a nákladné, především pro společnosti, které nejsou tak dobře technologicky vybavené. Neustále ***vyšší komplexita a složitost výrobních zařízení a***

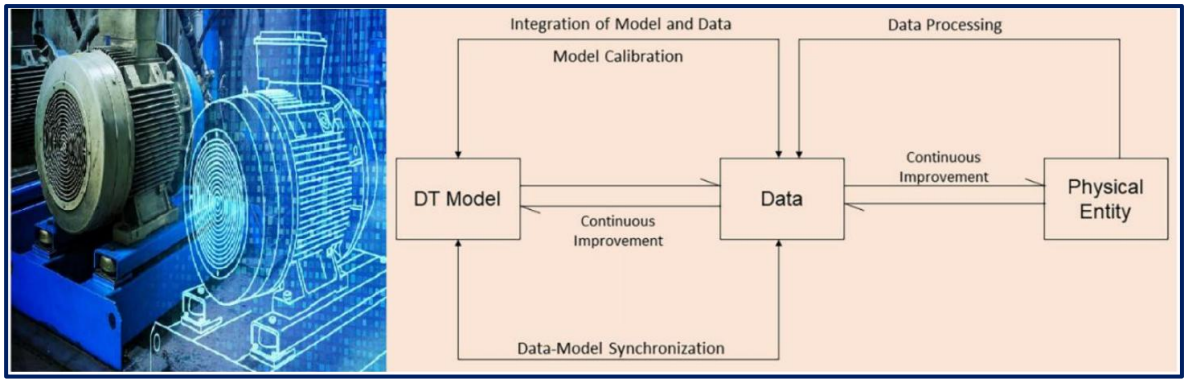
produktů také klade vyšší nároky na správné zavedení tohoto systému. Výzkum v oblasti e-Maintenance však stále pokračuje a její důležitost v průběhu času neustále roste (Štaffenová et al., 2021).

3.4 Digital Twin

Digital twin neboli digitální dvojče je koncept, kdy se v digitálním prostředí ***vytvoří digitální kopie nebo digitální verze*** nějakého fyzického objektu, stroje, zařízení nebo systému. Dnešní technologie je již schopná toto udělat do těch nejmenších detailů, které potřebujeme. Toto dvojče je následně ***napojeno na svůj reálný protějšek***, získávajíce data a jiné informace z reálného funkčního systému, samo s nimi může pracovat a následně virtuálně vytvářet a testovat scénáře, které by mohli nastat. To umožňuje otestovat funkce, reakce, chování a potenciální situace bez toho, abychom jakkoliv narušili fyzický objekt, který musí fungovat.

Tento proces nám pomáhá na základě otestovaných scénářů následně ***vylepšovat či upravovat dané zařízení*** tak, abychom předešli potenciálním poruchám nebo chybám anebo se alespoň připravili, až tento scénář nastane. Velmi jednoduchý praktický příklad může být, co se stane s výrobní linkou, pokud by nám vypadla elektřina. Digitální dvojče toto dokáže nasimulovat v kontrolovaném prostředí, bez narušení výroby. Dokáže ***okamžitě sesbírat všechna data***, a navíc tento scénář můžeme provádět opakovaně bez sebemenšího problému. Tím dokážeme získat veškerá potřebná data, umíme připravit se dopředu na vznik tohoto problému a ve chvíli, kdy nastane, tak jsme schopní přesně a okamžitě zareagovat.

Společně s tím jsme pak schopni dělat i ***opatření, aby tento scénář vůbec nenastal***. Mimo tento příklad, kdy se scénáře testují na dvojčeti, tak naopak i ***dvojče je schopné se získanými daty z reality provádět různé operace***, data upravovat a optimalizovat a pokud je potřeba, tak vrací tato vylepšená data reálnému objektu. Ten je pak schopný na jejich základě lépe pracovat a zlepšit svou funkci. Největší rozdíl mezi digital twin a jinou simulací je pak skutečnost, že ***není statický a díky konstantní komunikaci s reálným objektem a dostupností dat se model neustále učí*** a upravuje svoji funkci na základě přesné aktuální situace reálného objektu (Hussien Gomaa, 2024).



Obrázek 3-1: Digital Twin (Hussien Gomaa, 2024)

3.5 Čištění olejů

Oleje jsou součástí všech strojů, které jsou kolem nás a je jasné, že ***jejich likvidace je velmi složitá***. Musíme při ní být velmi opatrní, aby nedošlo ke znečištění životního prostředí, proces podléhá spousta předpisům a často může být i velmi nákladný. O to horší je to v případě, kdy nějaký stroj dokáže projít tisíci litrů oleje za kratší časový úsek. ***Kvalita daného oleje***, a to jak kvalita samotného produktu, tak čistota oleje samotného ve stroji, může zásadně ovlivnit kvalitu a efektivitu jeho funkce.

Proto jsou existují koncepty, které jsou schopné velmi ***dobře eliminovat znečištění oleje*** ve stroji a současně i jeho celkovou spotřebu. Jedná se například ***o elektrostatické čištění olejů***, které umožňuje na základě chemických procesů čistit olej přímo ve stroji i bez přerušení výroby. Náplň oleje se díky ní stává trvalou a vysoce snižuje objem průmyslového odpadu. Dalším konceptem/nástrojem je ***obtoková filtrace olejů***, která se vyrábí tak, aby se připojila přímo ke stroji. Každou chvíli je pak ze stroje odebíráno miniaturní množství oleje, aby nedošlo k ovlivnění funkce stroje, odebraný díl je profiltrován a následně navrácen do stroje. (Rada, 2022).

3.6 Závěry



- Aktuální trendy v řízení údržby dobře znázorňují, že je zde ***posun směrem k digitalizaci, podpoře rozhodování***, maximální využití dat a integrace s dalšími systémy.
- Nejvýznamnější jsou ***Maintenance 4.0, Human-centric maintenance a e-Maintenance***, které úzce souvisí s prediktivní údržbou.
- ***Digital twin*** má také využití v prediktivní údržbě, ale společně s ***čištěním olejů*** nejsou přímo předmětem tohoto dokumentu.

4. Datová analytika a řízení údržby



Účelem je charakterizovat datovou analytiku a to, jak významně zasahuje i do údržby. Často bývá součástí analytiky ***investic a údržby***, jelikož každý jeden vlastněný stroj, linka, zařízení nebo i budova je investicí společnosti, která může buď přinášet hodnotu anebo ji bude ztrácet-

4.1 Hlavní KPI a metriky

Samotná údržba má svá hlavní KPI a metriky, které se k ní obecně váží. Tyto ***metriky jsou pro údržbu relativně univerzální*** a jsou použitelné ve velkém množství různých průmyslových a dalších společností. ***Výběr metrik*** pak závisí i na konkrétním podniku a co potřebuje sledovat. Další uvedené metriky lze pokládat za základní.

4.1.1 MTBF – Mean Time Between Failures

MTBF znamená ***průměrný čas mezi poruchami***. Reprezentuje průměrnou dobu, po kterou zařízení operuje od poslední poruchy do další poruchy. Pomáhá určit spolehlivost zařízení a předpovídat potřeby údržby pro dané zařízení. Vypočítá se jako ***Celkový čas provozu / Počet poruch***. Pro jednoduchý příklad můžeme vzít stroj, který skutečně pracoval 500 hodin (plánovaný čas výroby mínus prostoje) za daný měsíc a za daný měsíc zaznamenal 4 poruchy. MTBF by v tu chvíli vycházelo jako 125 hodin. Vedoucí údržby podle toho pak mohou upravit pracovní plán. (MaintainX.com, 2025).

4.1.2 MTTF – Mean Time To Failure

MTTF znamená ***průměrný čas do selhání***. Měří průměrnou dobu, po kterou zařízení pracuje, než dojde k jeho selhání a je neopravitelné. Je to užitečná metrika pro neopravitelná či spotřební zařízení, pro které je typické, že mají danou dobu a pak musí být vyměněny (např.: žárovka, filtr apod.). Rozdíl od MTBF je v tom, že MTTF sleduje jen ***neopravitelné*** poruchy. Vypočítá se jako ***Celkový čas provozu / Počet zničených kusů***. Pro příklad máme krabici 100 žárovek, které by měly dohromady fungovat 50 000 hodin. Za tu dobu ale 25 z nich přestane fungovat. 50 000 děleno 25 nám pak dá MTTF s hodnotou 2000 hodin. Používá se k určení životnosti zařízení (MaintainX.com, 2025).

4.1.3 MTTR – Mean Time To Repair

MTTR znamená ***průměrný čas do opravy***. Měří průměrný čas, který je potřeba na opravu zařízení a jeho navrácení do funkčního stavu od poruchy. Tento kritický ukazatel přímo ovlivňuje dostupnost zařízení a kapacitu výroby. Vypočítá se jako ***Celkový čas všech oprav / Počet oprav***. Pokud by tým údržby strávil na jednom zařízení 20 hodin v průběhu 5 různých oprav, MTTR by pak vycházelo na 4 hodiny. U tohoto ukazatele naopak cílíme na nižší hodnoty, jelikož to znamená kratší opravy a efektivitu údržby. Tato hodnota ale může být ovlivněna komplexností daného zařízení (MaintainX.com, 2025).

4.1.4 OEE – Overall Equipment Effectiveness

OEE znamená ***celková efektivita zařízení***. Počítá celkovou efektivitu kombinací tří faktorů: dostupnost, výkon a kvalita. Celkově dává dobrou informaci o tom, jak kvalitně naše zařízení funguje. Vypočítá se pak jako ***dostupnost * výkon * kvalita***. Praktický příklad by byl, že můžeme mít zařízení, které má fungovat po směnu 8 hodin (480 minut). Naneštěstí zařízení hodinu nefungovalo (60 minut). ***Dostupnost*** by byla ***87,5 %*** (420/480). Mělo vyrábět 100 kusů za hodinu, ale vyrábělo jen 80. ***Výkon*** by byl ***80 %*** (80/100). Ze 7 hodin práce vyrobila 560 kusů, ale jen 540 bylo v pořádku. **Kvalita** by byla ***96 %.*** Pokud to vložíme do rovnice, vyjde nám ***OEE 67,4 %.*** Z popisu ukazatele je jasné, že chceme vyšší hodnoty a obecně OEE kolem 85 % je považováno za velmi dobré (MaintainX.com, 2025).

4.1.5 Equipment Downtime

Equipment Downtime sleduje ***celkové procento času, po které zařízení nefunguje*** z důvodu údržby, oprav nebo poruchy. Je to kritický ukazatel pro operativní efektivitu a spolehlivost zařízení. Vypočítá se jako ***(Celkový čas prostojů / Celkový čas možného provozu) * 100***. Vezměme, že zařízení nefunguje 20 hodin z celkový 200 hodin provozu. Celkový ***equipment downtime*** by pak byl 10 %. Zde opět chceme nižší procentuální hodnoty. Vyšší totiž znamenají delší čas, kdy zařízení nefungovalo (MaintainX.com, 2025).

4.1.6 Asset Utilization

Asset Utilization znamená využití aktiva (stroje). Reprezentuje, jak efektivně je zařízení využíváno v porovnání se svým plnohodnotným potenciálem. Ukazuje nám potom, zda je zařízení používáno neoptimálně, optimálně anebo dokonce je zatěžováno. Vypočítá se pak jako ***(Celkový výstup zařízení / Maximální možný výstup) * 100.*** Pro příklad máme zařízení schopné produkovat 1000 kusů za měsíc, ale tento měsíc vyprodukovalo pouze 750. Po vložení do vzorce zjistíme, že zařízení jede pouze na 75 % své potenciální kapacity a nefunguje tedy optimálně. Znovu zde cílíme pro vyšší hodnoty procent. Existuje ale i ***časová Asset Utilization***, která se počítá pomocí času, tedy ***(Reálný čas výroby / Celkový možný čas výroby) * 100***. Výpočet by byl stejný, jen bychom místo kusů použili například hodiny. Oba výpočty jsou možné, jen poskytují jiný pohled na efektivity aktiva (MaintainX.com, 2025).

4.1.7 RAV – Remaining Asset Value

RAV znamená ***zbývající hodnota aktiva***. Počítá aktuální hodnotu zařízení v porovnání k jeho původní nákupní hodnotě. Pomáhá především s rozhodováním, zda je cenově výhodnější udržovat zařízení v chodu a provádět údržbu anebo ho vyměnit. Vypočítá se jako ***(Aktuální hodnota aktiva / Původní nákupní hodnota) * 100.*** Pokud tedy vlastníme zařízení zakoupené za 20 000 Kč a jeho odhadovaná aktuální hodnota je 8 000 kč, tak nám RAV vyjde 40 %. Nižší hodnoty RAV naznačují, že se zařízení blíží ke konci svého použitelného života a hodnoty pod 50 % obecně znamenají předzvěst nutnosti jeho výměny. Proto by toto zařízení bylo potřeba brzy vyměnit a u RAV rozhodně chceme vyšší hodnoty (MaintainX.com, 2025).

4.1.8 SMCP – Standard Maintenance Cost per Unit

SMCP znamená ***standardní náklady na údržbu za kus***. Měří průměrné náklady potřebné na udržení každého jednotlivého zařízení funkční. Pomáhá nám určit efektivitu využití nákladů na údržbu. Vypočítá se jako ***Celkové náklady na údržbu / Počet zařízení***. Velmi jednoduchý příklad tedy bude, pokud máme náklady na údržbu ve výši 100 000 Kč a musíme se starat o 100 strojů, naše SMCP bude 1000 Kč. Zde chceme docílit nižších hodnot nákladů, které znamenají cenově efektivní údržbu. (MaintainX.com, 2025).

4.2 Data důležitá pro údržbu

Pro údržbu pak existuje ***velké množství dat***, které by měly být sbírány a zpracovávány, abychom měli co nejkvalitnější základnu pro deskriptivní i prediktivní analytiky a pro efektivní řízení údržby managementem. Soupis těchto potřebných dát popisuje pan Legát (2018) ve svém článku o údržbě v průmyslu 4.0 (Legát et al., 2018).

Podrobný rozpis dat údržby zahrnuje (Legát et al., 2018):

- ***Všeobecná data***
  - popis a data o umístění objektů,
  - popis úkolů údržby,
  - historie údržby,
  - zprávy o poruchách včetně podmínek zařízení, když porucha vznikla,
  - modifikace objektů,
  - informace o materiálech a náhradních dílech,
  - plánovaná a rozvrhovaná práce údržby,
  - pracovní příkazy,
  - informace a rady pro údržbu na základě nových znalostí a zkušeností,

- stav konfigurace zařízení a data o tomto stavu,
- servisní příkazy a pravidla vydávané výrobci.

- ***Další data***
  - ***Údaje časového charakteru***
    - doby používání, provozu do poruchy a mezi poruchami
    - intervaly prohlídek, údržeb a oprav
    - průběžné doby a pracnosti zásahů
    - časy prostojů (způsobené organizací, pauzami, poruchami, preventivní údržbou, přestavbami atd.)
  - ***Údaje o technickém stavu***
    - teploty
    - napětí a proud
    - opotřebení
    - vůle a průtoky
    - změna účinnosti zařízení
    - reálná výkonnost
  - ***Údaje ekonomického charakteru***
    - náklady na preventivní údržbu
    - náklady na diagnostiku, prohlídky a revize
    - náklady na zásahy po poruše
    - náklady na vzniklé prostoje
    - náklady vyvolané opotřebením
    - výrobní cena jednoho výrobku
    - náklady pracovní, materiální a režijní – interní i externí
    - náklady na zásoby náhradních dílů
  - ***Údaje o spolehlivosti***
    - **Kvantitativní**
      - doba provozu do poruchy a mezi poruchami
      - doba obnovy
      - rozdělení pravděpodobnosti poruchy
      - pravděpodobnost poruchy a bezporuchového stavu
      - doba nezjištěného poruchového stavu
      - doba administrativního zpoždění
      - doba údržby po poruše preventivní údržby
      - počet neshodných výrobků díky chybnému výrobnímu procesu
      - spotřeba náhradních dílů
    - ***Kvalitativní***
      - popis příčiny, odhalitelnosti a důsledku poruchy
      - způsob nahlášení, odstranění poruchy a zajištěnost údržby
      - data o zásobách náhradních dílů (Legát et al., 2018)

4.3 Závěry



- Existují ***metriky, které jsou relevantní*** a je třeba je v rámci řešení sledovat.
- Data ukazují ***rozsah dat***, jaké je možné v údržbě sledovat.
- V rámci řešení práce budou muset být ***cíleně vybrána*** ta, která budou umožňovat ***získat a vyhodnocovat zvolené metriky***.
- Skupina údajů o ***technickém stavu*** je nejobsáhlejší a nejnáročnější na samotnou analýzu a vyhodnocení.
- ***Ekonomické údaje*** pak vyžadují důkladnou spolupráci s oddělením financí a často adekvátní úpravu systému údržby (Legát et al., 2018).

5. Údržba a prediktivní analytika – Prediktivní údržba

Účelem je charakterizovat prediktivní údržbu. Je to údržba, na kterou je ***aplikována prediktivní analytika***, která je jedním ze stupňů pokročilé analytiky. Zaměřuje se na to určit, co se pravděpodobně stane v budoucnosti. Prognózování je zde definováno jako proces, který využívá analytických a datových poznatků a zkušeností k vytvoření předpovědí a odpovědí na otázky týkajících se různých obchodních potřeb.

Vstupem pro řešení je ***prediktivní analytika***. Její podstatné principy a charakteristiky obsahuje dokument na MBI-AF [Prediktivní analytika].

5.1 Prediktivní analytika a údržba

Prediktivní údržba je přístup ***založený na proaktivní údržbě***, který klade důraz na to, co vám chtějí vaše podniková zařízení a stroje sdělit. Tedy, jaká bude jejich životnost, jaký bude jejich stav, jaké mohou být jejich poruchy. Provádí to ***prostřednictvím predikování sledovaných KPI***, metrik a trendů. Pomáhá tedy předvídat do budoucna, ***jak bude vypadat jejich životní cyklus***. Tyto možnosti jsou způsobeny rozmachem dnešních digitálních technologií a počítačů, a především díky implementaci IoT, který pronikl i do údržby a pomohl vytvořit Maintenance 4.0, kterou je ***prediktivní údržba***. Pomáhá tak vytvářet a ***upravovat plán údržby*** a úkony do různorodých intervalů, nezávisle na aktuálním stavu zařízení, právě díky predikcím (Bharati et al., 2025; SAP.com, 2025).

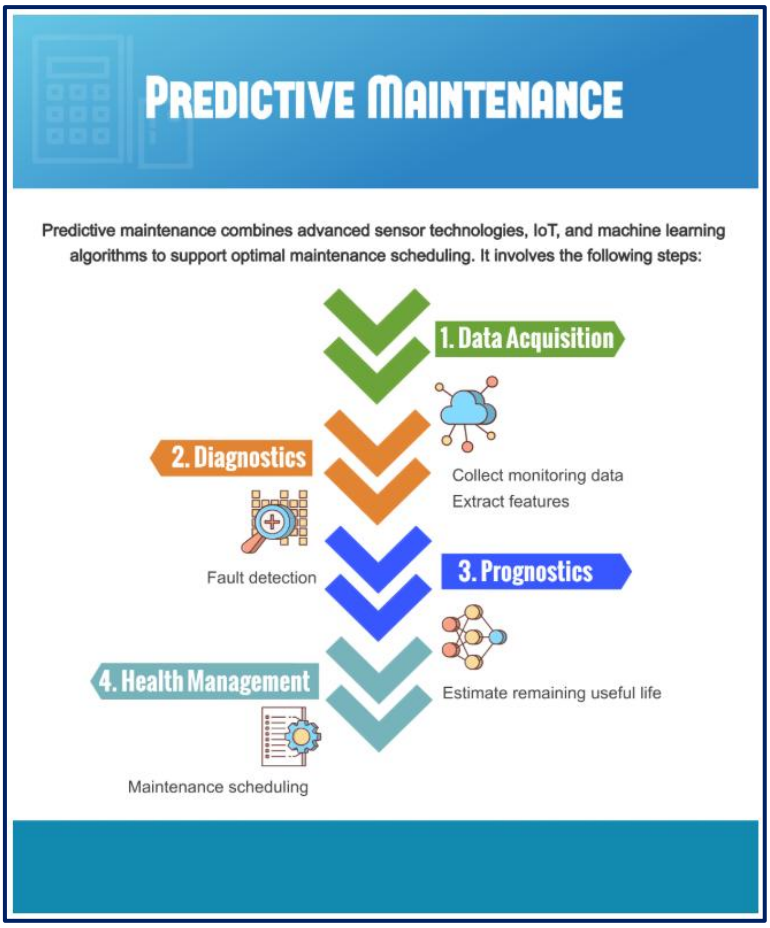
Při tvorbě prediktivní údržby jsou pak důležité ***základní analytické otázky***:

- Existuje dostupný schválený roční plán proaktivní údržby?
- Je plán údržby v souladu s platnou legislativou, doporučeními výrobce a zkušenostmi výroby?
- Jsou stanovena pravidla pro tvorbu plánu včetně kompetencí?
- Je plán v souladu s plánovanými odstávkami zařízení?
- Dochází v případě změny dodávek, služeb či finančních plánů v aktualizaci plánu?
- Existuje automatický provázanost mezi plánem údržby, výroby a financí?
- Využívají se nástroje prediktivní analytiky pro plánování údržby (o Novotný et al., 2024)?

5.2 Proces a hlavní kroky prediktivní údržby

Prediktivní údržba má obecně ***čtyři hlavní kroky***, které popisují postup aplikovatelný v každé průmyslové společnosti. Tento ***proces a jeho kroky je velmi obecný*** a specifika musí každá jedna společnost řešit individuálně a subjektivně na svých podmínkách a stavu.

Obrázek 5-1 4 hlavní kroky, a to získání a shromažďování dat, přenos a diagnostika těchto dat v podnikových systémech, prognózování a finální health management neboli podnikání akcí na základě prognóz (Guo, 2021; SAP.com, 2025).



Obrázek 5-1: 4 hlavní kroky prediktivní údržby (Guo, 2021)

5.2.1 Získávání dat (Data Acquisition)

Získávání dat a měření ***přímo ze zařízení a strojů*** zahrnují např. vibrace, tlak, teplota, hluk a další. To se provádí ***prostřednictvím čidel a senzorů*** instalovaných přímo na zařízení. Dokážou tak extrahovat holý signál a zpracovat ho do užitečné informace, která indikuje stav stroje. Z těchto holých signálů jsou pak schopné vytvořit potřebné informace (Guo, 2021; SAP.com, 2025).

5.2.2 Přenos a diagnostika dat (Diagnostics)

Zde dochází k přenosu dat do podnikových systémů rychlostí v rámci sekund či milisekund. Následně jsou nasbíraná ***data diagnostikována a vyhodnocována*** pomocí inteligentních nástrojů údržby a dochází k ***hledání poruch***, opotřebení nebo potenciálního selhání. Ve většině případů jsou tyto procesy prováděny jako ***klasifikační úlohy/metody***. Mezi nástroje, které data zpracovávají patří například rozhodovací stromy, k-nejbližších sousedů, ale i například Random Forest, umělá inteligence. Z celkového hlediska se snaží ***provést propracované analýzy*** systému a jeho objektů (stroje, zařízení atd.) (Guo, 2021; SAP.com, 2025).

5.2.3 Prognózování (Prognostics)

Dalším krokem a pro prediktivní údržbu tím ***nedůležitějším jsou právě prognózy***. Na základě zpracovaných dat inteligentními nástroji je cílem předpovědět budoucí stav monitorovaného systému a ***určit jeho efektivní délku života***, tedy jak dlouho, než se objeví nějaké selhání nebo porucha. Toto je krok, který dává prediktivní údržbě její podstatu, největší váhu a důvod, proč jsou tyto systémy a metody stále tolik prozkoumávány. Celkovou podstatou je zde zmírnit riziko selhání a prodloužit životnost systému a efektivitu (Guo, 2021; SAP.com, 2025).

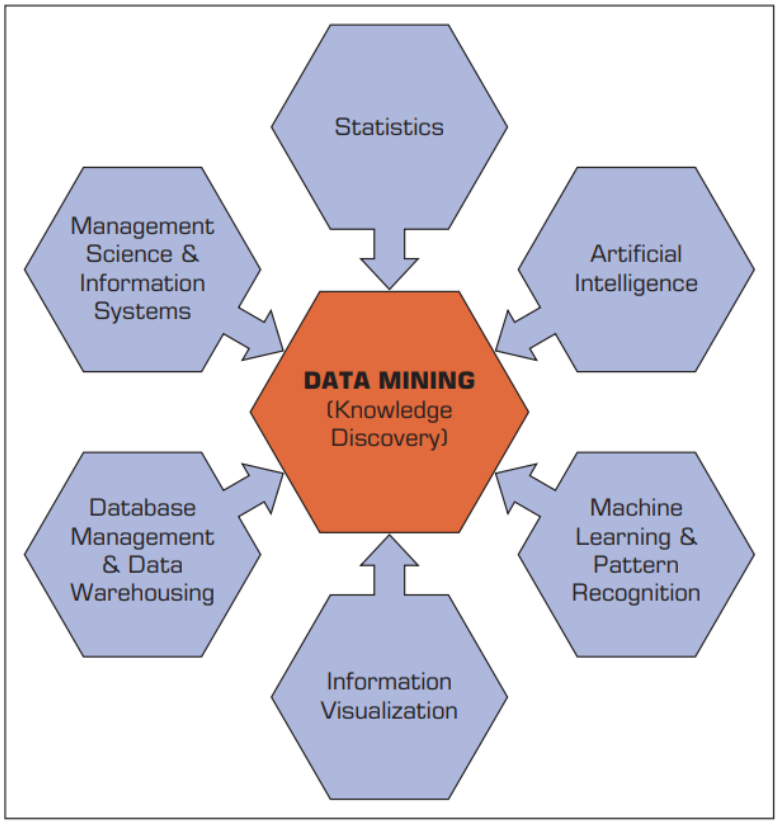
5.2.4 Health management

Finální a logický navazující krok je ***health management neboli podnikání akcí údržbou*** na základě získaných predikcí, aby zužitkovala získané výsledky. Dochází k co nejrychlejšímu přijetí opatření buď s automatizovanou reakcí anebo lidským zásahem. Hlavním ***cílem je řídit údržbu a její plán*** v maximálně optimalizovaném směru. To znamená ***maximalizovat dostupnost, spolehlivost a bezpečnost*** zařízení a stejně tak minimalizovat potřebu údržby a náklady na ni vynaložené (Guo, 2021; SAP.com, 2025).

5.3 Vybrané metody prediktivní údržby

5.3.1 Data Mining

Data mining je termín, který označuje ***proces prohledávání a získávání znalostí z obrovských souborů*** dat. Příbuzné termíny jsou extrakce znalostí, analýza vzorů/trendů, sběr informací či prohledávání dat. Více technicky je to pak proces, který využívá ***statistické, matematické a AI techniky*** pro identifikaci a extrakci užitečných informací a následné znalosti z obrovských souborů dat, která jsou v databázích a jsou to data společnosti z různorodých zdrojů, které spolu nějak souvisí. Spojuje do sebe mnoho pojmů jako je statistika, AI, strojové učení, informační systémy, databáze či věda managementu, jak ukazuje Obrázek 5-2 (Sharda et al., 2020).



Obrázek 5-2: Data mining (Sharda et al., 2020)

Jevy, které se pak ***v rámci data miningu hledají*** jsou asociace v datech, predikce na základě dat, hledání shluků (clustering) rozdělující data do přirozených skupin, hledání sekvenčních vztahů zavádějící data do časově seřazených a souvisejících událostí a mnoho dalšího. Standardizovaná metodika CRISP-DM pak definuje obecný postup, jak provádět data mining.

5.3.2 Machine learning

Machine learning, strojové učení je ***metoda, která má napodobit proces, jakým se učí lidský mozek***, ale využitá v počítačích. Snahou je naučit počítač, jak identifikovat souvislosti a spojení v datech

na množství příkladů a souvisejících dat. Současně jsou počítače schopné monitorovat aktivity v okolí a tím jsou schopné upravit své chování, aby se vypořádaly se změnami ve svém okolí. Je to vědecká metoda zabývající se ***navrhováním a vytvářením algoritmů, které umožňují počítači učit se*** na základě velkého množství dat, které přicházejí ze senzorů, z databází a mnoha dalších zdrojů. Hlavní cíl pak je opět provádět predikce, rozpoznávat souvislosti v datech a podporovat rozhodování lidí (Sharda et al., 2020).

Jsou k tomu zapotřebí počítačové programy, které se vždy ihned učí ve chvíli, kdy je před ně postavena nová situace nebo nový problém. Program pak ***sbírá stále více dat, analyzuje je a natrénuje se***, aby se dostal k řešení či závěru. Následným ukazováním příkladných řešení problémů je pak schopen ***vidět v nich elementy nebo aspekty***, které by běžné oko ani nepostřehlo. Strojové učení je dnes velmi aktuální téma, které je umožněno především díky neustálému zrychlování výpočetní rychlosti, dokonalejším a propracovanějším algoritmům, velmi rychlým databázím a stále rostoucích dostupným zdrojům Big Data (Sharda et al., 2020).

5.3.3 Deep learning

Deep learning, resp. „hluboké učení“, je podkategorií předchozí metody strojového učení. Je to metoda, která se ***ještě propracovaněji snaží napodobovat, jak funguje lidský mozek***. Využívá velmi složitou ***umělou neurální technologii*** a snaží se vypořádat s velmi komplexními problémy, na které i strojové učení a jiné technologie nestačí. Je to koncept, který zdánlivě doslova myslí, kdy s každou jednou další přijatou informací a novými daty upravuje své chování. Pomocí svých velmi pokročilých a složitých algoritmů dokáže řešit i dříve neřešitelné problémy (Sharda et al., 2020).

Deep learning je dnes ***využíván například v autonomních vozidlech***, kde pomáhá rozpoznávat systému auta silniční značky a překážky na vozovce. Současně hraje i kritickou roli v dnešních nejmodernějších telefonech, tabletech, robotech či chytrých domácnostech. Díky své rychlosti a efektivní funkčnosti je nejvíce využitelná ***aplikovaná v interaktivních systémech***, které musí fungovat a zpracovávat data v reálném čase, jako jsou obory strojového vidění, robotiky, rozpoznávání obrazu a rozpoznávání hlasu a zvuků. (Sharda et al., 2020).

5.4 Často používané algoritmy

Samotné algoritmy pro zpracování dat jsou pak vyvinuté technologie a prostředky, které se využívají v rámci metod, které jsme si popsali v předešlé kapitole. Opět je jich obrovské množství a k tomu od každé z nich pak mohou existovat různé variace založené na daném algoritmu. Proto není možné říct, že se v prediktivní údržbě používá ten a ten, a to platí vždy. Můžeme si ale pro nastínění popsané alespoň ty, které jsou využívané velmi běžně.

5.4.1 Neuronové sítě

Tyto modely jsou ***inspirovány biologickou strukturou mozku***. V posledních letech byly neuronové sítě prokázány jako velmi slibný systém pro ***provádění predikcí a klasifikačních aplikací*** podniku díky své schopnosti učit se z dat, neparametrické povaze (nemají rigidní předpoklady) a schopnosti generalizace naučených věcí. Algoritmus je ***součástí metody strojového učení*** a je založená na rozpoznávání vzorů v datech. Celkově se využívají v různorodých oblastech podniků pro právě rozpoznávání vzorů, předpovědi a klasifikaci a dnes už je klíčovým prvkem mnoha nástrojů pro data science a podnikovou analytiku (Sharda et al., 2020).

Bylo zjištěno, že ***model***, který byl podpořen výzkumy mozku a ***má podobnou strukturu jako lidský mozek,*** může jevit podobnou inteligentní funkcionalitu. Biologický mozek je sestaven z obrovského množství masivně propojených neuronů, které dokážou komunikovat se sousedními neurony přenosem a přijímáním elektrických a chemických signálů. Podobně tedy neuronová síť je ***sestavena z mnoha propojených jednoduchých procesních elementů*** tedy umělých neuronů (*artificial neurons*). Při zpracovávání informace pak tyto neurony navzájem komunikují, podobně jako neurony mozku. Proto poté dokáží vykazovat ***podobné schopnosti, jako je učení se***, úprava svého chování nebo tolerance dělání chyb. Díky jejich úspěchu a rozmachu pak začali vznikat i ***Deep neural networks***, které spadají pod metodu Deep learning (Sharda et al., 2020).

5.4.2 Random Forest

Random Forest je typ algoritmu, který je ***variací na metodu*** zvanou ***bagging***, který je jedním z nejběžnějších algoritmů. Jeho podstata je velmi jednoduchá, ale velmi silná. Vytváří ***větší množství rozhodovacích stromů*** na základě vzorků dat vytvořených s využitím ***bootstrapingu***. To je metoda vytváření ***fixního množství instancí/vzorků*** z originálních dat s opakováním záznamů (vybraný vzorek je vždy vrácen do původního datasetu a vytvořen nový, tedy se může stát, že se objeví již použité záznamy a některé vůbec), které jsou ***použity k trénování*** a zbytek zbylých dat je použit pro testování. Toto se zopakuje tolikrát, kolikrát je potřeba a kolikrát je nastaveno. Výsledky a predikce stromů z baggingu jsou pak ***zkombinovány do jednoho výsledku pomocí průměrování nebo hlasování*** (*voting*) (Sharda et al., 2020).

To, co je pak u Random Forest algoritmu modifikací a vlastností ***navíc*** oproti baggingu je, že nejen že využívá bootstrapingu, tedy náhodné vzorky, ale také ***ve vytvořeném stromě v každém jeho dělení provádí náhodný výběr proměnných.*** Proměnou si zde jednoduše můžeme představit jako jeden sloupec v datasetu, který obsahuje nějaké hodnoty stejného typu, které mají vliv na výslednou predikci. Bagging vezme vždy všechny proměnné a na jejich základě provádí predikci/rozhodnutí. Random Forest si náhodně ***vybere pouze určitou podmnožinu těchto proměnných a rozhodnutí*** provádí na jejím základě. To znamená, že Random Forest má náhodný nejen vzorek dat, ale i proměnné, které použije. Tyto parametry, jak se má chovat, jsou nastaveny při tvorbě modelu. Tento proces pak způsobuje, že Random Forest obecně produkuje přesnější předpovědi než právě bagging (Sharda et al., 2020).

5.4.3 k-nearest neighbor

K-nearest neighbor, nearest neighbor nebo také k-NN, resp. metoda k-nejbližších sousedů je jeden z ***nejjednodušších a nejintuitivnějších algoritmů***, které Machine learning nabízí. Na první pohled se může zdát i příliš jednoduchý pro vytváření kompetitivních predikcí, na rozdíl od jiných metod, které vyžadují přítomnost času a jsou výpočetně velmi intenzivní kvůli svým matematickým nárokům. Algoritmus ***se využívá pro buď klasifikaci anebo pro regrese***. Algoritmus funguje na základě, kdy jsou známé vyhodnocené případy/pozorování zaneseny například na graf, každý z těchto případů má svou hodnotu a s každým z těchto případů je spojené známé chování, výsledek nebo následky, které nastaly nebo hodnoty targetu (sledované proměnné). Když se objeví nový případ, určíme jeho hodnotu a na jejím základě ho také zaneseme do grafu. Co chceme sledovat jsou jeho ***nejbližší sousedé***, přesněji jaký byl výsledek/chování v případě těchto nejbližších sousedů. Případně u targetu průměrná hodnota těchto sousedů. ***Proměnná k je pak hyperparametr a vyjadřuje počet těchto nejbližších sousedů, které má model sledovat nastavený předem*.** Na základě porovnání se sousedy můžeme předpovídat/predikovat, jaký bude výsledek tohoto nového případu, protože bude pravděpodobně podobný, jako byl u jeho nejbližších sousedů. Jako jednoduchý ***praktický příklad*** řekněme, že na grafu budou zaneseny body s hodnotami 10, 13, 14 a 17, při nichž došlo k poruše zařízení a body s hodnotami 2, 4, 5 a 6, při kterých k poruše nedošlo. Pokud zaneseme na graf nový případ, který bude mít hodnotu 12, tak můžeme predikovat, že pravděpodobně dojde k poruše, jelikož jeho nejbližší sousedé jsou případy, kdy porucha vznikla. Rozhodování není prováděno pomocí matematických výpočtů, ale na základě vzdálenosti k bodům v prostoru tvořeném různými atributy. Další vlastností je, že model není tvořen dopředu, rozhoduje se až v okamžiku predikce (Hill & Berry, 2021; Sharda et al., 2020).

5.5 Obecně používané analýzy

Účelem zde je charakterizovat ***základní vybrané analýzy, které se provádějí v rámci prediktivní údržby***. Existuje jich velké množství, proto bychom se případně neměli zaměřovat jenom a pouze na ty zde uvedené. Ty jsou ale základní, se kterými se většinou setkáme, pokud jde o prediktivní údržbu.

5.5.1 Vibrační analýza

Vibrační analýza se používá jako součást prediktivní údržby, která ***kontroluje ve většině případů rotující části***, jako mohou být vadná ložiska, a nevyváženosti různých dílů. Při vibrační analýze se sledují a zaznamenávají ***vibrace vykazované takovými díly pomocí senzorů*** a odesílají do systémů společnosti. Následně při analýze těchto záznamů lze na základě amplitud a průběhu vibrací vyvodit v jakém je daná součást stavu. Pokud rozpoznáme odchylky od běžných stavů, tak můžeme rozpoznat abnormální situaci a provést ve chvíli stále fungujícího stroje nápravy, které předejdou budoucí

náhlé poruše, která by měla mnohem větší vliv a časový prostoj. (o Novotný et al., 2024; zeiss.cz, 2025).

5.5.2 Termografická analýza

Termografická analýza je založená na ***sledování teploty zařízení*** nebo jeho součástky. U většiny zařízení je například vysoká teplota vnímaná jako problém. Sledováním součástky, která je nějak spojena s činností, která by mohla generovat teplo, pak můžeme vyčíst, zda je na základě teploty daná součást v pořádku anebo by mohl nastat problém. V takovém případě ***může jevit jedna ze součástek opotřebení anebo přímo poškození***. Pro toto měření se využívají ***optické termokamery a kontaktní a bezkontaktní čidla***. Můžeme pak kontrolovat prvky, jako jsou části rozvaděčů, ložiska nebo třeba neadekvátní napětí gumových dopravníkových pásů. Zvyšování teploty znamená nesprávné usazení nebo nesprávný stav daného dílu, což by v průběhu času mohlo znamenat výrazný problém. Zde je vidět ukázkový příklad využití prediktivní údržby, kdy pokud sleduje změnu či růst teploty, tak dokáže odhadnout, že v určité době dojde k poruše způsobené tímto špatným stavem (o Novotný et al., 2024; Puškár, 2024).

5.5.3 Olejová analýza

Provozování průmyslových strojů bezpodmínečně vyžaduje i používání olejů a maziv, které tyto stroje potřebují ke své existenci a efektivní funkčnosti. Proto je také faktem, že provozování strojů bez pravidelných ***rozborů těchto olejů a zjišťování stavu, co tyto stroje konzumují***, může mít extrémně negativní dopad na jejich funkci. To může vést ke ztrátě jejich produktivity, což vede ke ztrátě tržeb a vysokým nákladům na opravy. Pravidelné ko***ntroly olejů mohou vést k prodloužení životnosti strojů, optimalizaci údržby a snížení nákladů***, protože pomohou identifikovat problém s oleji, než nastane. Tento proces se nazývá ***tribodiagnostika*** a zaměřuje se na analýzu tření. Snaží se v oleji najít přítomnost kovů, organických sloučenin nebo vody, které právě vyvolávají tření a které mohou znamenat opotřebení v různých částech stroje. Jejich nalezení před poruchou znovu znamená včasnou reakci a vyhnutí se nepříjemné poruše, kterou bychom bez prediktivní údržby nezaznamenaly. Nejčastější chyby, které jsou detekovány jsou vnik vody, poškozeni ložisek, použití nesprávného oleje, kaly v olejích nebo nedostatečná preventivní péče (DIAGNOSTIKASTROJU.cz, 2025; o Novotný et al., 2024).

5.5.4 Akustická analýza

Typ analýzy, která se zaměřuje na ***sledování zvukových vln, které vychází ze zařízení*** a na jejich základě se snaží určit potenciální poruchy. Využívá především mikrofony, které zachycují zvukové vlny vytvořené komponentami stroje nebo procesem výroby samotným (něco v zavedeném procesu ve stroji generuje nežádoucí zvuk). Každý stroj, pokud funguje správně a stabilně, produkuje konstantní a přesné zvukové projevy. ***Identifikováním odchylek, zvláštních frekvencí a pravidelností v těchto zvucích lze detekovat chyby v konstrukci, defekty, úniky materiálů*** ve stroji nebo nesprávnou manipulaci se strojem a provést okamžitou nápravu, která předejde poruše a ztrátě efektivity stroje. Další možností pak je i vysílat zvukový signál sám pomocí nějakého zařízení a sledovat, jak se daný signál vrátí, když projde stroje/zařízením. Akustická analýza je schopná detekovat opotřebení, nesprávné posazení komponent, přehřátí, blokace, praskliny, úniky, elektrické anomálie a další (NI.com, 2025; o Novotný et al., 2024).

5.6 Rizika a bariéry zavedení datové analytiky

Rizika a bariéry ***jsou jednou ze základních věcí***, kterou musíme zvážit při tvorbě datové analytiky v rámci prediktivní údržby. Tato rizika se pak mohou měnit v rámci konkrétních řešených případů a konkrétní společnosti.

5.6.1 Bezpečnost dat

Bezpečnost je jedním z hlavních aspektů, na kterou musí být brán ohled. U údržby platí, že toto může být ještě ***zásadnější,*** jelikož ***data o produktivitě, efektivitě a stavu strojů*** mohou znamenat ***velké bezpečnostní riziko***, pokud by se dostala do nesprávných rukou. Současně, jelikož množství dat získaných ze zařízení může být v průběhu času značné, většinou se pro jejich uchovávání používají cloudové služby. To znamená poskytování mnoha kritických dat externí firmě, což s sebou právě nese velké bezpečnostní riziko. (MBI tým, 2023).

5.6.2 Kvalita a dostupnost dat

Prediktivní údržba se do velké míry ***spoléhá na kvalitu a dostupnost dat***, které sahají dostatečně do historie. Pokud by data byla nedostatečná, tak to může znamenat nepřesná vyhodnocení, KPI, metriky a samozřejmě předpovědi. To by pro firmu bylo spíše škodlivé a nemělo by to žádný přínos. Mimo jiných problémů by to způsobilo ***zbytečné náklady, další náklady na provoz nefunkčního řešení*** či opravu, špatně vedenou údržbu a nastavení jejího plánu, nulovou podporu managementu společnosti a mnoho dalšího. Proto je běžnou praxí pro tato řešení vytvořit ještě systémy, které mají na starost správu dat a jejich čištění. Při tvorbě řešení datové analytiky pak může ***řešení*** právě dostupnosti a kvality dat zabrat ***velké množství celkového času***. Stejně tak sem patří získání kvalitních trénovacích dat pro prediktivní modely, která nejsou ihned dostupná přímo v systémech. (Guo, 2021; SAP.com, 2025).

5.6.3 Správa a konektivita zařízení IoT

Základní problém může být ***nedostatečná technologická základna společnosti***, aby jejich informační systém byl schopný zajistit potřeby datové analytiky. Z toho mohou plynou další navazující náklady pro získání potřebného vybavení. Pak je potřeba zaměřit se na jejich dostupnost, zabezpečení a interoperabilitu. Musí být ***chráněna proti kybernetickým útokům***, aby nebyly prostředkem, který umožní útočníkům se k datům dostat a způsobit škodu společnosti. Musí být správně zaintegrované do systému a mělo dostatečnou konektivitu. Toto zajistí plynulý a nerušený přenos real time dat do všech systémů, kde jsou potřeba, a nebudou představovat bezpečnostní hrozbu. Sem pak patří i senzory a čidla, u kterých musíme zajistit jejich správné umístění a ošetření, jelikož mohou pracovat i v nepřívětivých prostředích, která je mohou poškozovat. (Guo, 2021; SAP.com, 2025).

5.6.4 Výběr algoritmu

Algoritmus prediktivní analytiky je zcela ***zásadní nástroj***. Často ale bývá problém vybrat ten, který by byl v daném případě nejlépe funkční. Je potřeba zvážit konkrétní ***případy použití a charakteristiky dat, která má zpracovávat***, jelikož na tomto základě lze pak určit ten, který je schopný je nejlépe využít. Některé algoritmy by pak například v konkrétním případě byly naprosto nevhodné pro daná data a výsledky, které by tento model přinesl, by byly nepřesné a nevyužitelné. Jasný určující faktor je také, že ***některé algoritmy daná data nemusí být schopné zpracovat*** nebo není schopný vytvořit ukazatel, který potřebujeme (Guo, 2021; SAP.com, 2025).

5.6.5 Odpor lidí a integrace do strategie

Zavedení datové analytiky a prediktivní údržby vyžaduje jejich podporu všemi zúčastněnými a všemi, koho se bude týkat. V návaznosti musí být pak integrována do stávající strategie údržby a pracovních postupů. Odolnost vůči změnám mohou snadno způsobit selhání implementace a veškerého snažení. Z toho by vyplývaly velké ztráty zbytečně vynaložených nákladů, zmaření mnoha hodin návrhů a tvorby a uniklé budoucí benefity z datové analytiky. Je tedy potřeba silné vedení, které bude jasně a pozitivně propagovat datovou analytiku a její benefity. (SAP.com, 2025).

5.7 Závěry



- Existují obecně mnoho ***různých faktorů***, které je potřeba zohlednit při návrhu údržby pro následnou implementaci, její přijetí a dlouhodobé řízení.
- Tento ***počet roste*** v závislosti na zvoleném typu údržby, přičemž největší je, pokud cílíme na proaktivní a prediktivní údržbu. Každý z nich může ovlivnit, jakým směrem musí být řešení vedeno.
- Existují také ***základní rizika***, která musíme zvážit při tvorbě a implementaci datové analytiky. Mohou existovat i ta, která budou platit jen u jedné konkrétní firmy.
- Je třeba vycházet také z toho, že ***prediktivní údržba*** vyžaduje určité ***pokročilé stádium stavu údržby*** a vyspělost a velmi obtížně by bylo možné ji vytvořit tam, kde je realizován pouze základní reaktivní typ.

6. Vstupní analýza řešení

Účelem je prezentovat ***vstupní analýzu řešení datové analytiky*** a ***faktory***, které budou již přímo navazovat na hlavní řešení deskriptivní a prediktivní analytiky. Faktory budou formovat, jak by daný ***návrh a postup dosažení měly vypadat***, jak by mělo být řešení přizpůsobeno pro společnost daného typu a jaké aspekty by finální výsledné řešení mělo obsahovat pro maximální efektivitu.

Součástí je ***katalog uživatelů***, které řešení může zasáhnout, ***cíle a přínosy*** pro údržbu a společnost či ***rozbor datových zdrojů***.

V této části budou užity zkušenosti z praxe. Tyto informace budou muset být zobecňovány s ohledem na požadavky firem a nastavenou situaci.

6.1 Cíle a přínosy datové analytiky

Oba typy datové analytiky mají mírně odlišné cíle a potenciál lehce odlišného přínosu. Z pohledu ***deskriptivního typu*** je cílem vytvořit ucelený pohled na řízení údržby a jeho každodenní fungování. Znamená to například přehled o nejdůležitějších KPI a metrikách pro linky a stroje, efektivitě a fungování oddělení údržby, efektivitě využití zdrojů, ekonomické stránce oddělení a strojů či aktuálním stavu aktiv.

Řešení, které bude spolehlivě v ***reálném čase konsolidovat a zpracovávat veškerá data*** může velmi omezit vznik chyby i zvýšit rychlost, v jakém jsou schopni informace získat. Stejně tak bude deskriptivní řešení přínosem pro ***management společnosti***, kdy budou schopni konstantně sledovat stav svých nejdůležitějších aktiv, tedy výrobních linek a jednotlivých zařízení. Také budou mít přehled o finanční a ekonomické stránce strojů, což jim poskytne mnohem větší kontrolu a rychlost ovlivnit své finance a náklady.

Prediktivní typ datové analytiky pak zajistí ve společnosti ***vznik prediktivní údržby***, a tedy provádět akce a zásahy kontrolovaně dříve, než jsou nutné a skutečně vznikne problém. Predikce by povýšily fungování údržby, a tedy vyšší efektivitu a spolehlivost linek a strojů s omezením celkových prostojů a omezením poruch. Jelikož jsou poruchy hlavním prvkem, který omezuje produkci společností, tak by se predikce měly zaměřit primárně na ně.

6.2 Katalog uživatelů

Pro role je zásadní definovat především jejich ***funkční náplň a hlavní požadované znalosti*** týkající se podnikové analytiky. Obecně sem patří především role manažerů, analytiků a vývojářů. Tyto role se ale mohou lišit a jsou více konkretizovány v rámci každé společnosti, jelikož závisí na jejích specifických podmínkách. Tyto role ***určují, jaké budou mít v rámci řešení různí zaměstnanci a pozice funkce a kompetence,*** které potřebují v rámci své specifické pozice a příslušných úloh, které řeší. V řešení pak mohou být pro tyto role přiřazeny různá práva a omezení a jejich přístup může být subjektivně upraven. Katalog rolí je pak vhodné v průběhu času adekvátně upravovat v závislosti na vývoji organizace a analytiky ve společnosti (o Novotný et al., 2025).

6.2.1 Role zaměstnanec managementu společnosti

Obecně je zásadní, aby měl management jako jeden z prvních k řešení přístup, jelikož vnímá údržbu jako součást strategie podniku a oddělení, které přímo ovlivňuje prosperitu společnosti. Pro názornější vysvětlení omezení rolí managementu by následně mohly existovat následující role.

- ***CEO / Generální manažer*** – jakožto hlavní manažer společnosti se zajímá o běh celého podniku a hlavního prostředku výdělku a běhu společnosti (výrobních linek). V rámci řešení by tak měl mít přístup k výsledkům spolehlivosti a efektivity linek stejně tak jako oddělení údržby, které se za jejich běh zaručuje. Stav a výsledky hlavních aktiv pak může pomoci utvářet cíle a směřování celého podniku.
- ***Výrobní manažer*** – manažerská role zabývající se plánováním a koordinací veškeré výroby závodu. Stejně jako CEO by tak měl mít přístup k výsledkům stavu a efektivity zařízení a údržby,

aby na jejich základě mohl uzpůsobovat plán a zajištění maximální výroby na základě těchto výsledků či predikcí.

- ***CFO / Manažer financí / Ekonom*** – jak již bylo zmíněno, tak by měl mít přístup především k finančním výsledkům strojů a údržby a případně aktuálnímu stavu aktiv pro kontrolu, plánování vynaložených nákladů, plánování rozpočtu či návratnost investic.
- ***Správce majetku*** – jelikož jsou linky hlavní součástí majetku společnosti, tak by i správce majetku měl mít přístup ke stavu linek a strojů, poruchám a predikcím. Na základě výsledků a spolupráce s údržbou pak může také podniknout patřičné akce pro zajištění co nejlepšího stavu strojů, řešit mimořádné situace, analyzovat majetek a plánovat jejich životní cyklus.
- ***CIO / Informační manažer*** – jakožto vedoucí IT by měl přístup do datově analytického řešení, které součástí IT vybavení společnosti. Může tak provádět celkovou kontrolu řešení, podílet se na jeho vývoji i řešit budoucí rozšíření a aktualizace řešení.

6.2.2 Role manažer údržby

Role manažera oddělení údržby je pravděpodobně ***nejdůležitější rolí celého řešení***. Právě on zodpovídá za všechny zaměstnance a celkový běh údržby. Z jeho podstaty je zásadní, aby měl ***neomezený přístup k celému řešení*** a zároveň měl neomezené rozhodovací právo, jak bude řešení koncipováno a vybudováno. Sám je také součástí managementu společnosti, ale z pohledu údržby je nejzásadnějším jedincem ve společnosti pro řešení. Na základě sledování KPI a metrik a získaných informací bude moci ***řídit zaměstnance, nastavovat plán údržby, korigovat zásahy údržby*** a s přidanou hodnotou predikcí i přímo ovlivňovat spolehlivost a efektivitu údržby a strojů.

6.2.3 Role zaměstnanec oddělení údržby

Zaměstnanci s touto rolí jsou ti, kteří ***provádějí zásahy na linkách*** a budou reagovat na výstupy získané z řešení. Je potřeba, aby údržbáři, bez ohledu na jejich specializaci, ***měli přístup do řešení*** a byly schopni buď s konzultací s manažerem anebo přímo sami z vlastní iniciativy provést zásah, pokud by se objevil jakýkoliv problém. Na základě jejich vstupu a návrhů může být řešení formováno a tvořeno do podoby, aby vyhovovalo i jejich potřebám.

6.3 Závěry



- Pokud nebude rozhodnuto jinak, tak především ***role managementu, manažera údržby a zaměstnanců údržby*** musí mít do řešení přístup.
- Pro ***další zúčastněné strany***, které přímo nesouvisí s údržbou, se budou práva a přístupy řešit individuálně až po zavedení.

6.4 Analýza faktorů řešení

Faktory řešení jsou aspekty prostředí společnosti, které je potřeba respektovat při řešení datové analytiky. Tyto aspekty mohou mít ***vliv v průběhu tvorby řešení, při implementaci i při užívání*** již zavedeného řešení. Tyto faktory se mohou dělit do ***různých kategorií***, jako jsou faktory vnějšího okolí a faktory vnitřního prostředí, byznys faktory, faktory technologického vybavení a mnoho dalších. Jejich zvážením a přizpůsobením řešení na jejich základě se snažíme adekvátně dosáhnou vytyčených cílů a dosáhnout očekávaných efektů. (o Novotný et al., 2025).

6.5 Analýza datových zdrojů

V této části dokumentu jsou uvedeny ***příklady hlavních tabulek*** a následně ***vygenerovaných zdrojových dat***. Z tabulek se budou vytvářet faktové tabulky a dimenze pro deskriptivní část datové analytiky a prediktivní dataset pro prediktivní část datové analytiky.

6.5.1 Výrobní data

První zdrojovou tabulkou jsou ***data z výroby***, tedy běhu výrobních linek a zařízení získané v reálném čase. Data jsou zachycována ***po hodinách***, což dává dostatečnou granularitu pro datovou analytiku. Pokud by byla po dnech nebo více, tak by granularita byla nedostatečná. Časové okno, které je

v datech je ***za 3 roky od roku 2022 do konce roku 2024***, což je dostatečná doba a množství dat, aby počet historických záznamů byl vhodný pro deskriptivní i prediktivní analytiku na naučení modelů. Data celá pak za tuto dobu obsahují **přes *315 000 záznamů***.

Data obsahují ***následující informace***:

- hodnoty ze senzorů získaných z linky, které zahrnují například teplotu, vibrace či tlak,
- hodnoty jsou po hodině s identifikací každého zařízení na každé lince,
- ukazatel, který zachycuje, zda zařízení bylo danou hodinu funkční,
- hodnoty vyprodukovaných kusů (včetně počtu celkem, NOK a OK kusů),
- hodnoty časové zachycující,
    - plánovanou produkci,
    - dobu od poruchy,
    - aktuální stáří zařízení.

6.5.2 Poruchy

Druhá je tabulka ***poruch***, která zachycuje ***všechny vzniklé poruchy za celé sledované období*** 3 let. Tyto poruchy jsou závislé na výrobních datech předchozího zdroje a každá vzniklá porucha musí vycházet a být popisována hodnotami naměřenými na lince v konkrétní časový okamžik. Výrobní data pak zachytí, že je zařízení v poruše a neběží. Celkem za 3 roky obsahuje ***přes 430 zaznamenaných poruch*** na všech zařízeních, což dává dobrý základ pro datovou analytiku fungování a spolehlivosti oddělení údržby i efektivity a spolehlivosti jednotlivých linek a zařízení.

Data obsahují:

- identifikaci linky a zařízení,
- časovou značku vzniku poruchy,
- typ poruchy, která vznikla,
- kritičnost poruchy,
- identifikaci zasahující osoby,
- časové hodnoty zachycující
    - čas reakce,
    - čas opravy,
    - downtime
    - konkrétní časy začátku a konce opravy, kdy zařízení opět fungovalo.

6.5.3 Údržba

Posledním zdrojem je tabulka ***celé údržby***, která zachycuje ***všechny zásahy údržby za sledované období*** *3 let*. Tabulka vychází z obou předchozích zdrojů a obsahuje jak ***reaktivní*** údržby, tedy reakce na náhlé poruchy předchozího zdroje, tak ***preventivní*** zásahy krátkodobé a dlouhodobé. Tabulka obsahuje přes ***1500 zásahů***, což opět dává dobrý základ pro datovou analytiku fungování a spolehlivosti údržby. Kombinací jednotlivých zdrojů získáme široký základ pro množství sledovatelných KPI a metrik pro popsání celé údržby. Mimo jiné jsou zde pak obsaženy i informace o použitých dílech na dané zásahy, které pomohou se sledováním stavu náhradních dílů a ekonomickou stránkou oddělení.

Tabulka obsahuje ***následující data:***

- identifikaci postižených linek a strojů,
- identifikaci typu zásahu,
- identifikaci konkrétní poruchy, pokud byl zásah reaktivní,
- časové údaje zahrnující,
    - start zásahu,
    - konec zásahu,
    - celkovou dobu zásahu,

- identifikaci zasahující osoby,
- personální náklady na zásah,
- informace o použitých náhradních dílech zahrnující,
    - identifikaci použitého dílu,
    - množství dílu,
    - náklady na použité díly.

6.5.4 Zdroje pro dimenzionální tabulky

Dříve popsané zdroje i jiné evidence v rámci společnosti pak budou poskytovat dostatečný základ pro dimenzionální tabulky potřebné pro deskriptivní analytiku. Díky nim bude možné provádět potřebné náhledy a filtrace nad fakty ve faktových tabulkách vycházejících opět z předešlých popsaných zdrojů. Mezi tyto dimenze budou patřit například následující.

- Linky
- Zařízení
- Čas
- Technici
- Typ poruchy
- Typ zásahu údržby a další

6.6 Závěry



- Příklady představují ***relativně rozsáhlou základnu*** velkého množství dat s dobrou granularitou, která jsou ***dostatečná*** pro provádění jak deskriptivní, tak prediktivní analytiky.
- Množství dat je ***dostatečné na naučení prediktivních modelů*** a lze z nich vytvořit velký prediktivní dataset.
- Tyto zdroje společně popisují ***na sebe navazující události***, kterými jsou průběh výroby jednotlivých linek a strojů, na nich vzniklé poruchy a na ně navazující reaktivní údržby společně s pravidelnými preventivními zásahy.
- Zdroje obsahují ***různorodé typy dat***, která nám dávají základ pro sledování většího množství KPI a metrik důležitých pro údržbu.
- Data současně pokrývají ***dostatečně dlouhé období*** pro popsání delšího historického života údržby a strojů a dostatečný základ pro naučení prediktivních modelů.

7. Návrh a tvorba řešení – deskriptivní úlohy

Účelem je charakterizovat návrh a tvorbu samotného datově analytického řešení. Aktuální část se bude zabývat ***deskriptivní částí,*** zahrnující definice jednotlivých KPI a metrik, definice a popis jednotlivých dimenzionálních a faktových tabulek, na jejich základě tvorba multidimenzionálního modelu a vyhodnocení vzniklých reportů či dashboardů.

Tvorba bude založena na předchozích kapitolách. K tvorbě reportů se využije aplikace Power BI.

7.1 Definice KPI a metrik

KPI neboli *Key Performance Indicator* představuje ***hlavní ukazatel, dle kterého se měří výkon*** a úspěšnost společnosti. Závisí na subjektu, který se snažíme popisovat. Je zásadní, aby tyto KPI byla sledovatelná z pohledu různých dimenzi analytického řešení, aby nepředstavovalo jen statické číslo. Každé KPI musí být ***jednoznačně definováno v rámci celého analytického řešení*** a mělo by být sledováno ve stejných jednotkách. Pokud chceme mít více KPI jednoho typu, tak je potřeba, aby bylo každé ***separátně definováno, navzájem odlišeno*** a jednoznačně odlišně pojmenováno (Kimball & Ross, 2013).

V následujících tabulkách jsou uvedeny hlavní charakteristiky jednotlivých metrik

Table 7-1: MTBF

ID	MTBF
Název	MTBF – Mean Time Between Failures
Popis	Průměrný čas mezi jednotlivými poruchami
Zdroj	F_EquipmentOperation, F_Failure
Výpočet	Celkový čas funkčního provozu / Celkový počet poruch
Jednotka	Čas - Hodiny (h)
KPI	ANO

Table 7-2: Celkový počet poruch

ID	celkovy_pocet_poruch
Název	Celkový počet poruch
Popis	Celkový počet poruch za sledované období
Zdroj	F_Failure
Výpočet	Počet záznamů v F_Failure
Jednotka	Kusy (ks)
KPI	ANO

Table 7-3: MTTR

ID	MTTR
Název	MTTR – Mean Time To Repair
Popis	Průměrný čas od vzniku poruchy do opravy zařízení
Zdroj	F_Failure
Výpočet	Průměr downtime na poruchu
Jednotka	Čas – Minuty (min)

KPI	ANO

Table 7-4: Celkový downtime

ID	celkovy_downtime
Název	Celkový downtime
Popis	Celkový čas prostojů (downtime) za sledované období
Zdroj	F_Failure
Výpočet	Suma celkových downtime na poruchy v F_Failure
Jednotka	Čas – Hodiny (h)
KPI	ANO

Table 7-5: : Equipment downtime

ID	equipment_downtime
Název	Equipment downtime
Popis	Procentuální zastoupení downtime vůči celkovému času možné výroby
Zdroj	F_EquipmentOperation, F_Failure
Výpočet	Celkový downtime / Celkový možný čas výroby
Jednotka	Procenta (%)
KPI	ANO

Table 7-6: Availability (dostupnost)

ID	availability
Název	Dostupnost (Availability)
Popis	Procentuální vyjádření, jak byl stroj dostupný vůči času plánu výroby
Zdroj	F_EquipmentOperation
Výpočet	Celkový funkční čas / Celkový plánovaný čas výroby
Jednotka	Procenta (%)
KPI	ANO

Table 7-7: Quality (kvalita)

ID	quality
Název	Kvalita (Quality)
Popis	Procentuální vyjádření OK kusů vůči celkem vyrobených kusů
Zdroj	F_EquipmentOperation
Výpočet	Celkový počet OK kusů / Celkový počet vyrobených kusů
Jednotka	Procenta (%)
KPI	ANO

Table 7-8: Quality (kvalita)

ID	performance

Název	Celkový výkon (Performance)
Popis	Procentuální vyjádření celkem vyrobených kusů vůči teoretickému maximálnímu možnému počtu vyrobených kusů
Zdroj	F_EquipmentOperation, D_Machine
Výpočet	Celkový počet vyrobených kusů / Teoretický maximální počet vyrobených kusů
Jednotka	Procenta (%)
KPI	ANO

Table 7-9: OEE – Overall Equipment Effectiveness

ID	OEE
Název	OEE - Overall Equipment Effectiveness
Popis	Efektivita celé výroby vůči svému maximálnímu potenciálu
Zdroj	F_EquipmentOperation, D_Machine
Výpočet	Dostupnost (Availability) * Kvalita (Quality) * Výkon (Performance)
Jednotka	Procenta (%)
KPI	ANO

Table 7-10: Celkem vyrobeno kusů

ID	celkem_vyrobeno_kusu
Název	Celkem vyrobeno kusů
Popis	Celkový počet vyrobených kusů za sledované období
Zdroj	F_EquipmentOperation
Výpočet	Suma celkem vyrobených kusů v F_EquipmentOperation
Jednotka	Kusy (ks)
KPI	NE

Table 7-11: Teoretická maximální výroba

ID	teoreticka_maximalni_vyroba
Název	Teoretická maximální výroba
Popis	Maximální teoretický počet vyrobených kusů za celý plánovaný čas výroby
Zdroj	F_EquipmentOperation, D_Machine
Výpočet	Celkový počet hodin plánu výroby * Maximální hodinový výstup stroje
Jednotka	Kusy (ks)
KPI	NE

Table 7-12: Celkový počet zásahů

ID	celkovy_pocet_zasahu
Název	Celkový počet zásahů
Popis	Celkový počet zásahů provedených oddělením údržby

Zdroj	F_Maintenance
Výpočet	Počet záznamů v F_Maintenance
Jednotka	Kusy (ks)
KPI	NE

Table 7-13: Celkový počet preventivních zásahů

ID	celkem_zasahu_preventivni
Název	Preventivní zásahy
Popis	Celkový počet preventivních zásahů provedených oddělením údržby
Zdroj	F_Maintenance
Výpočet	Počet preventivních záznamů v F_Maintenance
Jednotka	Kusy (ks)
KPI	NE

Table 7-14: Celkový počet reaktivních zásahů

ID	celkem_zasahu_reaktivni
Název	Reaktivní zásahy
Popis	Celkový počet reaktivních zásahů provedených oddělením údržby
Zdroj	F_Maintenance
Výpočet	Počet reaktivních záznamů v F_Maintenance
Jednotka	Kusy (ks)
KPI	NE

Table 7-15: Průměrná délka zásahu

ID	prumerna_delka_zasahu
Název	Průměrná délka zásahu
Popis	Průměrná doba trvání všech zásahů za sledované období
Zdroj	F_Maintenance
Výpočet	Průměr délky trvání v F_Maintenance
Jednotka	Čas – Minuty (min)
KPI	NE

Table 7-16: Průměrná doba reakce na poruchu

ID	prumerna_doba_reakce
Název	Průměrná doba reakce (poruchy)
Popis	Průměrná doba reakce na poruchy
Zdroj	F_Failure
Výpočet	Průměr doby reakce v F_Failure
Jednotka	Čas – Minuty (min)
KPI	ANO

Table 7-17: Ztráta z prostojů

ID	ztrata_z_prostoju
Název	Ztráta z prostojů
Popis	Celková Ztráta vzniklá všemi prostojů.
Zdroj	F_Failure, D_Machine
Výpočet	Celkový Downtime na zařízení * Hodinová ztráta zařízení
Jednotka	Kč
KPI	ANO

Table 7-18: Celkové náklady údržby

ID	celkove_naklady_udrzba
Název	Celkové nakłady údržby
Popis	Celkové náklady oddělení údržby
Zdroj	F_Maintenance
Výpočet	Suma nákladů na jednotlivé zásahy v F_Maintenance
Jednotka	Kč
KPI	ANO

Table 7-19: Celkový finanční dopad

ID	celkovy_dopad
Název	Celkový finanční dopad
Popis	Celkový finanční dopad vzniklý z nákladů údržby a ztrát.
Zdroj	F_Maintenance, F_Failure, D_Machine
Výpočet	Celkové náklady údržby + Ztráta z prostojů
Jednotka	Kč
KPI	ANO

Table 7-20: RAV (Remaining Asset Value)

ID	RAV
Název	RAV (Remaining Asset Value)
Popis	Zbývající hodnota aktiva vůči jeho původní hodnotě.
Zdroj	F_EquipmentOperation, D_Machine
Výpočet	1 - (Aktuální stáří zařízení / Očekávaná životnost zařízení)
Jednotka	Procenta (%)
KPI	ANO

7.2 Dimenze

Dimenze obsahují ***kontext k datům ve faktových tabulkách***, které obsahují reálná data třeba právě výroby. Vysvětlují otázky typu “kdo, co, kde, kdy, jak a proč”, která dokážou přesně popsat vzniklá data nějaké události (například vznik poruchy) a filtrovat je. Obsahují často ***přesné textové údaje*** (název, popis, cena, typ apod.) narozdíl od faktových tabulek, které obsahují cizí klíče k dimenzím

(primární klíč dimenze), přes které se tyto údaje váží k události. Událost poruchy tedy bude mít cizími klíči navázané například dimenze stroje či času, což odpoví právě na otázky “kdy a kde”. Takovýmto dimenzionálním modelováním je pak ***možné dynamicky filtrovat*** celé reporty přes jednotlivé dimenze – například poruchy na stroji 5 v měsíci červen a k nim příslušné metriky (Kimball & Ross, 2013).

Další tabulky, obdobně, jako v případě metrik zachycují ***základní charakteristiky významných dimenzí*** pro řízení údržby.

Table 7-21: Čas

ID	D_Time
Název	Čas
Popis	Podrobné záznamy času umožňující sledování ukazatelů v čase i na směny
Typ	Časová
Struktura	Datum – Rok – Čtvrtletí – Měsíc – Týden – Den – Hodina - Směna
Atributy	timestamp, date_key, date, year, quarter, month, month_name, week, day, day_of_week, day_of_week_number, hour, is_weekend, shift_id, shift_name, shift_start_hour, shift_end_hour

Table 7-22: Stroje

ID	D_Machine
Název	Stroje
Popis	Stroje a jejich příslušné údaje
Typ	Snowflake
Struktura	Linky – Stroje, Typ stroje – Stroje, Kategorie stáří stroje – Stroje
Atributy	machine_id, line_id, machine_type_id, manufacturer, installation_date, acquisition_cost, depreciation_rate, max_units_per_hour, planned_hours_per_day, hourly_production_loss, expected_lifetime_hours, machine_age_years, asset_age_id

Table 7-23: Typ Stroje

ID	D_MachineType
Název	Typ stroje
Popis	Typy strojů, které společnost vlastní a jsou na linkách
Typ	Snowflake
Struktura	Typ stroje – Stroje
Atributy	machine_type_id, machine_type_name, description

Table 7-24: Linky

ID	D_Line
Název	Linky
Popis	Linky, na kterých jsou jednotlivá zařízení.
Typ	Snowflake
Struktura	Linky – Stroje
Atributy	line_id, line_name, plant, description

Table 7-25: Kategorie stáří zařízení

ID	D_AssetAge
Název	Kategorie stáří zařízení
Popis	Pomáhá rozdělovat a kategorizovat stroje do skupin dle jejich stáří.
Typ	Snowflake
Struktura	Kategorie stáří stroje – Stroje
Atributy	asset_age_id, age_bucket_name, age_min_years, age_max_years

Table 7-26: Typ poruchy

ID	D_FailureType
Název	Typ poruchy
Popis	Typy poruch, které mohou vzniknout.
Typ	Star
Atributy	Failure_type_id, failure_type, description

Table 7-27: Typ údržby

ID	D_MaintenanceType
Název	Typ údržby
Popis	Typ zásahu údržby, který byl proveden.
Typ	Star
Atributy	maintenance_type_id, maintenance_type

Table 7-28: Technici

ID	D_Technician
Název	Technici
Popis	Zaměstnanci oddělení údržby
Typ	Star
Atributy	technician_id, technician_name, specialization, seniority_level, hourly_rate

Table 7-29: Kritičnost

ID	D_Severity
Název	Kritičnost
Popis	Typy kritičnosti poruchy
Typ	Star
Atributy	severity_id, severity_level, severity_description

Table 7-30: Náhradní díly

ID	D_SparePart
Název	Náhradní díly
Popis	Náhradní díly, které jsou k dispozici.

Typ	Star
Atributy	part_id, part_name, cost_per_unit, minimum_stock, current_stock

7.3 Znázornění vazeb mezi dimenzemi a KPI

Vztahy mezi jednotlivými KPI a dimenzemi ***určí, k jakým KPI jsou navázána data jednotlivých dimenzí a jaké následně zajistí filtraci a agregaci dat*** pro daná KPI. Zde jsou ***jen ukazatele definované jako KPI*** a nikoli všechny metriky, které byly popsány či které pak vzniknou ve výsledných reportech.

Table 7-31: Vztah KPI a dimenzí

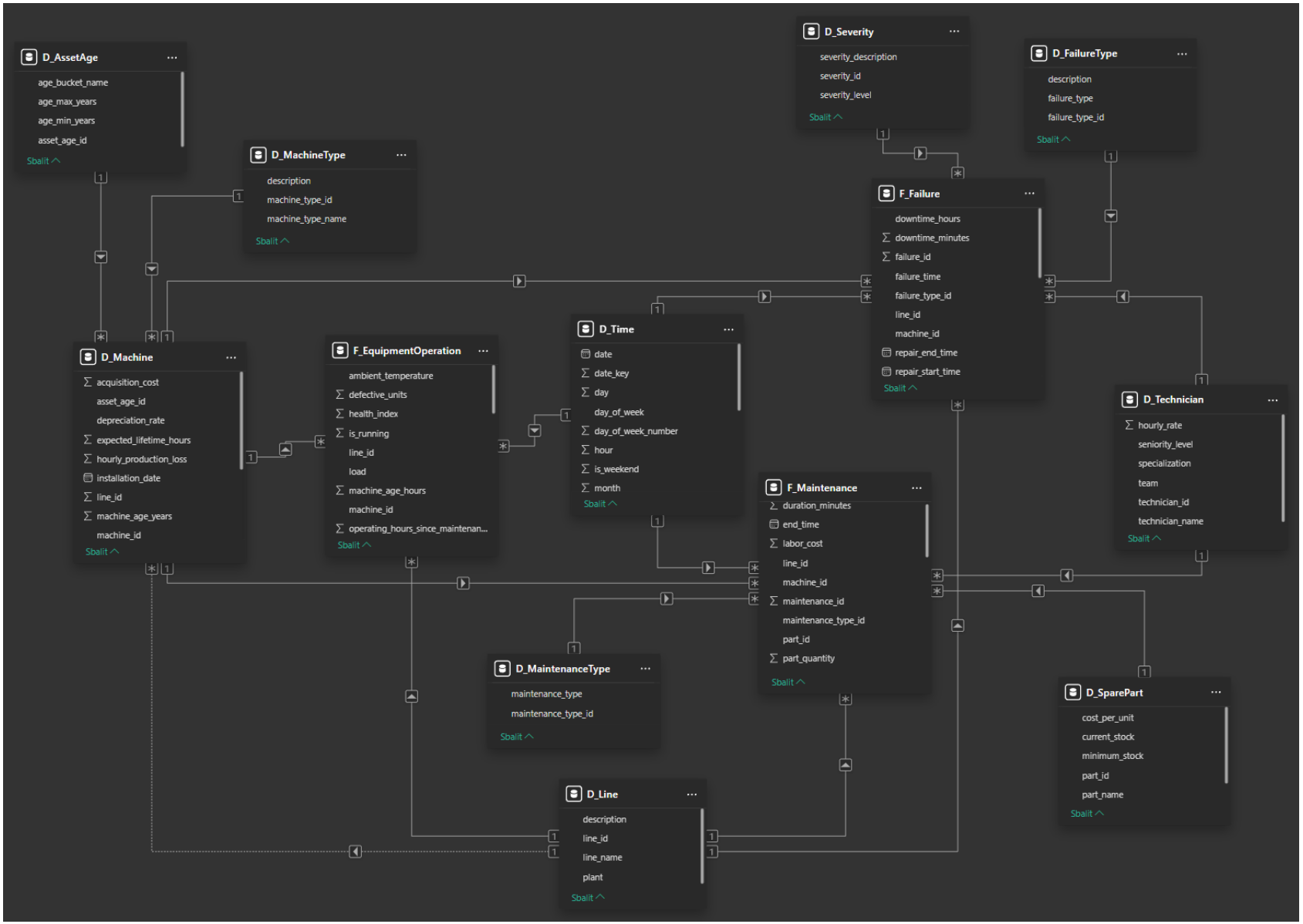
Dimenze → / KPI ↓	Čas	Stroje	Typ stroje	Linky	Kategorie stáří stroje	Typ poruchy	Typ údržby	Technici	Kritičnost	Náhradní díly
MTBF	X	X	X	X	X	X			X	
Celkový počet poruch	X	X	X	X	X	X		X	X	
MTTR	X	X	X	X	X	X		X	X	
Celkový downtime	X	X	X	X	X	X			X	
Equipment downtime	X	X	X	X	X	X			X	
Dostupnost (Availability)	X	X	X	X	X					
Kvalita (Quality)	X	X	X	X	X					
Výkon (Performance)	X	X	X	X	X					
OEE	X	X	X	X	X					
Průměrná doba reakce	X	X	X	X				X		
Ztráta z prostojů	X	X	X	X	X					
Celkové nakłady údržby	X	X	X	X	X		X			X
Celkový finanční dopad	X	X	X	X	X		X			X
RAV (Remaining Asset Value)	X	X	X	X	X					

7.4 Dimenzionální model

Faktová tabulka ***F_EquipmentOperation*** vznikla ***ze zdrojové tabulky výrobních dat***. Obsahuje veškeré informace ***s hodinovou granularitou***, jak probíhala výroba za sledované období. Mimo cizích klíčů dimenzí obsahuje informace o stavu strojů, zda je funkční, produkovaných kusech, době operování celkem a od poruchy a další.

Faktová tabulka ***F_Failure*** vznikla ***ze zdrojové tabulky poruch***. ***Granularita je v rovině událostí na základě každé poruchy***. Obsahuje informace o dobách a času poruch, časech reakcí, technicích či kritičnosti poruchy. Zachycuje i informace o tom, jaká část z downtime byla reakce a jaká byla reálně doba celé opravy.

Faktová tabulka ***F_Maintenance*** vznikla ***ze zdrojové tabulky údržby***. V ní jsou zachycené veškeré zásahy údržby zahrnující reaktivní vznikající z náhlých poruch i preventivní, které jsou prováděny pravidelně na základě nastaveného plánu údržby. ***Granularita je na úrovni událost každého jednoho zásahu.***


D_AssetAge
age_bucket_name
age_max_years
age_min_years
asset_age_id
Sbalit
D_MachineType
description
machine_type_id
machine_type_name
Sbalit
D_Severity
severity_description
severity_id
severity_level
Sbalit
D_FailureType
description
failure_type
failure_type_id
Sbalit
F_Failure
downtime_hours
downtime_minutes
failure_id
failure_time
failure_type_id
line_id
machine_id
repair_end_time
repair_start_time
Sbalit
D_Machine
acquisition_cost
asset_age_id
depreciation_rate
expected_lifetime_hours
hourly_production_loss
installation_date
line_id
machine_age_years
machine_id
Sbalit
F_EquipmentOperation
ambient_temperature
defective_units
health_index
is_running
line_id
load
machine_age_hours
machine_id
operating_hours_since_maintenan...
Sbalit
D_Time
date
date_key
day
day_of_week
day_of_week_number
hour
is_weekend
month
Sbalit
D_Technician
hourly_rate
seniority_level
specialization
team
technician_id
technician_name
Sbalit
F_Maintenance
duration_minutes
end_time
labor_cost
line_id
machine_id
maintenance_id
maintenance_type_id
part_id
part_quantity
Sbalit
D_MaintenanceType
maintenance_type
maintenance_type_id
Sbalit
D_SparePart
cost_per_unit
current_stock
minimum_stock
part_id
part_name
Sbalit
D_Line
description
line_id
line_name
plant
Sbalit

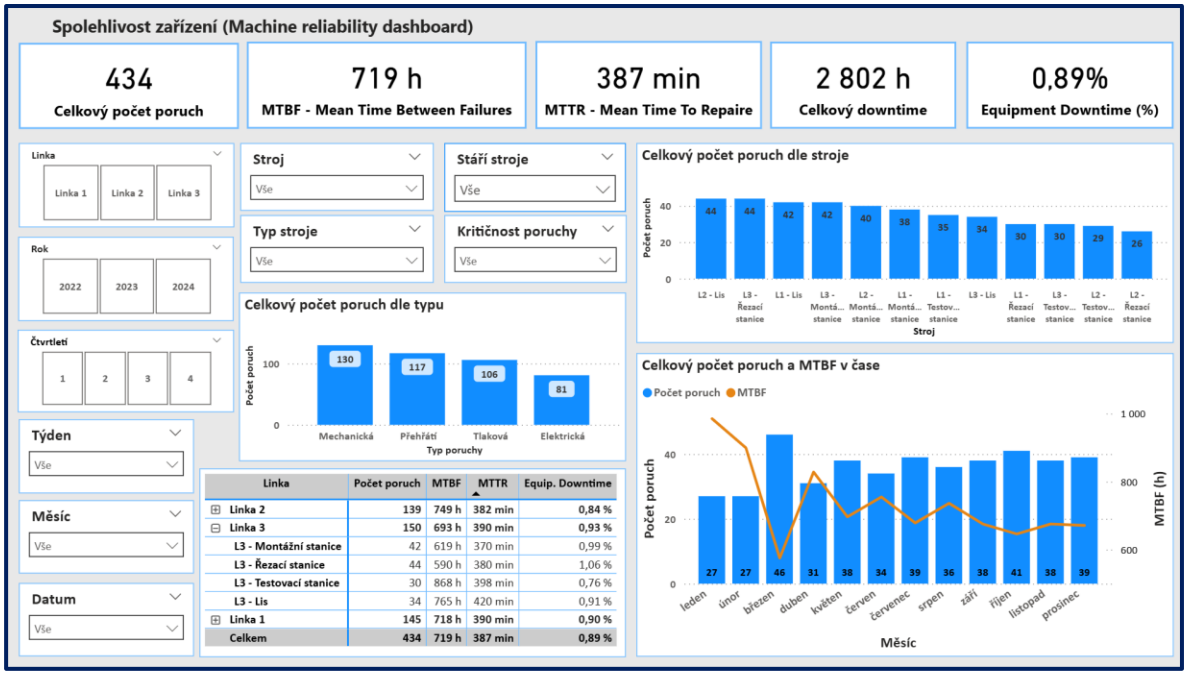
7.5 Reporty deskriptivní části

Řešením je ***5 analytických reportů*** v aplikaci Power BI. Jednotlivé reporty jsou rozděleny dle témat a zobrazují důležitá KPI, která jsou zásadní pro celou společnost, ale především pro oddělení údržby.

Reporty jsou rozděleny následovně.

- Spolehlivost strojů.
- Výkonnost strojů.
- Efektivita oddělení údržby.
- Ekonomika údržby.
- Stav aktiv.

7.5.1 Report Spolehlivost



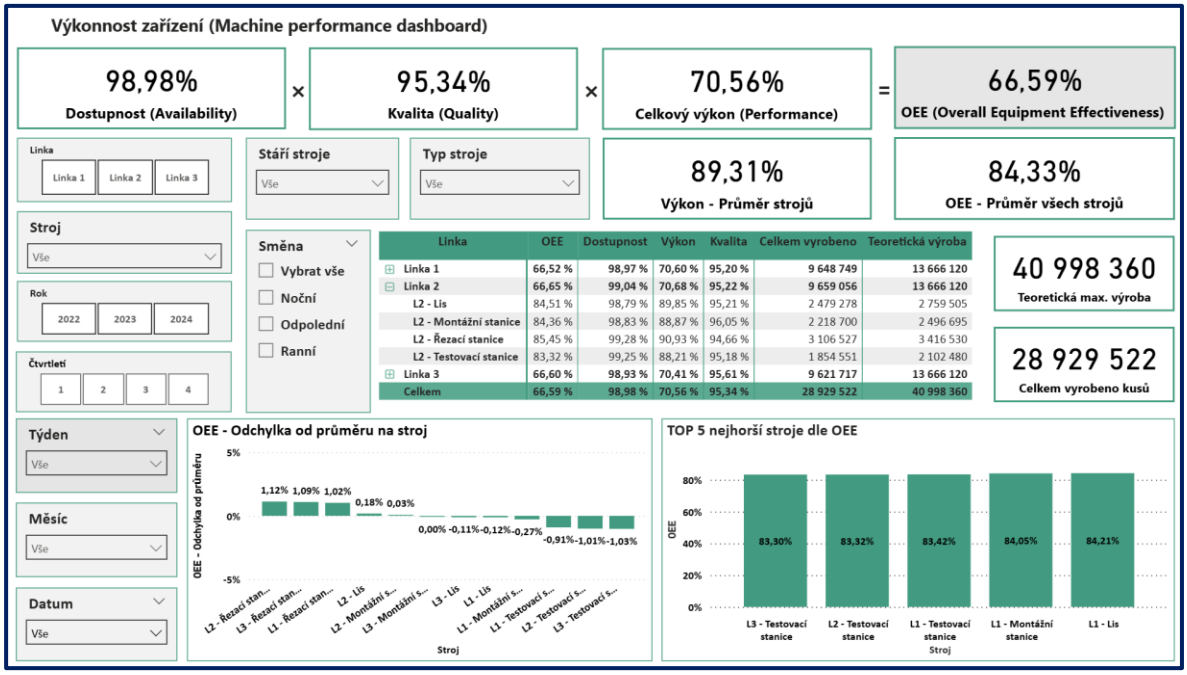
Obrázek 7-1: Report Spolehlivost (zdroj: autor)

Report ***Spolehlivost*** (Obrázek 7-1) zachycuje spolehlivost strojů. Jeho dominantou jsou hlavní KPI v horní části. Zachycují celkový ***počet vzniklých poruch***, ***MTBF – Mean Time Between Failures*** (průměrný čas mezi poruchami), ***MTTR – Mean Time to Repair*** (průměrný čas do opravy), ***celkovou hodnotu downtime*** za celé období a ***Equipment Downtime*** (procento downtime z celkového času výroby).

Graf celkový počet poruch dle stroje říká, jaké stroje jsou nejproblémovější a naopak. Další je ***celkový počet poruch a MTBF v čase***, kde lze sledovat, jak se v čase vyvíjelo množství poruch a jak se v návaznosti na ně vyvíjelo MTBF. Větší množství poruch znamená klesající MTBF, tedy horšící se spolehlivost, což je na grafu viditelné. Poslední graf přehledně znázorňuje ***rozdělení všech poruch dle jejich typu***, kdy v našem případě je největší množství mechanických poruch. Ve spodní části shrnující je matice, která zobrazuje všechny hodnoty přehledně na jednotlivé linky a stroje. V základu matice zobrazuje celkově linky a symbolem plus u nich pak přejdeme e do většího detailu k jí příslušným strojům. Z ní můžeme jednoduše porovnávat jednotlivé linky a stroje vůči sobě a přímo vidět korespondující hodnoty jednotlivých zařízení.

Všechny ukazatele jsou filtrovatelné pomocí časových úseků (rok, měsíc, datum atd.), pomocí strojů a linek, typů strojů, stáří stroje (starší stroje mají například kratší MTBF) a kritičnosti samotných poruch.

7.5.2 Report Výkonnost



Obrázek 7-2: Report Výkonnost (zdroj: autor)

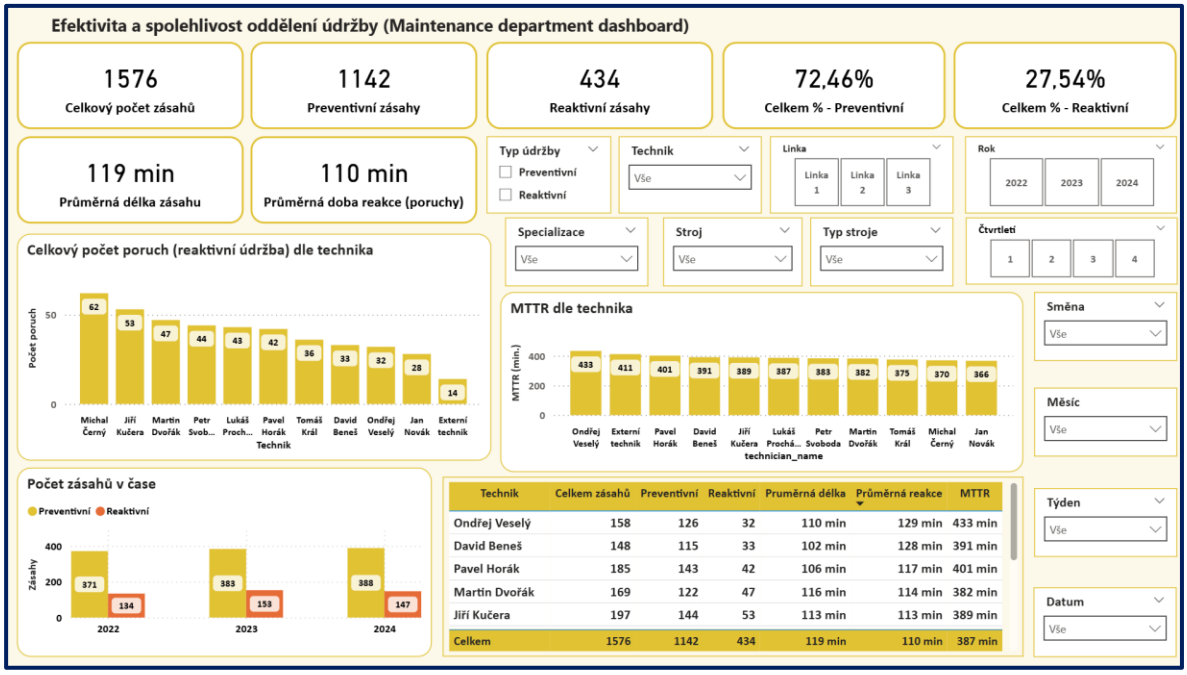
Report ***Výkonnost*** (Obrázek 7-2) zobrazuje výkonnost strojů. Jeho dominantou jsou opět hlavní KPI, která zachycují na sebe navazující ***Dostupnost***, ***Kvalitu*** a ***Výkon***. Společným součinem těchto KPI získáme hlavní KPI ***OEE – Overall Equipment Effectiveness*** (celková efektivity výroby). Tato KPI jsou zásadní, jelikož vyjadřují efektivitu celkové výroby a počítají se z různých aspektů výroby.

Dostupnost se získává z celkových časů výroby a plánu. ***Kvalita*** se získává z celkem vyrobených a OK kusů. ***Výkon*** se získává z celkových vyrobených kusů a teoretické maximálně možné výroby kusů. Každé zařízení má ***jiný počet maximální hodinové výroby***, a tedy ne každé zařízení má stejnou váhu dle počtu výroby kusů.

Například je zařízení s maximálním počtem kusů za hodinu 90 a jiné s počtem 130. Pro jedno je výstup 90 dokonalý a pro druhé slabý. Pokud zprůměrujeme veškeré stroje s jejich potenciálním a reálným výstupem, tak jsou hodnoty procent nižší (70 %). Pokud ale spočítáme každé zařízení zvlášť samo o sobě (procenta bude mít vyšší) a poté uděláme jednoduchý průměr jejich jednotlivých výkonů, procenta jsou značně vyšší (90 %). To znamená, že každé zařízení má dobrý výkon kolem 90 %, ale celková výroba je slabší a mohla by být více optimalizovaná. Ukazatel výkonu pak ovlivňuje i OEE, pro které platí stejné vlivy jako pro výkon.

Report obsahuje i ***celkový počet vyprodukovaných kusů*** *a* ***teoretickou maximální možnou výrobu***. Ve spodní části jsou grafy, kdy levý znázorňuje ***odchylku strojů od průměrného OEE*****,** takže můžeme sledovat, které stroje jsou lepší, které horší a jaké jsou rozdíly v jejich efektivitě. Pravý graf pak zřetelně znázorňuje ***TOP 5 nejhorších strojů dle OEE***, které společnost vlastní. Nakonec je uprostřed tabulka, která popisuje detailně jednotlivá KPI linek a strojů a můžeme je navzájem přehledně porovnávat. KPI můžeme opět filtrovat pomocí času (roky, měsíce, týdny atd.), linek, strojů, typů strojů, stáří stroje (starší stroje mají například horší OEE) a zde i pomocí typu směny.

7.5.3 Report Efektivita oddělení údržby



Obrázek 7-3: Report Efektivita oddělení údržby (zdroj: autor)

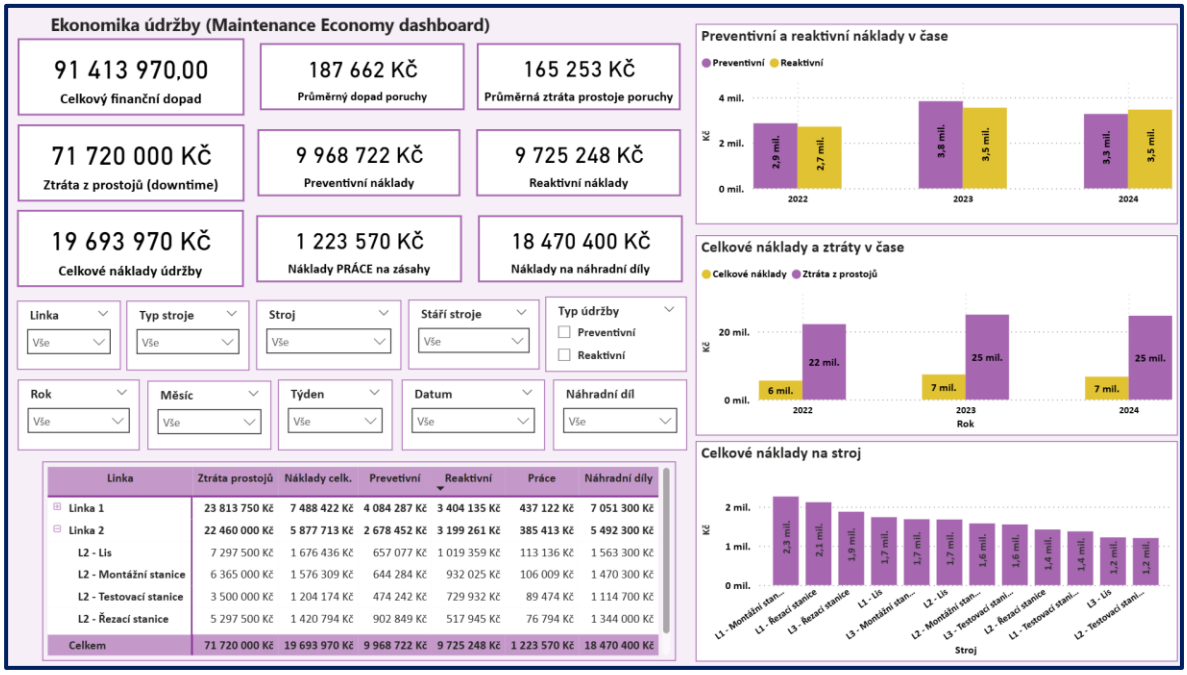
Report ***Efektivita oddělení údržby*** (Obrázek 7-3) se zaměřuje již na samotné oddělení, a nikoliv přímo na linky a stroje. Hlavní součástí jsou metriky, které popisují práci techniků, zásahy oddělení a další časové údaje. Hlavními ukazateli jsou ***celkový počet zásahů***, ***celkový počet preventivních a reaktivních zásahů*** (reakce na poruchy), ***procentuální poměr těchto typů zásahů***, ***průměrná délka zásahu*** a ***průměrná reakční doba*** (ta se vztahuje jen na poruchy, jelikož preventivní zásahy reakci nemají).

Například z celkového počtu téměř 1600 zásahů je zde vidět, že naprostá většina (72,5 %) je preventivních, což je správné, jelikož ve společnostech s preventivní údržbou by měl převažovat počet preventivních zásahů. To pak znamená, že preventivní zásahy mají svůj vliv a snižují počet nečekaných poruch. Pokud by bylo více reaktivních či by poměry byly vyvážené, tak by bylo jasné, že preventivní údržba není nastavena správně nebo není prováděna správně. Průměrná délka zásahu je počítána z preventivních i reaktivních a metriku lze filtrovat. Průměrná doba reakce je pak jen z poruch a říká, jak rychle dokážou průměrně zaměstnanci oddělení údržby zareagovat na poruchu.

Report obsahuje grafy, kdy první zobrazuje ***celkový počet poruch na zaměstnance***, což popisuje podíl jednotlivých zaměstnanců na řešení poruch a jejich pracovní vytížení. Stejně tak podobné informace poskytuje graf zobrazující ***MTTR dle technika*** (průměrná doba do opravy na technika), což poskytne informace, kdo dokáže provádět opravy efektivně a rychle. Graf může tak poskytovat informaci, kdo by se měl například zlepšit. Poslední graf popisuje ***počet preventivních a reaktivních zásahů v čase***, což může podávat informaci, jak se vyvíjí stav strojů (pokud by degradací časem rostly reaktivní zásahy) či zda oddělení pracuje správně.

Poslední vizuál je shrnující tabulka, kde můžeme sledovat všechny ukazatele na zaměstnance s konkrétními čísly a sledovat výkon jednotlivců. Report lze opět filtrovat v čase, na základě linek, strojů, typů stroje, ale i dle směny, technika, jeho specializace a typu zásahu. Součástí je i pro potřeby práce ***zastupující externí technik***, který reprezentuje všechny externí zásahy, které jsou součástí kombinované organizační struktury údržby.

7.5.4 Report Ekonomika údržby



Obrázek 7-4: Report Ekonomika údržby (zdroj: autor)

Report ***Ekonomika údržby*** (Obrázek 7-4) se zaměřuje na ekonomickou, finanční a nákladovou stránku celé údržby, což zahrnuje jak stroje, tak oddělení údržby samotné. K tomu příslušní několik hlavních KPI, kterými jsou ***Celkový finanční dopad*** (náklady na údržbu a ztráty z prostojů), ***Ztráta z prostojů*** za celé sledované období a ***Celkové náklady na údržbu***. Od těchto se pak odvíjí ukazatele ***Průměrný dopad poruchy*** (průměr ztráty a nákladů), ***Průměrná ztráta poruchy*** (z prostojů), ***Preventivní náklady*** *a* ***Reaktivní náklady*** (společně tvoří celkové náklady), a nakonec kombinace ***Náklady práce na zásahy*** a ***Náklady na náhradní díly***.

Pro náklady na práci platí, že ***se zde počítá jen konkrétní náklad na odvedenou práci*** pro daný zásah, tedy kombinace délky zásahu/opravy a hodinové sazby zaměstnance. Není zde tedy řešena personální stránka zaměstnanců ve smyslu jejich celkových mezd. To způsobuje, že náklady na práci jsou v reportu markantně nižší než náklady na náhradní díly. Jelikož ale hodnota závisí na délce zásahu, tak dle ní můžeme hodnotit i efektivitu techniků (nižší náklady mohou znamenat méně poruch či rychlost a efektivitu jejich práce).

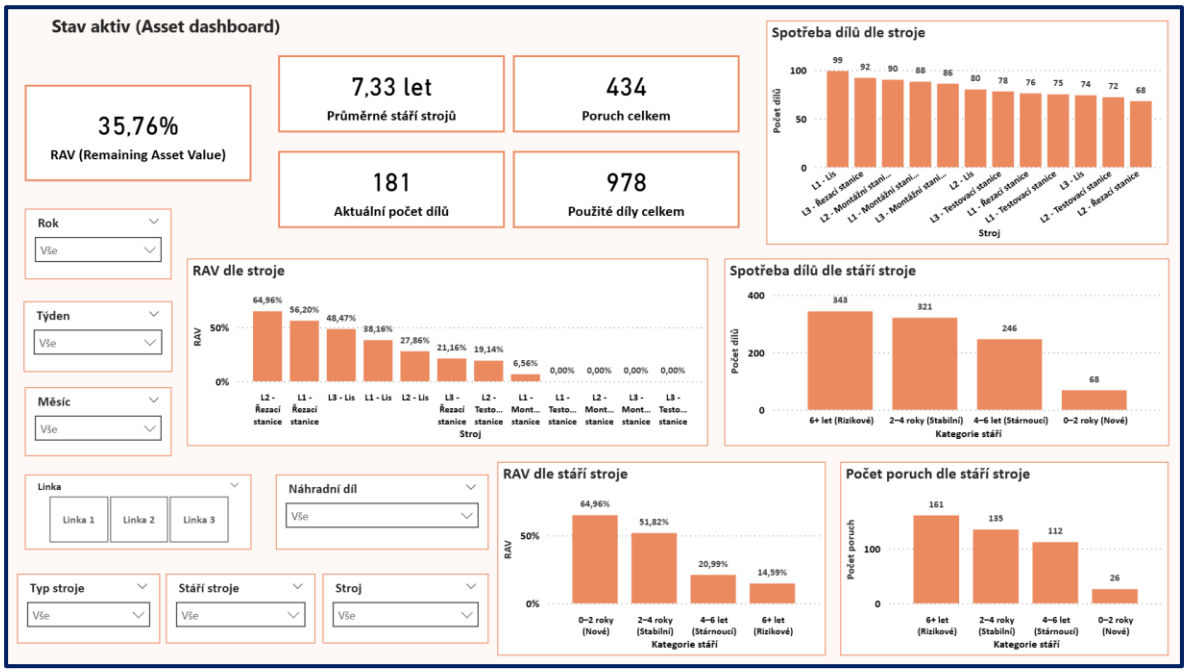
Ztráty z prostojů jsou počítány na základě celkového downtime a nastavené hodinové ztráty každého stroje, pokud stroj nefunguje. Rozdíl ***průměrného dopadu*** *a* ***průměrné ztráty*** pak vychází jako průměrné náklady na poruchu a vývojem těchto hodnoty pak můžeme sledovat, zda se nám zvyšuje ztráta (vyšší downtime, méně efektivní údržba) nebo se nám naopak zvyšují náklady, ale ztráty jsou jinak podobné a podle toho provést potřebná opatření. Také můžeme sledovat, že i přesto, že (na základě předchozího reportu) je počet reaktivních zásahů třetinový, tak generuje podobné náklady jako všechny preventivní zásahy. To vyvolává potřebu eliminovat poruchy, zefektivňovat preventivní zásahy a snažit se eliminovat reaktivní náklady, což může mít velký pozitivní vliv.

Graf ***rozdělení nákladů na preventivní a reaktivní zásahy v čase*** umožňuje sledovat vývoj nákladů, porovnávat je a dělat závěry, zda je využití nákladů efektivní a dává smysl a zda údržba funguje správě či nikoliv (postupné zvyšování reaktivních může znamenat buď špatnou funkci údržby anebo naopak postupnou degradaci strojů a více poruch). Další popisuje porovnání ***celkových nákladů a celkové ztráty v čase***. Sleduje se tak, zda se ztráta a náklady zvyšuje či snižuje anebo zda více nákladů na údržbu sníží ztrátu či nikoliv. Poslední graf ***znázorňuje celkové náklady na stroj***, kde můžeme sledovat jaký stroj generuje nejvíce nákladů a které jsou v tomto ohledu naopak nenáročné.

Poslední vizuál je ***tabulka***, kde je možné ztráty a všechny ***náklady sledovat přehledně*** na jednotlivé linky a stroje a sledovat, jaký stroj je náročnější, jaký má vyšší či nižší finanční nároky a jak je rozložená ztráta. Report je pak možné opět filtrovat v čase, dle linek, strojů, typů stroje a stáří stroje, na

základě typů údržby, a nakonec i na konkrétní náhradní díly, kdy vidíme, kde byly použity, kdy a s jakými náklady.

7.5.5 Report Stav aktiv



Obrázek 7-5: Report Stav aktiv (zdroj: autor)

Report ***Stav Aktiv*** (Obrázek 7-5) se zaměřuje jak na aktuální stav vlastněných strojů, tak i na využití náhradních dílů. Hlavní KPI je ***RAV – Remaining Asset Value***, které popisuje procentuální zbytkovou hodnotu jednotlivých strojů vůči jejich původní hodnotě. KPI je pak filtrovatelné. Další ukazatele jsou pak ***Průměrné stáří strojů***, ***Celkový počet poruch*** (toto KPI by mělo být i zde u stavu aktiv), ***Aktuální počet dílů***, který ukazuje celkový počet vlastněných dílů a ukazatel je filtrovatelný na jednotlivé díly a poslední je ***Celkový počet použitých dílů***, z kterého lze sledovat spotřebu náhradních dílů, která linka potřebuje více, jaké a případně jak často. Grafy popisují RAV a náhradní díly. První je ***Spotřeba dílů dle stroje***, kde můžeme sledovat, jaký stroj potřebuje nejvíce dílů a je nejproblémovější a případně jaké díly. Další je graf ***Spotřeba dílů dle stáří stroje***, kde můžeme jasně vidět, že starší stroje potřebují více náhradních dílů. Vedle něho vidíme ***RAV na jednotlivá stroje***, kde je vidět, že některá zařízení jsou již procentuálně po své životnosti a jejich RAV je tedy 0 %, ale přesto stále fungují a využívají se. Dole pak vidíme graf ***RAV dle stáří stroje***, kde je vidět, že nejnovější stroje mají správně nejvyšší RAV a nejstarší mají naopak RAV již velmi nízké. Poslední graf je ***Počet poruch dle stáří stroje***, kde je jasně vidět, že starší stroje mají logicky nejvyšší počet poruch a nejnovější mají naopak poruch minimu. Report můžeme opět filtrovat na základě času, linek, strojů, typů stroje, stáří stroje a samozřejmě náhradních dílů.

7.6 Závěry návrhu a tvorby deskriptivního řešení

- Reporty mají vytvářet ***široký ucelený obrázek o celém fungování údržby***, zahrnující jak stroje a jejich výstupy, tak fungování oddělení údržby.
- Řešení datové analytiky ve formě reportů vytvářejí ***celkový obraz o oddělení údržby a fungování strojů,*** s vyobrazením nejdůležitějších KPI a současně se zaměřující i na ekonomickou stránku, kdy všechny reporty je možné filtrovat na základě nastavených dimenzí.
- Reporty musí být ***funkční, intuitivní, interaktivní, pracující s nejaktuálnějšími daty.***

8. Návrh a tvorba řešení – prediktivní úlohy



Účelem je kapitoly specifikace návrhu a tvorby řešení datové analytiky – prediktivní částí. Ta přímo navazuje na předchozí deskriptivní část a v rámci typů řešení je další pokročilou úrovní, jak řešit datovou analytiku.

Součástí této kapitoly bude definování cílů prediktivní analytiky neboli definice vytyčených predikcí, analýza dat z pohledu prediktivní analytiky, Preprocessing, Target Engineering, tvorba prediktivních modelů a další. Korektní a kvalitní vytvoření prediktivní části může dát dobrý základ pro ***vznik prediktivní údržby***.

Použitá data jsou synteticky ***vygenerovaná*** pomocí python scriptů, ale byla tvořena na základě reálných struktur a reálných jevů platících v průmyslových výrobních společnostech.

8.1 Definice cílů prediktivní analytiky – predikční úlohy

Jedná se o ***tři predikce***, které se zaměřují na ***předvídání vzniku poruch***, kde porucha vznikne, co to bude za poruchu a současně, jaká je ***pravděpodobná životnost stroje*** do další poruchy.

8.1.1 Predikce – Porucha do 72 hodin

Hlavní je zde predikce ***vznik poruchy do 72 hodin*** neboli 3 dny. Tato predikce má největší váhu, jelikož je jejím účelem přímo předpovědět, ***kde by mohla vzniknout porucha v daném časovém okně 72 hodin,*** což dává oddělení údržby jasné upozornění, které zařízení je potenciálně rizikové. Může tak podniknout kroky ke kontrole stroje, opravě stroje a potenciálně eliminaci poruchy. Pokud poruchu z určitých důvodů eliminovat nelze, může pak alespoň podniknout potřebné kroky ke snížení škody a snížení možných ztrát.

Okno 72 hodin (3 dny) bylo nastaveno z důvodu, že je to ***adekvátní rozsah***, kdy je možné na poruchu zareagovat. Pokud bychom nastavili 24 hodin (1 den), tak okno může být příliš krátké a reakce a eliminace může být nemožná. Pokud bychom naopak nastavili 168 hodin (týden), tak je okno příliš velké na to, aby byl model schopný poruchu vůbec identifikovat. ***Získávaná data z výroby*** a změny hodnot sledovaných dat se velmi často mohou začít měnit až v kratším rozsahu času před poruchou, týden by byl příliš, a porucha by byla neviditelná. Predikce se v tomto případě ***zakládá především na sledování fyzikálních hodnot*** ze stroje (teplota, tlak, vibrace atd.), ***na časových hodnotách stroje*** (stáří stroje, doba od poruchy atd.), ***počtu poruch*** apod.

8.1.2 Predikce – Typ poruchy

Druhá predikce, která přímo navazuje na první predikci, je ***predikce typu poruchy***. V závislosti na první predikci, která předpoví potenciální vznik poruchy, pak tato vezme identifikované rizikové stroje a na základě jejích dat předpoví, ***jaký typ poruchy by na tomto stroji pravděpodobně vznikl***. Typy poruch jsou pak rozdělené na mechanické, elektrické, přehřátí a tlakové, na základě sledovaných hodnot. Oddělení údržby může získat představu, co se ve stroji může porouchat a kde by měli příčiny poruchy a vadné části hledat. Predikce se znovu zakládá především ***na sledovaných datech fyzikálních veličin***, které přímo identifikují typ poruchy, ***na časových hodnotách***, jelikož různé stroje s různým stářím mohou vykazovat jiné poruchy, ***na typech historicky vzniklých poruch*** atd.

8.1.3 Predikce – Remaining Useful Life (RUL)

Třetí predikce má za úkol předpovídat ***Remaining Useful Life (RUL)***, které se udává v jednotkách času. V případě údržby se RUL definuje jako ***zbývající doba / životnost stroje do další poruchy***. Nejde tedy o celkovou životnost stroje za celou jeho existenci, ale pouze za jak dlouho dosáhne poruchového stavu. Predikce nepřímo navazuje na první predikci, kdy stroje identifikované jako rizikové do 72 hodin, by měly mít nejkratší hodnoty RUL. Predikce pak bude dávat informaci oddělení údržby, ***které stroje jsou v dobrém stavu a mají potenciál dlouhého fungování*** před vznikem poruchy, a naopak které stroje už se blíží k bodu očekávané poruchy. U těchto pak může údržba podniknout

potřebné úkony, aby se pokusila RUL stroje zvýšit, oddálit poruchu a obecně reagovat na stroj před jeho selháním.

8.2 Analýza dat z pohledu prediktivní analytiky (EDA)

Tato část textu se zaměřuje na ***průzkumnou analýzu predikčního datasetu***, tedy dat, která tvoří základ pro následné prediktivní modelování a všechny predikce. Tento krok je obecně označován jako ***Exploratory Data Analysis (EDA)*** a představuje úvodní fázi práce s daty, jejímž cílem je jejich detailní pochopení. V rámci této analýzy jsou identifikovány základní charakteristiky dat, případné chyby či anomálie, stejně jako vztahy mezi jednotlivými proměnnými. ***Porozumění datům*** je klíčovým předpokladem pro jejich správné analytické zpracování a následnou tvorbu kvalitních prediktivních modelů.

Zde je nutné zdůraznit, že ***v této fázi je analýza prováděna, nad již existujícím prediktivním datasetem, který je částečně připraven pro účely modelování.*** Dataset proto obsahuje nejen původní proměnné, ale také nově vytvořené odvozené vstupní proměnné (features) a cílové proměnné (targets). Tyto proměnné byly vytvořeny v rámci procesu přípravy a generování dat a jejich detailní popis bude uveden v následujících kapitolách ***Feature Engineering*** a ***Definice cílových proměnných (Target variables)***.

8.2.1 Základní charakteristiky datasetu

Hlavní prediktivní dataset byl tvořen ze základních ***zdrojových výrobních dat,*** která obsahují data výroby po hodinách. Zahrnují aktuální data o stavu stroje, hodnoty fyzikálních veličin, hodnoty výstupu stroje a časové hodnoty (stáří stroje, čas od poruchy). Byly odstraněny některé hodnoty ze zdrojových dat, která jsou pro prediktivní modely nepotřebné, přinášely by nepotřebnou informace či model spíše mátly.

Finální dataset obsahuje ***315 372 záznamů a 27 sloupců*** (včetně již vytvořených features a target sloupců). ***Základních sloupců*** ze zdrojových dat je pak 13 a zbytek jsou sloupce nově vytvořené pro potřeby modelů. Tabulka 8.1 ukazuje rozložení sloupců s jejich datovým typem a počtem chybějících hodnot, modře podbarvené jsou sloupce nově vytvořené pro modely.

Table 8-1: Sloupce s datovým typem a počtem chybějících hodnot

Sloupec	Datový typ	Počet NaN hodnot
timestamp	datetime64[us]	0
machine_id	int64	0
temperature	float64	0
vibration	float64	0
pressure	float64	0
load	float64	0
ambient_temperature	float64	0
is_running	int64	0
produced_units	int64	0
defective_units	int64	0
operating_hours_since_maintenance	int64	0
machine_age_hours	int64	0
health_index	float64	0
health_delta_24h	float64	288
rolling_temp_mean_24h	float64	276
rolling_temp_std_24h	float64	276
rolling_vibration_mean_24h	float64	276
rolling_vibration_std_24h	float64	276

Sloupec	Datový typ	Počet NaN hodnot
rolling_pressure_mean_24h	float64	276
rolling_pressure_std_24h	float64	276
running_ratio_24h	float64	276
hours_since_last_failure	float64	15939
failure_count_last_7d	int64	0
failure_count_last_30d	int64	0
target_failure_72h	int64	0
target_failure_type	str	284124
target_RUL_hours	float64	5047

Lze sledovat, že všechny ***základní sloupce nemají chybějící hodnoty***, což je způsobeno hlavně simulovaným generováním dat. To je dobrý jev kvality dat pro prediktivní modely. Chybějící hodnoty jsou především u nově vytvořených sloupců, což je způsobeno hlavně způsobem, jak byly sloupce tvořeny (bude popsáno v dalších kapitolách). Většina hodnot má číselnou hodnotu (int64, float64). Výjimkou je pouze timestamp, což je časová značka, a target proměnná, která určuje typ poruchy. Ta ze své podstaty bude zachycená textem.

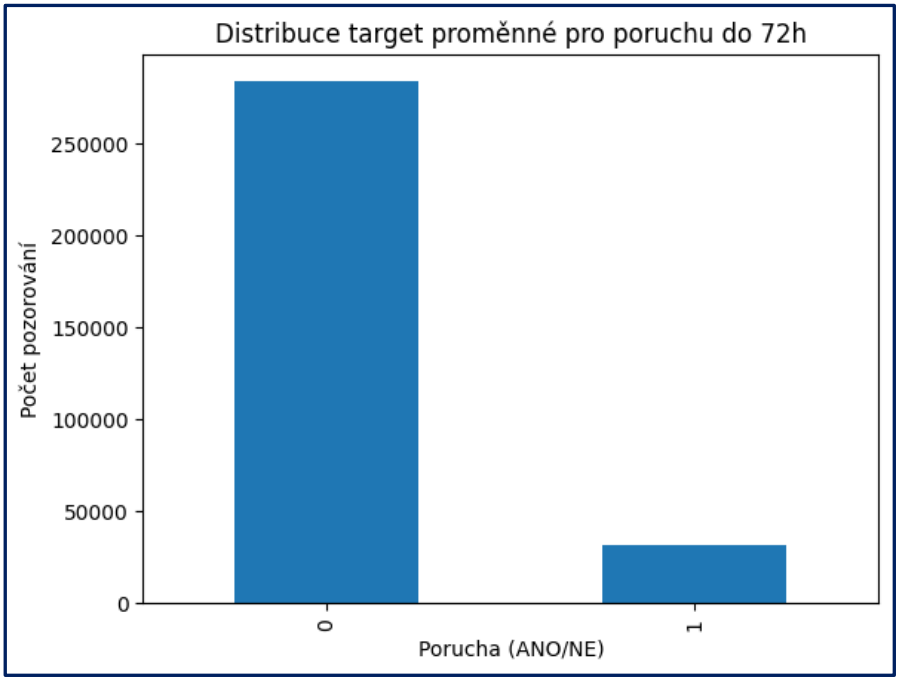
Na základě deskriptivních statistik pak nebyly v datasetu pozorovány ***žádné zjevně nerealistické hodnoty***, anomálie či extrémy, které by narušovaly použitelnost datasetu pro prediktivní modelování. Rozsahy jednotlivých proměnných působí realisticky a konzistentně. Hlavním z důvodů může být syntetický způsob generování dat, který zpravidla vede k vyšší čistotě dat a nižšímu výskytu náhodných odchylek oproti reálným provozním datům reálné společnosti.

8.2.2 Cílové proměnné

Dále byly zkoumány cílové (target) proměnné a jejich rozložení. Sem patří ***proměnné pro indikátor poruch v horizontu 72 hodin, ukazatel typu poruchy a ukazatel RUL do každé poruchy***. Podrobně budou popisovány v kapitole Target Engineering, která se na ně bude přímo zaměřovat.

V případě cílové proměnné pro poruchu do 72 hodin bylo zjištěno přibližné rozložení v poměru 90:10. Přibližně 90 % záznamů tedy odpovídá stavu, kdy u daného stroje nedojde k poruše v následujících 72 hodinách, zatímco přibližně 10 % záznamů představuje stav, kdy se stroj, již nachází v tomto rizikovém časovém okně před poruchou. Toto rozložení souvisí s celkovou četností poruch v datech. Vyšší počet poruch by vedl také k vyššímu zastoupení pozitivní třídy a tím i k vyváženějšímu rozdělení cílové proměnné. Současně však není cílem vytvářet dataset s nepřirozeně vysokým počtem poruch, protože takový stav by neodpovídal realistickému provozu výroby.

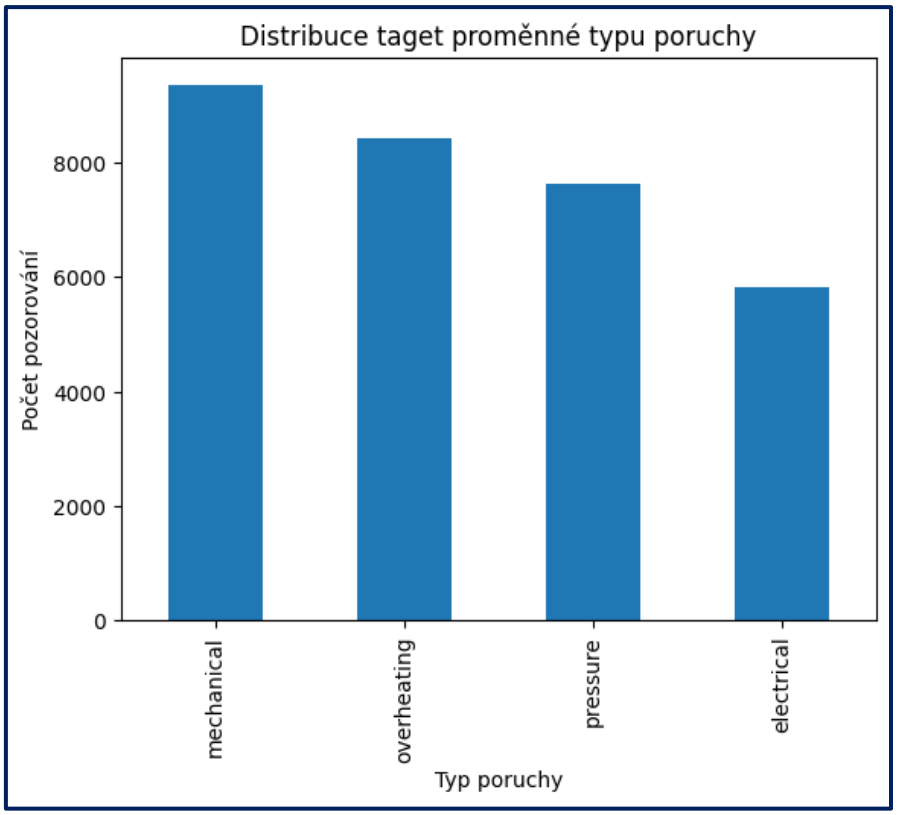
Rozložení cílové proměnné ukazuje další obrázek.



Obrázek 8-1: Graf distribuce cílové proměnné pro poruchu do 72h (zdroj: autor)

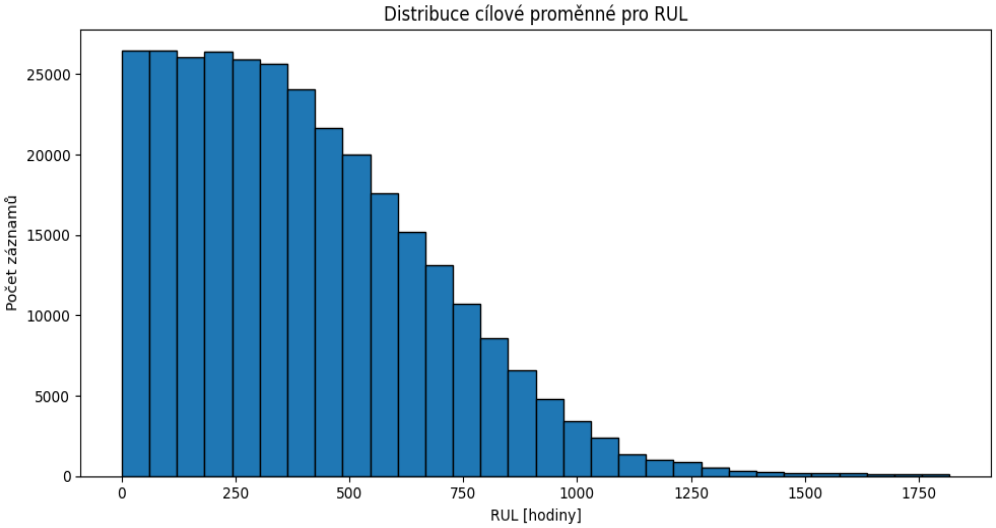
Z distribuce cílové proměnné vyplývá, že dataset je z pohledu ***této proměnné nevyvážený***. Tento stav je však v oblasti prediktivní údržby očekávatelný, protože ***poruchové stavy se vyskytují výrazně méně často*** než běžný provoz zařízení. Menšinová třída proto reprezentuje přirozeně vzácnější jev, což je třeba zohlednit při následném modelování i vyhodnocení modelu.

Další analyzovanou cílovou proměnnou je ***typ poruchy***, který určuje ***typ poruchového stavu při zaznamenaném selhání*** zařízení za celé sledované období. U této proměnné naopak nebyla zjištěna zásadní nevyváženost a žádný z typů poruchy není zcela v menšině. Zároveň ale zastoupení není zcela rovnoměrné, což je důležité pro model, který bude mít ***více záznamů pro častější typ poruchy***. Můžeme tedy konstatovat, že některé typy poruch jsou častější než jiné, ale žádný typ není jednoznačně vzácný. To je z pohledu modelování příznivé. Distribuci proměnné dokumentuje následující obrázek.



Obrázek 8-2: Graf distribuce cílové proměnné pro typ poruchy (zdroj: autor)

V rámci analýzy byla dále posouzena ***distribuce cílové proměnné pro RUL***, která vyjadřuje zbývající dobu do další poruchy stroje v hodinách. Z histogramu i základních deskriptivních statistik vyplývá, že hodnoty této proměnné jsou rozloženy v poměrně širokém intervalu přibližně od 0 do 1800 hodin. Rozdělení proměnné ***není rovnoměrné a mnohem větší koncentrace hodnot se nachází spíše v nižších a středních intervalech,*** zatímco velmi vysoké hodnoty jsou zastoupeny minimálně. Tento charakter rozdělení odpovídá tomu, že s blížící se poruchou se hodnota RUL přirozeně snižuje, zatímco vyšší hodnoty reprezentují stavy, kdy je porucha ještě časově vzdálenější či bezprostředně po opravě. Proměnná tak vykazuje dostatečnou variabilitu a představuje vhodný základ pro regresní modelování zbývajícího RUL. Rozložení proměnné ukazuje další obrázek.



Obrázek 8-3: Graf distribuce cílové proměnné pro RUL (zdroj: autor)

8.2.3 Závěry analýzy dat z pohledu prediktivní analytiky



- ***Datové typy*** proměnných jsou ***v pořádku*** a problém, který bude řešen v Data preprocessing, jsou chybějící NaN hodnoty některých proměnných.
- ***Distribuce cílových proměnných*** je celkově v pořádku a použitelná pro prediktivní modely.
- Určitá ***nevyváženost*** je pouze u cílové proměnné pro poruchu do 72 hodin a zde je tato nevyváženost očekávaná z důvodu realistické vzácnosti poruchového stavu.

8.3 Feature Engineering

Feature Engineering specifikuje, jak byly tvořeny nové vstupní proměnné. Termín ***feature*** označuje obecně ***jakoukoliv proměnnou v datasetu, která slouží jako vstup pro trénování*** prediktivního modelu. Kvalita těchto proměnných má zásadní vliv na výslednou výkonnost modelu, jelikož model je omezen informací obsaženou ve vstupních datech. Feature Engineering představuje ***proces, jehož cílem je zvýšení informační hodnoty dat*** prostřednictvím tvorby nových proměnných na základě původních dat. Tyto proměnné mohou vznikat například ***agregací či transformací***. Tento proces je nedílnou součástí přípravy dat pro modelování a významně přispívá ke zlepšení přesnosti a kvality prediktivních modelů (Murel & Kavlakoglu, 2024).

8.3.1 Klouzavé a rozdílové proměnné

Klouzavé proměnné, resp. *rolling features* jsou proměnné, které ***zachycují krátkodobý historický trend a vývoj existující proměnné*** za určitý omezený časový úsek ***do minulosti*** a agregují chování proměnné. V tomto případě byly klouzavé proměnné nastaveny v horizontu minulých ***24 hodin***. Znamená to, že vždy pro konkrétní záznam se pro konkrétní původní proměnnou (tlak, teplota atd.) vezmou všechny hodnoty za posledních 24 hodin. V následujícím záznamu opět 24 hodin, ale již posunuté o hodinu. Takto postupně tvoří ***krátkodobý trend vývoje*** a mohou zvýšit schopnosti modelu zachytit trendy a tvořit přesnější predikce. Všechny klouzavé proměnné znázorňuje Table 8-2.

Table 8-2: Klouzavé proměnné

Sloupec	Datový typ
health_delta_24h	float64
rolling_temp_mean_24h	float64
rolling_temp_std_24h	float64
rolling_vibration_mean_24h	float64
rolling_vibration_std_24h	float64
rolling_pressure_mean_24h	float64
rolling_pressure_std_24h	float64
running_ratio_24h	float64

Vzhledem k časovému charakteru dat jsou všechny odvozené proměnné konstruovány s ohledem na jejich časovou posloupnost. Pro klouzavé proměnné byly vybrány hodnoty, které znázorňují ***průměr (mean)*** a ***standardní odchylku od průměru (std – standard deviation) od průměru***). Například pro teplotu nám původní proměnná (*temperature*) z výroby řekne pouze aktuální hodnotu v danou hodinu. Klouzavou proměnnou průměru (*rolling_temp_mean_24h*) ale získáme informaci, jaká byla ***typická teplota za posledních 24 hodin*** a klouzavá proměnná odchylky (*rolling_temp_std_24 h*) znázorní, ***jak moc hodnota kolísala*** v tomto krátkodobém horizontu. Například velké výkyvy a nestabilita odchylky mohou začít naznačovat, že je něco špatně. Dokážou přesněji odhalit, že se stroj začíná chovat jinak, i když základní hodnota teploty zatím není extrémní. Tímto způsobem zvyšujeme informační hodnotu datasetu a dáváme modelům nové features, na kterých se může učit.

Další proměnnou je ***running_ratio_24h***, které není založené na fyzikální veličině, ale ***zakládá se na ukazateli is_running***. To je proměnná, která jasně říká, zda v danou hodinu stroj fungoval či nikoli. *Running_ratio_24h* následně znázorňuje ***podíl času, po který stroj fungoval za 24 hodin***. Hodnoty blíže k 1 znamenají, že stroj běžel téměř celou dobu, a naopak hodnoty směrem k 0 znamenají, že stroje téměř nefungoval. Proměnná může signalizovat, že vyšší intenzita provozu souvisí s vyšším opotřebením a vyšší vytížení bude vést k vyššímu riziku degradace. Nepravidelný provoz pak také může indikovat nestandardní stav stroje.

Poslední je časová ***rozdílová*** proměnná ***health_delta_24h***, která vychází z proměnné *health_index*. *Health_index* popisuje celkový stav stroje v daný moment a mění se v závislosti na „zdraví" stroje. *Health_delta_24h* znázorňuje ***změnu / rozdíl tohoto celkového stavu*** oproti stavu před 24 hodinami. Pokud se pak index rychle zhoršuje, model stav stroje a jeho vývoj vidí přímo.

Zde je potřeba zdůraznit, že ***health_index je uměle vytvořená proměnná***, která byla použita pro realistické generování poruch a jejich rozložení v čase v datech. V části preprocessingu se ukáže, že tato feature nebude použita, jelikož je uměle generovaná a reálně není zcela běžné, že by je společnosti sledovaly. Naopak *health_delta_24h* nebyla vyloučena. Je z něho sice odvozená, ale představuje změnovou charakteristiku, která je realisticky zastupitelná jiným výpočtem či transformací z reálně měřitelných veličin, která bude reprezentovat změnu stavu stroje. Tato naopak použita bude.

U těchto klouzavých proměnných ***se na začátku datasetu přirozeně objevují chybějící hodnoty***. Důvodem je, že jejich výpočet vyžaduje dostupnost historických pozorování za předchozích 24 hodin. Na počátku sledování jednotlivých strojů však tento časový kontext ještě není k dispozici, a proto nelze tyto proměnné pro první záznamy vypočítat. Nejedná se tedy o chybu datasetu, ale o očekávaný důsledek konstrukce časově odvozených proměnných.

8.3.2 Historické poruchové proměnné

Historické poruchové proměnné jsou ty, které ***vycházejí z počtu poruch a časových údajů***. Přidávají modelu další kontext ohledně života stroje, rozložení poruch či jejich rizikovosti. Proměnné, které sem patří, znázorňuje Table 8-3.

Table 8-3: Historické poruchové proměnné

Sloupec	Datový typ
hours_since_last_failure	float64
failure_count_last_7d	int64
failure_count_last_30d	int64

Proměnná ***hours_since_last_failure*** zachycuje počet hodin stroje uplynulých od poslední poruchy. To umožňuje modelu vidět kontext, v jaké fázi cyklu mezi poruchami se stroj právě nachází anebo doby, jaké mají různé stroje standartně mezi poruchami, což naznačuje obecný stav a poruchovost. Stroje s vysokými dobami mezi poruchami budou z obecného hlediska méně rizikové než ty s pravidelnějšími poruchami. Zde jsou v ***datasetu NaN hodnoty***, jelikož na začátku datasetu nelze přiřadit hodnotu, dokud nevznikne první porucha stroje.

Poslední proměnné ***failure_count_last_7d*** a ***failure_count_last_30d*** zachycují celkový počet poruch daného stroje za 7 dní a za 30 dní. 7 dní představuje kratší časový horizont a 30 dní naopak delší spíše střednědobý horizont. Poruchová historie stroje přidává další kontext technického stavu a poruchové intenzity strojů, kdy stroje s vysokou poruchovostí za daný časový horizont budou více rizikové a naopak. Proměnné pomáhají hodnotit každý stroj nejen podle aktuálních senzorů, ale přidávají kontext historické spolehlivosti stroje.

8.3.3 Závěr Feature Engineering



- V rámci Feature Engineering se vytvoří ***množství nových pomocných proměnných*** na základě původních dat v datasetu.
- Tyto proměnné pomáhají ***přidávat nový krátkodobý a dlouhodobý kontext chování*** pro učení prediktivních modelů a zvyšují informační hodnotu datasetu.

- Tento kontext především rozšiřuje původní data ***o historické stavy strojů a přidává vyšší nadhled,*** než je pouze aktuální stav.
- Nové proměnné pomohu ***modelům se lépe učit a potenciálně zvýšit kvalitu*** modelů, než by byla dosažena pouze s původními proměnnými.

8.4 Definice cílových proměnných (Target variables)

Tato podkapitola se zaměřuje na ***tvorbu a definici cílových proměnných*** (***Target Engineering***), které představují proměnné pro následně predikované modely. Cílová proměnná určuje, ***jaký jev bude model predikovat***. Model se v procesu učení učí ***vztah mezi vstupními proměnnými (features) a cílovou proměnnou (target),*** přičemž cílem je následně co nejpřesněji předpovídat její hodnotu na základě vstupních dat. Cílová proměnná může být buď přímo dostupná ve zdrojových datech, nebo je nutné ji odvodit a vytvořit na základě existujících informací. V kontextu prediktivní údržby je její definice klíčová, jelikož přímo ovlivňuje interpretaci výsledků i praktické využití modelu.

V rámci datasetu byly vytvořeny tři target proměnné, kdy každá je připravena pro jednu z nastavených predikčních úloh. Target proměnné ukazuje Table 8-4.

Table 8-4: Cílové proměnné

Sloupec	Datový typ
target_failure_72h	int64
target_failure_type	str
target_RUL_hours	float64

8.4.1 Target – Porucha do 72 hodin

První cílová proměnná je ***target_failure_72h***, což je cílová proměnná pro ***hlavní predikci poruchy do 72 hodin***. Proměnná znázorňuje, zda od aktuálního časového bodu nastane porucha na daném stroji v následujících 72 hodinách. Proměnná tak označuje ***záznamy, které spadají do rizikového období před vznikem poruchy***, na základě, kterých se model učí identifikovat vzory vedoucí k poruše.

Samotná úloha je ***klasifikační typ*** a target je ***binární proměnnou***, což znamená, že zobrazuje pouze dva stavy (ANO / NE či 0 / 1). V datasetu je tvořen tak, že ***pro každou jednotlivou poruchu jsou nalezeny záznamy v časovém horizontu 72 hodin*** před poruchou a pro ně je target proměnná nastavena jako 1. Pro ostatní záznamy target obsahuje 0. Model se pak učí na features v datasetu, které jsou právě v tomto časovém okně a sleduje jejich vývoj.

8.4.2 Target – Typ poruchy

Cílová proměnná pro druhou prediktivní úlohu typu poruchy je ***target_failure_type***. Proměnná obsahuje vždy konkrétní daný typ poruchy, jak byla vzniklá porucha klasifikována (mechanická, elektrická atd.). Je přímo navázaná na první prediktivní úlohu a tento typ poruchy je tedy určen pouze u záznamů, které mají ***target_failure_72h*** jako 1. Na rozdíl od první predikce je zde nemožné u ostatních záznamů s hodnotou 0 určit typ poruchy a z toho důvodu target sloupec obsahuje ***velké množství NaN hodnot***. Je to ale jeden z předpokládaných jevů target sloupce a pro účely trénování modelu budou využívány pouze záznamy, u kterých je cílová proměnná definována.

Typ úlohy je ***multiclass klasifikace***, jelikož se určuje z většího množství kategorií. Hodnota samotná je tedy ***kategoriální***. Model se pak učí na záznamech, které mají poruchu do 72 hodin a mají určený typ poruchy. Na základě vývoje features se pak snaží určit pravděpodobný typ poruchy, který vznikne.

8.4.3 Target – RUL

Poslední cílová proměnná je ***target_RUL_hours***, která přísluší ***predikci Remaining Useful Life (RUL)***. Proměnná obsahuje vždy hodnotu počtu hodin do následující poruchy. To znamená, že po

opravě stroje po poruše je v následujících záznamech určen přesný počet hodin až do následující poruchy. To znamená, že ***hodnoty postupně klesají***, hodinu před poruchou je počet hodin 1 a v moment poruchy 0. Konkrétněji je výpočet nastaven jako rozdíl mezi aktuálním časem a časem další poruchy. To je možné udělat díky tomu, že historicky známe, kdy další porucha vznikne. Hodnota RUL je při trénování modelu ***odvozena z historických dat***, přičemž poté model samotný využívá pouze vstupní proměnné dostupné v daném časovém okamžiku.

Zde se jedná o ***spojitý typ proměnné*** a modelem se snažíme předpovídat přesný počet hodin, za který by měla nová porucha vzniknout. Nejedná se již o klasifikační úlohu, ale spíše o ***regresní*** úlohu. Predikce pomáhá ***odhadovat zbývající dobu provozu*** a u snižujících se RUL konkrétních strojů může údržba provést potřebné akce, než porucha opravdu nastane.

8.4.4 Závěr definice cílových proměnných



- V rámci definice cílových proměnných jsou nastaveny ***3 target proměnné***, které jasně popisují jev / úlohu, kterou budeme chtít predikovat.
- Nyní je možné ***určit prediktivním modelům, co je jejich target proměnná***, na kterou se mají zaměřit. Bez jejich existence by nebylo možné prediktivní modely vůbec začít tvořit a získat jakékoliv výsledky.

8.5 Data preprocessing

Další ***krok tvorby prediktivního řešení*** je Data preprocessing neboli ***příprava dat***, aby byla vhodná pro použití a správnou funkci prediktivních modelů. ***Kvalita dat*** je jeden z hlavních faktorů, které pak ovlivňují kvalitu jakéhokoliv datově analytického řešení. Preprocessing zahrnuje ***jakýkoliv úkon na surových zdrojových datech***, který má za úkol je připravit pro další zkoumání a analýzy. Může zahrnovat odstranění chybějících hodnot, odstranění duplicit, opravování datových typů, ošetření extrémních hodnot, škálování a normalizaci hodnot a mnoho dalšího (Pykes, 2025).

Jelikož byla ***data synteticky generována a kontrolována***, tak už zdrojová data byla v relativně dobré kvalitě a měly pohlídány aspekty, jako jsou datové typy, extrémy apod. Některé běžné kroky byly realizovány již ve Feature a Target Engineering při tvorbě datasetu. Pro všechny prediktivní úlohy je pak použit stejný prediktivní dataset, a proto ***preprocessing pro všechny úlohy zahrnuje velmi podobné kroky,*** až na pár výjimek. Preprocessing proto bude popsán souhrnně, a pokud se bude něco pro konkrétní úlohu lišit, tak to bude patřičně zdůrazněno.

8.5.1 Seřazení datasetu

Jako základní krok pro všechny predikce po načtení dat je seřazení datasetu. To je potřeba, jelikož je dataset závislý na čase a je tedy potřeba, aby byl chronologický.

```
df = df.sort_values("timestamp")
```

8.5.2 Tvorba dataframe pro každou úlohu

Pro každou jednotlivou úlohu je vytvořen samostatný dataframe, který na základě nastavení obsahoval pouze data, která úloha správně potřebovala. V následujícím výpisu kódu jsou za sebou jednotlivě řádky kódu, které daný dataframe vytvořily.

Výpis 1: Definice dataframe pro predikce

```
# Porucha do 72 hodin
df_model = df.dropna(subset=["target_failure_72h"])

# Typ poruchy
df_type = df[df["target_failure_72h"] == 1].copy()

# RUL
df_rul = df.dropna(subset=["target_RUL_hours"]).copy()
```

Pro první predikci do 72 hodin je pouze formálně a pro jistotu vzat dataframe, který nebude mít hodnoty NaN. Pro predikci typu poruchy jsou vybrány pouze záznamy, které mají target_failure_72h hodnotu jako 1, což ***vybere jen záznamy, kde je daný typ poruchy***. Pro predikci RUL jsou vybrány pouze záznamy, které neobsahují v target hodnoty NaN.

8.5.3 Odstranění NaN hodnot hours_since_last_failure

Byly odstraněny NaN hodnoty sloupce určující počet hodin od poslední poruchy. Došlo k nahrazení NaN hodnot pomocí stáří stroje (*machine_age_hours*). Je vhodnou hodnotou, kterou lze dosadit, a pokud před první poruchou datasetu žádná porucha nevznikla, tak je stáří stroje validní hodnota.

```
df_model["hours_since_last_failure"] =
df_model["hours_since_last_failure"].fillna(df_model["machine_age_hours"])
```

8.5.4 Výběr vhodných features

Byly vybrány features, které se pro predikce hodí, jelikož nemají použitelnou a validní informaci anebo by naopak modelu spíše škodily. Úprava byla provedena ne pomocí konkrétního výběru proměnných pro model, ale pomocí výběru proměnných, které pak použity být nemají. Také sem patřily i target sloupce, dle kterých predikováno být nesmí. Pro zjednodušení a zachování konzistence byly pro jednotlivé modely použity převážně shodné vstupní proměnné.

Výpis 2: Výběr feature pro prediktivní modely

```
features = df_model.columns.drop([
    "timestamp",
    "machine_id",
    "target_failure_72h",
    "target_failure_type",
    "target_RUL_hours",
    "health_index",  # vynechat
    "produced_units",
    "defective_units"
]).tolist()
```

Zde je potřeba zdůraznit, že byla vynechána i proměnná *health_index*, jak již bylo popsáno. *Health_index* popisoval celkový stav stroje v daný moment a mění se v závislosti na „zdraví“ stroje. Je to však uměle vytvořená proměnná pro potřeby tvorby dat, která byla používána pro generování poruch ve zdrojových datech, k jejich realistickému rozložení a změnám fyzikálních hodnot. V praxi není zcela běžné, že by společnosti tento údaj sledovali, a proto by mohlo být na škodu ho v řešení pro modely nechávat. Proměnná pak měla v některých případech velkou váhu, což by mohlo prakticky uškodit ještě více.

8.5.5 Odstranění zbylých NaN hodnoty

V datasetu pak stále zbyly NaN hodnoty u klouzavých proměnných, které byly zapříčiněny způsobem, jakým byly features tvořeny. Celkem jich však bylo pouze kolem 280, než měly hodnoty možnost být vypočteny na začátku datasetu. To je zanedbatelné množství vůči velikosti datasetu, a proto byly záznamy s NaN hodnotami odstraněny.

```
df_model = df_model.dropna(subset=features)
df_model = df_model.reset_index(drop=True)
```

Také byl resetován index řádků, aby šly čistě popořadě od 0 po odstranění záznamů.

8.5.6 Rozdělení dat na trénovací a testovací

Následně bylo provedeno rozdělení dat pro potřeby modelů. V tomto případě ale nebyl použit klasický model v pythonu *train_test_split*, jelikož ten rozdělí data náhodně na základě daného nastaveného poměru. To by byl zde problém, jelikož data obsahují časovou posloupnost záznamů. Kdyby došlo

k náhodnému rozdělení, tak by byly záznamy zpřeházeny, nedržely posloupnost vývoje proměnných v čase, a také by se pak model mohl učit na datech z roku 2024 a testovat na datech z roku 2022. Je potřeba, aby se model učil z minulosti na starých záznamech a testoval na nových záznamech. Proto bylo provedeno vlastní rozdělení na základě časové značky, které rozdělí data 80/20 dle časového bodu. 80 % dat bude trénovacích a 20 % dat bude testovacích. Každá úloha má příslušný target *y*.

Výpis 3: Rozdělení datasetu na trénovací a testovací dle času

```
# Seřazení dat
df_model = df_model.sort_values("timestamp")

# Split podle času
split_time = df_model["timestamp"].quantile(0.8)
train = df_model[df_model["timestamp"] < split_time]
test = df_model[df_model["timestamp"] >= split_time]

# Vytvoření X a y
X_train = train[features]
y_train = train["target_failure_72h"]

X_test = test[features]
y_test = test["target_failure_72h"]
```

8.5.7 Škálování hodnot

Nakonec bylo pouze pro potřeby modelu logistické regrese, která byla zkoušena jako jedna z variant modelu, použito i škálování hodnot, které mělo za úkol vyrovnat rozdíly mezi hodnotami fyzikálních veličin. Škálování je pro logistickou regresi vhodný nástroj. Každá veličina se totiž pohybuje v jiných řádech a model je na toto typicky citlivý. Může být však rovnou řečeno, že logistická regrese v žádné z úloh jako nejlepší model nevyšla.

```
# škálování
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)
```

8.5.8 Závěr Data preprocessing



- V rámci Data preprocessingu se provádí několik důležitých kroků, aby byla ***zajištěna maximální čistota datasetu***.
- Tyto základní kroky mohou velmi ***pomoci kvalitně prediktivního modelu***.
- Dataset nyní ***neobsahuje prázdné hodnoty*** a každá úloha má připravena přesně ta dat, která potřebuje ke správnému fungování.
- Současně jsou ***data rozdělená na trénovací a testovací*** část, což je absolutní nutnost pro trénování a hodnocení jakéhokoliv prediktivního modelu.

8.6 Tvorba a trénování prediktivních modelů

Pro každou úlohu je vybráno a ***vyzkoušeno několik různých modelů***, které jsou vhodné pro daný typ úlohy. Z těchto modelů byl pak vybrán ten, který byl na základě hodnotících metrik po porovnání nejlepší. Všechny modely jsou ***trénovány na předem připravené sadě vstupních proměnných*** a testovány na časově oddělené sadě testovacích dat získané rozdělením v preprocessingu. U vybraných modelů je ***provedeno ladění hyperparametrů*** s cílem zlepšit jejich výkonnost. Jsou testovány různé kombinace parametrů a následně byl vybrán model s nejlepšími výsledky.

8.6.1 Modely – Porucha do 72 hodin

První úloha je typ ***klasifikační úlohy***, která má predikovat ***binární proměnnou stavu***, zda do 72 hodin vznikne či nevznikne porucha na základě vstupních proměnných. Pro tento typ predikce byly testovány modely, které jsou na toto vhodné. Modely, které ***byly testovány jsou následující***.

- Logistická regrese
- Random Forest
- XGBoost

8.6.1.1 Logistická regrese

Logistická regrese je jeden ze ***základních klasifikačních modelů***. Je obdobou lineární regrese, ale na rozdíl od ní je ***cílovou proměnnou modelu kategorie (kategoriální proměnná)*** a ne číselná hodnota. Na počátku byla vyvinuta právě pro predikce binárních proměnných, ale dnes se již vyvinula i ***pro multiclass klasifikace***. Princip modelu je v zásadě takový, že se snaží dojít k matematické funkci, která by vyjadřovala vztah mezi vstupními proměnnými a cílovou proměnnou, dle které by pak dokázal odhadovat i budoucí hodnoty. (Sharda et al., 2020).

8.6.1.2 Random Forest

Je to model, který tvoří ***množství rozhodovacích stromů*** na základě náhodně vybraných vzorcích dat s opakováním a ***v každém stromě má současně náhodný výběr vstupních proměnných***. Pro klasifikační úlohy je to jeden z nejznámějších a nejpoužívanějších modelů, který je vhodný i pro aktuální prediktivní úlohu (Sharda et al., 2020).

8.6.1.3 XGBoost

XGBoost (Extreme Gradient Boosting) je prediktivní model založený na ***kombinaci většího množství rozhodovacích stromů***. Vychází z ***principu gradient boostingu***, který je využit také u dalších úloh. Využívá metod boostingu, kdy ***model postupně přidává nové rozhodovací stromy***, přičemž každý další strom se snaží korigovat / snižovat chyby předchozího souboru stromů. Tímto způsobem dochází k postupnému snižování predikční chyby a ke zvyšování kvality výsledného modelu a samotných predikcí. ***XGBoost se často využívá při práci s tabulkovými daty***, kde obecně dosahuje velmi dobrých výsledků. (Stryker et al., 2024).

8.6.2 Modely – Typ poruchy

Druhá úloha je typ *multiclass klasifikace*, která ***má predikovat kategoriální proměnnou typu poruchy*** (mechanická, elektrická atd.). Je to tedy obdobný typ jako předchozí typ úlohy, ale cílová proměnná již není binární (pouze dvě hodnoty), ale ***má více hodnot, které může cílová proměnná nabýt***. Lze pro ni použít podobné modely, které byly testovány pro první úlohu. Tento typ predikce je ze své podstaty značně složitější než binární klasifikace.

- Logistická regrese
- Random Forest
- Histogram Gradient Boosting
- XGBoost

8.6.2.1 Logistická regrese

Stejně jako u první úlohy byl jako základní model testována logistická regrese. Ta je v dnešní době přizpůsobená, aby byla ***schopná predikovat nejen binární, ale i multiclass klasifikace***. Opět se snaží najít matematickou funkci, která by vysvětlovala cílovou proměnnou na základě vstupních. Dle povahy úlohy bude tato funkce / predikce značně složitější.

8.6.2.2 Random Forest

Opět zde lze použít Random Forest klasifikační model. Z obecného hlediska funkce modelů by měl Random Forest potenciálně ***dosáhnout lepších výsledků u multiclass klasifikace*** než Logistická regrese. Z povahy modelu ho zde však nelze vynechat.

8.6.2.3 Histogram Gradient Boosting

Histogram Gradient Boosting je ***založený na principu boostingu***. Tvoří ***sekvenčně větší množství rozhodovacích stromů, kdy další strom je vždy trénován tak, aby opravoval chyby předchozího***. Model vytvoří při učení první rozhodovací strom a predikci, na kterém jsou vyhodnoceny výsledky a spočítána *residua*, což jsou chyby modelů získané porovnáním predikcí a skutečnosti. Další model má za úkol udělat minimalizaci residuí předchozích stromů.

Postupným ***iterativním zlepšováním jednotlivých stromů*** je získávána maximální přesnost modelu a predikce. Výsledná predikce je pak kombinací všech vytvořených stromů. Histogram verze pak nepracuje s každou jednou hodnotou zvlášť, ale převádí vstupní proměnné do intervalů (*bins*), čímž vytváří histogramy. Je vhodný pro větší datasety. Tento přístup výrazně ***snižuje výpočetní náročnost*** a zrychluje trénování modelu při zachování srovnatelné přesnosti. Histogram Gradient Boosting často dosahuje kvalitních výsledků při práci s tabulkovými daty (Sharda et al., 2020).

8.6.2.4 XGBoost

Stejně jako u předchozí úlohy bude testován XGBoost, který je obecně ***velmi kvalitní pro práci s tabulkovými daty*** a je velmi dobrý, když jde o klasifikační úlohy. XGBoost představuje více optimalizovanou verzi Gradient Boostingu. Vzhledem k povaze multiclass klasifikace, jakožto složitější klasifikace, je nutné otestovat i pokročilejší modely.

8.6.3 Modely – RUL

Třetí úloha již není klasifikační úlohou, ale ***regresní úlohou***. Z tohoto důvodu je potřeba testovat jiné než předchozí klasifikační modely. Výsledkem regresní úlohy je konkrétní ***číselná hodnota***, a nikoliv stav či kategorie. Má za úkol ***predikovat Remaining Useful Life každého stroje***, což je pravděpodobný počet hodin do další poruchy. ***Testované modely*** jsou následující.

- Dummy Regressor
- Lineární regrese
- Random Forest Regressor
- Histogram Gradient Boosting Regressor
- XGBRegressor

Regrese je jeden z nejzákladnějších statistických modelů. Používá se pro nalezení a ***vyjádření vztahu a závislostí vstupních proměnných vůči vysvětlující / cílové proměnné***. Tento vztah se snaží jednoduše vyjádřit pomocí matematické funkce / rovnice, která jasně definuje vztah proměnných a následně může být použita pro predikci budoucích cílových proměnných. Tedy pokud do rovnice vložíme konkrétní hodnoty vstupních proměnných, měl by nám vyjít konkrétní výsledek cíle (Sharda et al., 2020).

8.6.3.1 Dummy Regressor

Dummy Regressor je ***základní regresní model***. Byl použit jako základní referenční model pro porovnání, ***zda další testované modely dosahují skutečně lepších výsledků***. Není určený pro použití ke skutečným predikcím (scikit-learn.org, 2026a).

8.6.3.2 Lineární regrese

Lineární regrese je základní regresní model, který ***se snaží hledat lineární vztahy v datech***. Model vezme vstupní proměnné a snaží se najít lineární vztah a závislost vůči cílové proměnné (*target*). Tento vztah se pak snaží popsat co nejpřesnější matematickou funkcí, pomocí které je pak schopný odhadovat cílové proměnné i pro nová vstupní data (Sharda et al., 2020).

8.6.3.3 Random Forest Regressor

Random Forest Regressor je ***regresní varianta modelu Random Forest***, který patří mezi metody založené na kombinaci většího množství rozhodovacích stromů. Model vytváří ***více stromů, jejichž výsledky jsou následně agregovány***. Na rozdíl od klasifikační varianty, kde dochází k hlasování o finální predikci jednotlivých stromů, je u regresní úlohy výsledná hodnota určena jako průměr predikcí všech stromů. Díky této kombinaci je model schopen zachytit i složitější nelineární vztahy v datech a zároveň dosahuje dobré stability. (Kavlakoglu, 2021; scikit-learn.org, 2026d).

8.6.3.4 Histogram Gradient Boosting Regressor

Histogram Gradient Boosting Regressor je model založený na gradient boostingu určený pro regresní úlohy. Jedná se ***o optimalizovanou variantu základního Gradient Boosting Regressor***, která je vhodná pro větší datasety s velikostí zhruba nad 10 000 záznamů. Hlavní rozdíl je v tom, že histogram verze nepracuje s každou jednou hodnotou zvlášť, ale převádí vstupní proměnné do intervalů (*bins*), čímž vytváří histogramy. Tento přístup výrazně ***snižuje výpočetní náročnost a zrychluje trénování*** modelu při zachování srovnatelné přesnosti. (Kaabar, 2025; scikit-learn.org, 2026c).

8.6.3.5 XGBRegressor

Poslední testovaný model, který ***je regresní variantou modelu XGBoost. XGBRegressor*** je zaměřený na regresní úlohy a funguje na principu rozhodovacích stromů, které jsou ale opět zaměřeny na hledání nelineárních vztahů, podobně jako Random Forest Regressor. Závěr tvorby a testování prediktivních modelů

Postupně byly popsány všechny modely, které byly testovány pro jednotlivé prediktivní úlohy. Díky rozdílnosti úloh ***musely být vždy vybrány ty, které jsou pro úlohu vhodné***. Z testovaných modelů bude následně vybrán ten s nejlepší kvalitou a přesností. Konečné shrnutí úloh a testovaných modelů můžete vidět v následující tabulce.

Table 8-5: Prediktivní úlohy a testované modely

Úloha	Typ úlohy	Testované modely
Porucha do 72h	Binární klasifikace	Logistická regrese, Random Forest, XGBoost
Typ poruchy	Multiclass klasifikace	Logistická regrese, Random Forest, Histogram Gradient Boosting, XGBoost
RUL	Regrese	Dummy Regressor, Lineární regrese, Random Forest Regressor, Histogram Gradient Boosting Regressor, XGBRegressor

8.7 Vyhodnocení a validace modelů

Nyní jsou jednotlivé ***modely porovnány na základě metrik kvality modelu*** a vybrán ten nejlepší. Tento nejlepší model bude následně použit pro samotné predikce konkrétní úlohy a může představovat základ pro zavedení prediktivní údržby do průmyslové společnosti.

8.7.1 Způsob vyhodnocení modelů

Z důvodu, že se u každé úlohy jedná o jiný typ predikce, ***je potřeba je jinak posuzovat***. Klasifikační úlohy mají vlastní kvalitativní metriky, které ale potom nejsou stejné pro regresní úlohy. Všechny modely byly ***trénovány a testovány na předem rozděleném souboru dat***, který byl rozdělený dle času. Modely tak byly trénovány na starších datech na trénovací sadě dat a následně testovány na datech novějších na testovací sadě dat, aby se učily z minulosti. Data byla v preprocessingu předem seřazena, aby byla zachována časová posloupnost vzniku dat. ***Metriky,*** které byly použity ***pro vyhodnocení kvality modelů*** jsou.

Porucha do 72 hodin a Typ poruchy (klasifikace)

- ***Accuracy*** – vyjadřuje schopnost modelu predikovat správnou hodnotu / kategorii nových dat. Zjistí se tak, že se skutečné hodnoty cílové proměnné v testovacích datech porovnají vůči predikcím modelu. Metrika ale může být zavádějící u nevyvážených dat, proto je potřeba hodnotit i jiné (Sharda et al., 2020).
- ***Precision*** – vyjadřuje podíl ***správně*** predikovaných pozitivních případů vůči celkovému počtu případů, které model označil jako pozitivní (správně i špatně). Udává tedy, jak spolehlivé jsou pozitivní predikce modelu (Sharda et al., 2020).
- ***Recall*** – vyjadřuje podíl ***správně*** zachycených pozitivních případů vůči celkovému počtu skutečně pozitivních případů (zachycených i nezachycených). Udává, kolik model zachytí ze všech skutečně vzniklých poruch (Sharda et al., 2020).

- ***F1 – score*** – vyjadřuje harmonický průměr mezi ***precision a recall***. Chceme u něho dosáhnout vyšších hodnot a reprezentuje vyváženost mezi schopností predikovat správně pozitivní případy a zachytit všechny skutečně pozitivní případy (scikit-learn.org, 2026b).

Remaining Useful Life (RUL) (regrese)

- ***MAE (Mean Absolute Error)*** – vyjadřuje průměrnou chybu predikcí vůči skutečným hodnotám cílové proměnné. Zde chceme nízké hodnoty (Waples, 2025).
- ***RMSE (Root Mean Squared Error)*** – vyjadřuje také průměrnou chybu jako MAE, ale je přísnější, kdy větší chyby mají mnohem větší váhu a je citlivější na extrémní chyby (Kasourova, 2026).
- ***R2*** – koeficient determinace vyjadřující, jak moc dokáže regresní model vysvětlit cílovou proměnnou a její variabilitu na základě vstupů. Hodnoty k 1 znamenají, že model dokáže velmi přesně vysvětlit vztahy v datech a naopak (Siddiqui, 2026).

8.7.2 Vyhodnocení – Porucha do 72 hodin

U predikce poruchy do 72 hodin se jedná ***o binární klasifikaci***, která má pomocí natrénovaného modelu upozornit na potenciální poruchu na určitém stroji v horizontu 3 dnů. Kvalitní model ***má potenciál upozornit oddělení údržby***, které může včas podniknout kroky k připravení se na poruchu, či úplné eliminaci poruchy. Table 8-6 prezentuje všechny metriky kvality testovaných modelů.

Table 8-6: Výsledky kvality modelů predikce poruchy do 72h

Úloha	Accuracy	Precision	Recall	F1
Logistická regrese	**0,90**	**0,52**	**0,99**	**0,68**
Random Forest	**0,92**	**0,57**	**0,95**	**0,71**
Random Forest - Tuned	**0,93**	**0,60**	**0,93**	**0,73**
XGBoost	**0,91**	**0,55**	**0,94**	**0,70**

V rámci testování modelů bylo provedeno i ladění modelu Random Forest, který je zachycen jako Random Forest – Tuned (termín *tuning* označující ladění hyperparametrů modelu). ***Na základě všech metrik kvality modelů byl vybrán jako nejlepší prediktivní model právě laděný Random Forest***.

Model dosahoval nejvyšší hodnoty ***F1*** (0,73), což znamená, že dosahoval nejvyšší vyváženosti mezi precision a recall, tedy predikovat správně pozitivní případy poruch a zároveň zachytit maximum skutečně vzniklých poruch.

Dále model dosahoval největší ***accuracy*** (0,93). Ta byla u všech modelů vysoká, což bude způsobeno tím, že je predikce pouze binárního typu a nevyvážeností, kdy většina záznamů patří do skupiny bez poruchy. Accuracy ale může být u nevyváženého datasetu zavádějící, a proto je potřeba hodnotit i ostatní metriky. Dle ní ale model predikoval správnou hodnotu velmi úspěšně.

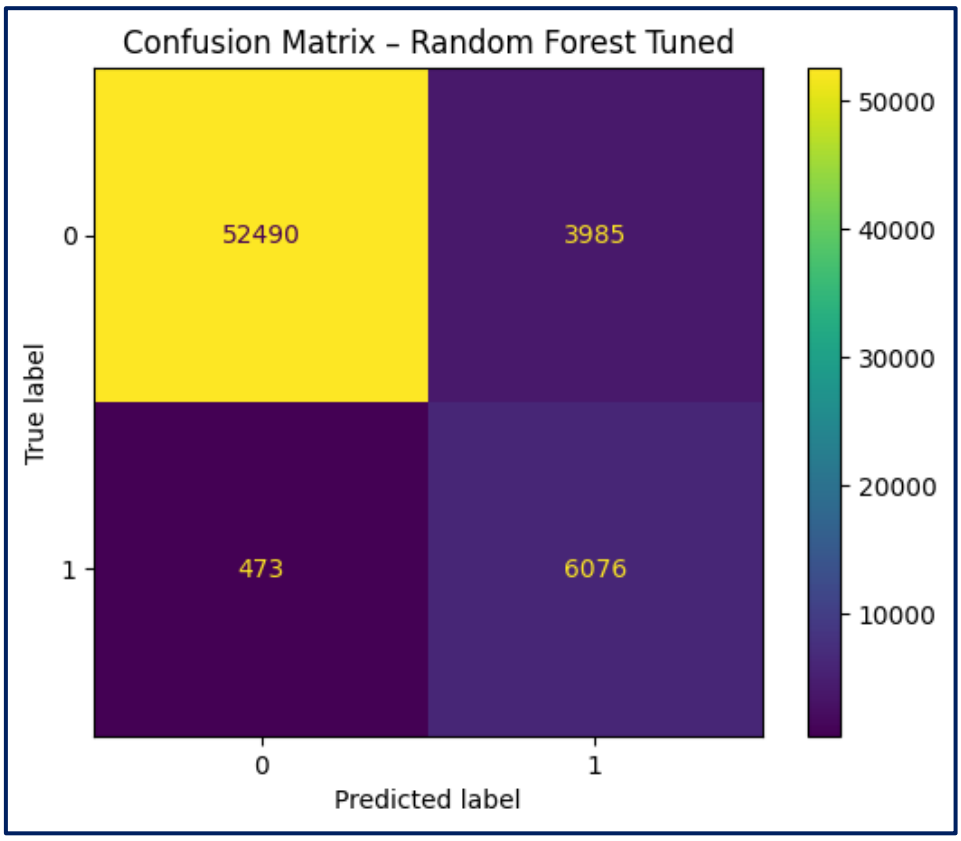
Také měl model nejvyšší ***precision*** (0,60), což znamená, že model má nejlepší schopnost správně identifikovat poruchy a zároveň omezovat falešné poplachy poruch.

Recall (0,93) je mírně nižší než u ostatních modelů, ale rozdíl je minimální a hodnota je stále vysoká. Logistická regrese sice dosáhla nejvyššího recall, ale její precision (0,52) zároveň indikuje, že bude mít mnohem větší množství falešně pozitivních poruch.

Zvážením kontextu všech metrik tedy měl model ***laděného Random Forest nejlepší výsledky a bude použit jako hlavní model pro následné predikce poruchy do 72 hodin*** v rámci návrhu řešení.

Pro další interpretaci výsledků byla zobrazena i ***Confusion Matrix, která zobrazuje vztah mezi skutečnými a predikovanými hodnotami,*** viz další obrázek. Na její diagonále má model velkou míru správně klasifikovaných případů. Množství poruch, které skutečně vznikly, ale model je nezachytil je relativně nízký (vlevo dole).

Naopak ***množství falešně pozitivních predikcí*** je vyšší (vpravo nahoře), ale prakticky je vždy lepší, aby model raději potenciální poruchu zachytil a falešně upozornil, aby oddělení bylo připravené. V kontextu prediktivní údržby je tento přístup žádoucí, jelikož je výhodnější potenciální poruchu zachytit a případně vytvořit falešné upozornění než poruchu zcela přehlédnout. Náklady na preventivní kontrolu jsou totiž výrazně nižší než ztráty způsobené nečekanou poruchou a prostojem. Model tak vykazuje ***vhodné opatrnější chování***, kdy preferuje vyšší míru zachycení poruch i za cenu zvýšeného počtu falešných poplachů.



Obrázek 8-4: Confusion Matrix pro model predikce poruchy do 72h (zdroj: autor)

Pro další doplnění pak byla vyhodnocena i hodnota ***AUC (Area Under Curve)***, která slouží k posouzení kvality modelu rozlišovat mezi oběma kategoriemi. Hodnota AUC vyšla ***0,977***, což indikuje velmi dobrou kvalitu modelu a naznačuje, že model má ***velmi dobrou schopnost rozlišovat poruchové a neporuchové stavy.***

8.7.3 Vyhodnocení – Typ poruchy

U druhé úlohy predikce ***typu poruchy jde o multiclass klasifikaci***, která má v návaznosti na první predikci určit, jaký bude typ poruchy, který pravděpodobně vznikne. Má za úkol dát oddělení informaci, na jaký typ se připravit, kde hledat poruchu či co by mohlo selhat. Ta na základě toho může podniknou potřebné kontroly a kroky k identifikaci a napravení daného typu poruchy. ***Tato predikce je již složitější než binární a má tedy horší výsledky než ta první***. Výsledky jednotlivých zkoušených modelů obsahuje Table 8-7.

Table 8-7: Výsledky kvality modelů predikce typu poruchy

Úloha	Accuracy	Precision (macro)	Recall (macro)	F1 (macro)	F1 (weighted)
Logistická regrese	**0,36**	**0,36**	**0,39**	**0,34**	**0,32**
Random Forest	**0,38**	**0,38**	**0,37**	**0,37**	**0,38**
Histogram Gradient Boosting	**0,42**	**0,43**	**0,43**	**0,42**	**0,41**
Histogram Gradient Boosting - Tuned	**0,43**	**0,44**	**0,44**	**0,43**	**0,43**
XGBoost	**0,42**	**0,44**	**0,43**	**0,42**	**0,42**

Pro úlohu bylo testováno více modelů a díky její vyšší složitosti je očekávané, že výsledky nebyly tak kvalitní, jako u úlohy první. ***Pro model Histogram Gradient Boosting bylo testováno opět ladění modelu a tento model nakonec na základě všech metrik vyšel jako nejlepší.*** Všechny jeho metriky dosahovaly nejlepších hodnot, a proto byl v tomto případě nejvhodnější volbou.

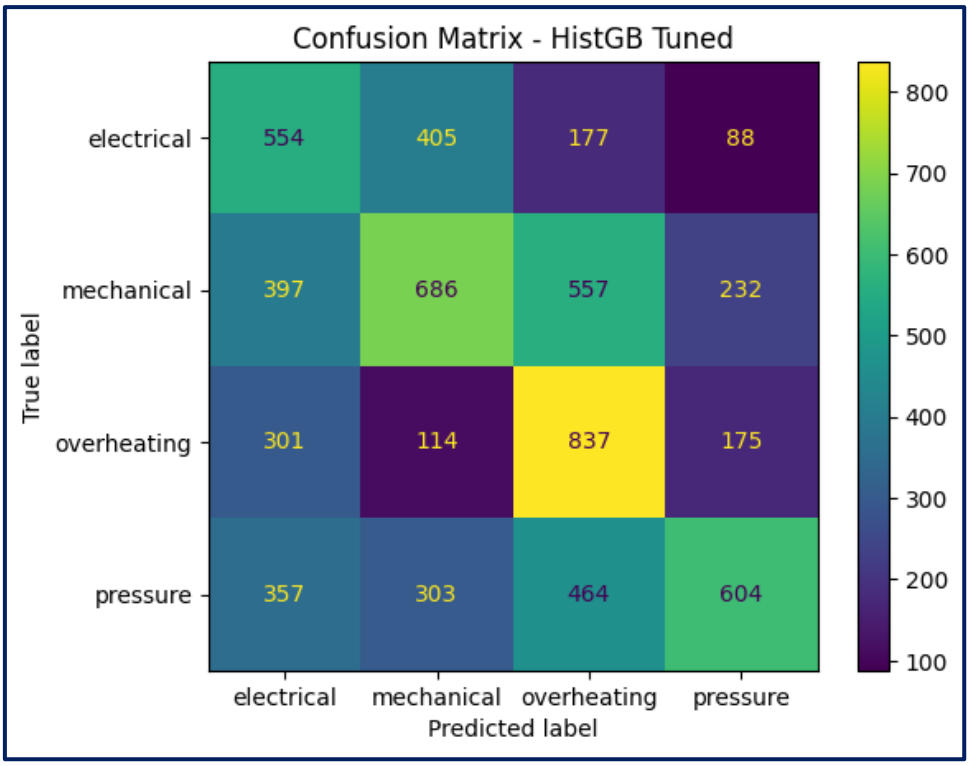
Metrika ***F1*** byla v tabulce zobrazena jako *macro* (0,43), která nezohledňuje rozložení jednotlivých kategorií v cílové proměnné a počítá průměr každé kategorie se stejnou váhou, a *weighted* (0,43), která naopak bere v potaz rozložení tříd v datasetu a vážený průměr nám může dát odlišnou hodnotu. Hlavní metrika je zde *macro*, jelikož chceme, aby model hodnotil každý typ poruchy rovnoměrně a ne, aby přinášel větší váhu více zastoupenému typu. Laděný Gradient Boosting měl nejlepší vyváženost precision a recall, tedy schopnosti skutečně správně klasifikovat případy typu poruchy a zároveň nejvíce skutečných typů.

Model měl dále nejvyšší ***accuracy*** (0,43), tedy nejčastěji správně predikoval vzniklý typ poruchy. Z pohledu jednotlivých metrik dosahoval zároveň nejvyšších hodnot ***precision*** (0,44) a ***recall*** (0,44), což indikuje, že model nejen nejvíce minimalizuje chybné klasifikace mezi třídami, ale zároveň dokáže zachytit většinu skutečně vzniklých typů poruch. Hodnoty precision a recall jsou velmi vyrovnané, což potvrzuje stabilní chování modelu napříč jednotlivými třídami.

Složitost úlohy ze své povahy snižuje výsledky všech modelů, ale v konečném důsledku byl ***laděný Histogram Gradient Boosting nejkvalitnějším testovaným modelem***, který bude použit pro jako ***hlavní model pro následné predikce typu poruchy*** v rámci návrhu řešení.

Dále byla pro model zobrazena ***Confusion Matrix, která zobrazuje vztah mezi skutečnými a predikovanými hodnotami typu poruchy.*** Je zobrazena na dalším obrázku. Matice zobrazuje, že model nejlépe klasifikuje typ poruchy *overheating*, ale zároveň však dochází k častým záměnám s typy poruch *mechanical* a *pressure*. Další větší záměny jsou mezi *mechanical* a *electrical*. Tento jev může být způsoben tím, že některé typy poruch jsou si v určitých aspektech podobné, a pokud rozdíly ve fyzikálních veličinách nejsou výrazné, tak mohou mít podobné indikátory.

Zde může být limitující faktor, že rozdíly a ***odchylky fyzikálních veličin jsou ze své podstaty malé, a tedy rozdíly nemusí být dostatečně výrazný*** pro jednoznačné rozlišení. V praxi může poruchu indikovat i relativně malá odchylka od běžných hodnot. Pokud by byly rozdíly hodnot markantnější, tak by byly indikátory poruch zřetelnější. Tyto výraznější rozdíly by pak sice mohly vést k lepší rozlišitelnosti tříd, ale na úkor realismu dat a byly by uměle měněny ve prospěch modelu.



Obrázek 8-5: Confusion Matrix pro model predikce typu poruchy (zdroj: autor)

8.7.4 Vyhodnocení – RUL

Třetí a poslední predikce je regresní úloha, která ***má za úkol předpovídat Remaining Useful Life (RUL), tedy celkový zbývající čas životnosti zařízení do další poruchy***. Jedná se o regresní úlohu z důvodu, že chceme predikovat konkrétní hodnotu. Cílem je, aby mělo oddělení informaci o tom, jaký počet hodin zhruba zbývá každému stroji do další poruchy, a tak mít přesně informaci, na které stroje se zaměřit dříve a která by měla ještě zůstat v provozu. ***U strojů s predikovaným nízkým RUL pak mohou předem provést kroky pro včasnou opravu, přípravu na poruchu či její úplnou eliminaci***. Výsledky kvality jednotlivých modelů obsahuje Table 8-8. Hodnoty chyb MAE a RMSE jsou v hodinách, jelikož RUL samotné je v hodinách.

Table 8-8: Výsledky kvality modelů predikce RUL

Úloha	MAE (h)	RMSE (h)	R^2
Dummy Regressor	**210,3**	**256,9**	**- 0,000033**
Lineární regrese	**96,4**	**119,6**	**0,783367**
Random Forest Regressor	**63,4**	**85,4**	**0,889461**
Histogram Gradient Boosting Regressor	**62,8**	**80,5**	**0,901662**
XGBRegressor	**61,6**	**77,7**	**0,908535**

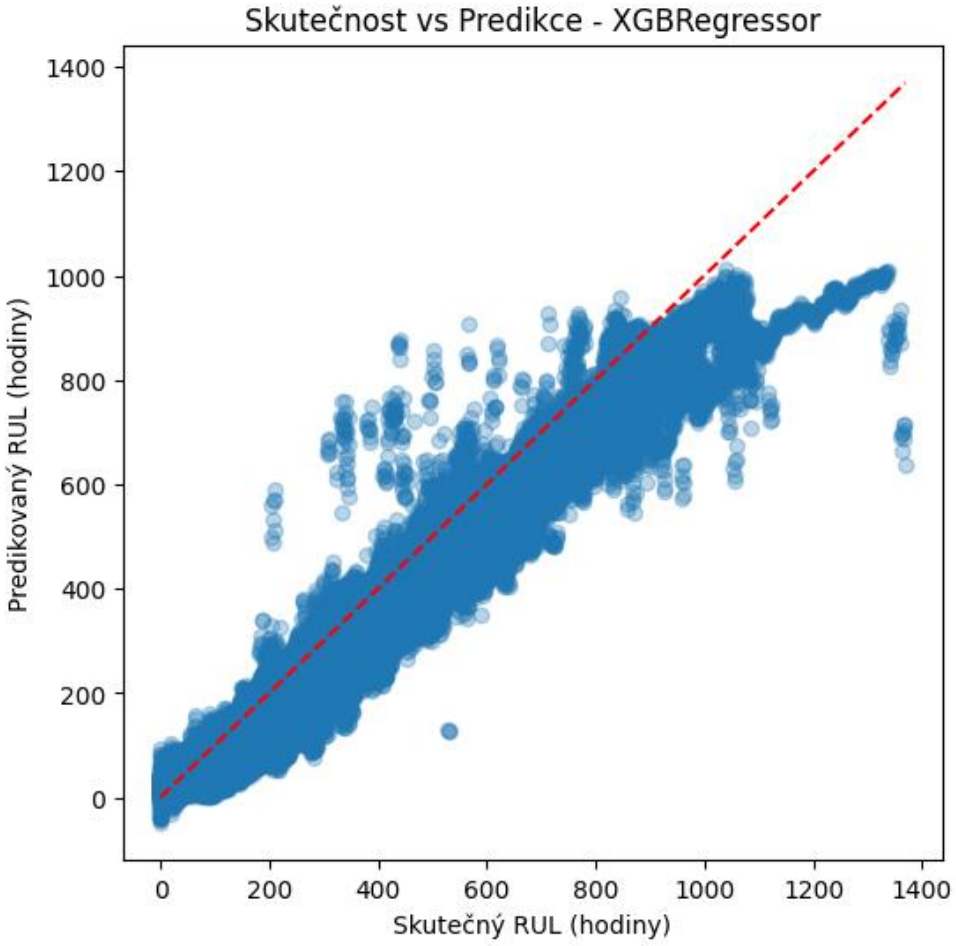
V rámci úlohy bylo testováno pět modelů, kdy první Dummy Regressor byl použit pouze jako referenční k porovnání ostatních modelů, zda se jejich kvalita zlepšuje. ***Po otestování byl jednoznačně nejlepším modelem XGBRegressor, který měl ve všech metrikách nejlepší hodnoty*.**

Model měl nejnižší hodnotu ***MAE*** (61,6 h), což znamená, že jeho predikce RUL byly nejpřesnější vůči skutečným hodnotám RUL. Interpretace je taková, že se predikce RUL v průměru odchýlí o 2,5 dne. Stejně tak měl nejnižší hodnotu metriky ***RMSE*** (77,7 h), což opět značí, že byly predikce nejpřesnější. Vzhledem k tomu, že RMSE více penalizuje větší chyby, tak dle něho se predikce v průměru odchýlí až o 3 dny, ale odchylka nebyla výrazně vyšší.

Hodnota R^2 (0,908535) blízká k 1 pak značí, že model dokáže dobře vysvětlovat a popisovat variabilitu cílové proměnné pomocí vstupních proměnných v datasetu. To znamená, že na základě změn vstupních proměnných dokáže velmi dobře popsat, jak se cílová proměnná změní.

Na základě všech hodnotících metrik tak vyšel ***XGBRegressor jako nejlepší varianta a bude použit jako hlavní model pro predikci Remaining Useful Life (RUL).***

Dále byl pro interpretaci vytvořen bodový graf, který zobrazuje vztah mezi predikovanými a reálnými hodnotami RUL, viz další obrázek. Jednotlivé ***predikce se dobře nacházejí okolo diagonály*** znázorňující reálné hodnoty, což potvrzuje dobré predikční schopnosti modelu. Většina predikcí se nachází pod diagonálou, což znamená, že model predikce spíše podhodnocuje a ***predikce jsou menší než skutečné hodnoty***. To je z praktického hlediska lepší, jelikož je výhodnější, když nám model udá nižší RUL, což vede k tomu, že se oddělení začne zabývat strojem dříve, než je jeho skutečné RUL. Pokud by byl scénář opačný, tak by porucha vznikla dříve než predikované RUL, a to by byla mnohem rizikovější chyba modelu. U vyšších hodnot je patrný větší rozptyl predikcí, což naznačuje menší přesnost modelu při predikci delšího RUL.



Obrázek 8-6: Bodový graf skutečnosti vs. predikcí RUL (zdroj: autor)

8.7.5 Závěr vyhodnocení a validace modelů



- Pro ***predikci poruchy do 72 hodin*** byl nejlepším modelem laděný Random Forest, který měl nejvyšší F1 metriku a nízký počet nezachycených poruch při lepší vyváženosti precision a recall.
- U ***predikce typu poruchy*** byl nejlepším modelem laděný Histogram Gradient Boosting, který překonal ostatní modely a dosáhl nejlepší celkové výkonnosti při určování typu poruchy.
- V poslední predikci ***zbývajícího RUL*** byl nejlepší model XGBRegressor, který měl nejnižší hodnoty průměrných chyb a vysokou hodnotu R^2, která značí dobrou schopnost vysvětlovat variabilitu cílové proměnné.
- Tyto vítězné tři modely budou ***použity pro celkové řešení prediktivní analytiky oddělení údržby*** a budou poskytovat predikce použitelné pro zlepšení fungování oddělení. Zajistí tak tři důležité predikce, kdy budou moci identifikovat včas poruchu, zjistit její typ a zjistit, jaký má každý stroj zbývající Remaining Useful Life.

Table 8-9: Vybrané nejlepší prediktivní modely

Úloha	Vybraný model	Důvody výběru
Porucha do 72h	Laděný Random Forest	Nejlepší F1, precision a vysoký recall / AUC
Typ poruchy	Laděný Histogram Gradient Boosting	Nejlepší accuracy, F1 (macro i weighted) a recall
RUL	XGBRegressor	Nejnižší MAE a RMSE a nejvyšší R^2

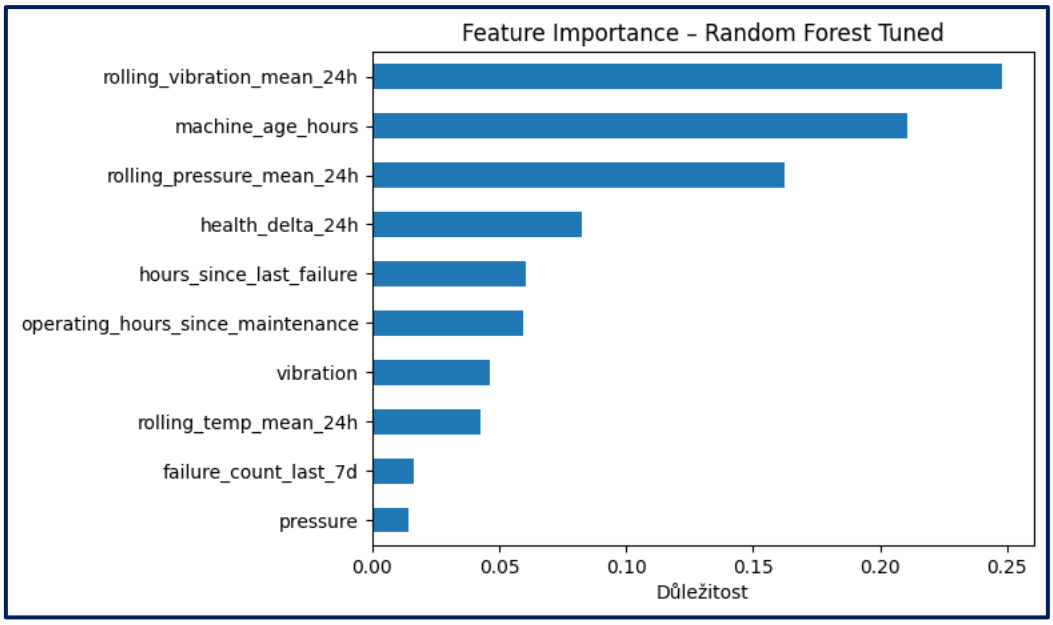
8.8 Interpretace a vizualizace výsledků

V další interpretaci a jsou ***popsány i samotné výsledné predikce modelů***. U každého ***nejlepšího prediktivního modelu se hodnotí nejen jeho kvalita, ale i reálné předpovědi,*** které z něho vychází. V tomto případě reálné predikce, které může oddělení údržby využít. Ty musí být pro reálné využití zobrazeny v přijatelné intuitivní formě, aby byl reálný přínos modelu využitelný pro běžného uživatele. To je v tomto případě například zaměstnanec oddělení údržby či vedoucí, který kontroluje pravděpodobné zbývající RUL strojů apod.

8.8.1 Interpretace a výsledky – Porucha do 72 hodin

Pro první úlohu predikce poruchy do 72 hodin byl nejlepším modelem ***laděný Random Forest***. Měl nejlepší výsledky a na základě úlohy dává binární předpovědi pro každý stroj, který by mohl mít potenciálně do 3 dnů poruchu, což nepředstavuje jistotu, ale odhad založený na pravděpodobnosti, kterou model odhaduje na základě naučených dat.

V rámci dalšího popisu modelu byla provedena ještě ***analýza feature importance***, která udává, jaké vstupní proměnné mají největší váhu pro model, když provádí predikce. Její graf je zobrazen na dalším obrázku.



Obrázek 8-7: Feature importance pro laděný Random Forest (zdroj: autor)

Z grafu vychází, že ***nejdůležitější proměnné jsou rolling_vibration_mean_24h, machine_age_hours a rolling_pressure_mean_24h.*** To znamená, že predikce se velmi zakládá jak na časových hodnotách, tak na změnách hodnot fyzikálních veličin ve výrobě. Současně jsou na ***vrcholu vibrace,*** což je shodné s realitou, jelikož vibrace jsou často považovány za jeden z nejlepších a nejvčasnějších indikátorů opotřebení a potenciální poruchy. Model tedy používá dobré relevantní hlavní proměnné pro predikci poruch. Nakonec můžeme vyvodit, že ***tvorba nových klouzavých proměnných byla přínosným krokem,*** jelikož většina zobrazených nejdůležitějších proměnných jsou právě tyto a mají tedy pro model větší váhu než aktuální hodinové hodnoty stavu.

V konečném důsledku můžeme konstatovat, že model ***vykazuje dostatečnou kvalitu*** a predikce ***zakládá na správných vstupních proměnných,*** které jsou dobrým indikátorem poruchy.

8.8.1.1 Výsledné predikce poruchy do 72 hodin

Dále přejdeme k interpretaci samotných predikcí modelu, které budou poskytovat pomocnou informaci pro fungování oddělení údržby a včasnou reakci na poruchy. ***Výsledné predikce*** zobrazuje následující obrázek, což je přehledná tabulka, která jasně indikuje, ***které stroje model identifikoval jako rizikové, a které rizikové nejsou.***

	Stroj ID	Stroj název	Predikce poruchy (0/1)	Predikce poruchy (ANO/NE)	Riziko poruchy do 72 h (%)	Úroveň rizika
0	1	L1 - Řezací stanice	1	ANO	77.1	VYSOKÁ
1	12	L3 - Testovací stanice	1	ANO	71.9	VYSOKÁ
2	11	L3 - Montážní stanice	1	ANO	50.9	STŘEDNÍ
3	3	L1 - Montážní stanice	0	NE	10.8	NÍZKÁ
4	5	L2 - Řezací stanice	0	NE	2.7	NÍZKÁ
5	7	L2 - Montážní stanice	0	NE	1.7	NÍZKÁ
6	6	L2 - Lis	0	NE	0.1	NÍZKÁ
7	10	L3 - Lis	0	NE	0.1	NÍZKÁ
8	9	L3 - Řezací stanice	0	NE	0.1	NÍZKÁ
9	2	L1 - Lis	0	NE	0.1	NÍZKÁ
10	8	L2 - Testovací stanice	0	NE	0.1	NÍZKÁ
11	4	L1 - Testovací stanice	0	NE	0.1	NÍZKÁ

Obrázek 8-8: Tabulka výsledných predikcí poruchy do 72h (zdroj: autor)

Výstup modelu je navržen tak, aby byl snadno interpretovatelný i pro uživatele bez znalosti datové analytiky. Výsledná tabulka obsahuje identifikaci každého stroje pomocí ***ID*** i dle jeho ***názvu*** pro

okamžité určení problémového stroje. Dále jsou již ***konkrétní predikce***, které jsou zobrazeny jak ve formě **0 / 1**, tak intuitivněji pro uživatele pomocí ***ANO / NE***. Tyto hodnoty znamenají, který stroj je rizikový do 72h a který ne. Dále je sloupec **„*Riziko poruchy do 72h (%)*“**, který vyjadřuje, jaká je dle modelu pravděpodobnost, že na základě vývoje vstupních dat vznikne porucha do 72 hodin. Poslední je sloupec **„*Úroveň rizika*“**, což je záměrně přidaný vizuální intuitivní sloupec, který na první pohled zhodnotí výši rizika. Tabulka je seřazená dle rizikovosti, tedy intuitivně jsou nejrizikovější stroje na vršku tabulky.

V praxi by to fungovalo tak, že zaměstnanec údržby přistoupí k výsledkům predikce a okamžitě identifikuje stroje, které jsou rizikové, že do 72h vznikne porucha. Na tomto základě může být nastaven plán a jasné kontroly a akce pro zjištění vad a příčin možné poruchy na těchto strojích. Současně na základě pravděpodobnosti poruchy může být určeno pořadí, ve kterém je potřeba rizikové stroje řešit. Z aktuálních predikcí z nejnovějších dat jsou tedy ***rizikové stroje L1 – Řezací stanice, L3 – Testovací stanice a L3 – Montážní stanice***. Tyto stroje by měly být prioritně řešeny a měly by se podniknout potřebné kroky. Z pravděpodobnosti pak vyplývá, že zcela ***nejrizikovější je L1 – Řezací stanice***, což znamená, že to je primární stroj, který musí být zkontrolován a řešen jako první.

Výsledky predikce tedy mohou pomoci regulovat fungování oddělení údržby a správnými kroky pomoci zefektivnit reakce na poruchy, snížit jejich dopady a případně je zcela eliminovat, což může odstranit vzniklé prostoje a ztráty z nich vzniklých.

8.8.2 Interpretace a výsledky – Typ poruchy

Pro druhou úlohu predikce typu poruchy byl nejlepším modelem ***laděný Histogram Gradient Boosting***. Kvůli složitosti multiclass úlohy byly výsledky jednotlivých testovaných modelů celkově nižší než u úlohy první, ale vítězný model dosáhl u všech metrik nejvyšších hodnot a byl tak nejlepší pro predikci typu poruchy.

Predikce funguje tak, že je ***navázána na první predikci a k rizikovým strojům s potenciální poruchou*** do 72 hodin se snaží predikovat, jaký typ poruchy by měl pravděpodobně vzniknout. Predikce tedy převezme jen vybrané stroje s predikcí poruchy do 72 hodin ANO a k nim následně přiřadí pravděpodobný typ poruchy.

8.8.2.1 Výsledné predikce typu poruchy

Konkrétní výstupy predikce typu poruchy jsou zobrazeny v tabulce na následujícím obrázku. Ke každému rizikovému stroji první predikce je druhou přiřazen typ poruchy a jeho pravděpodobnost.

	Stroj ID	Stroj název	Predikce poruchy (ANO/NE)	Riziko poruchy do 72 h (%)	Úroveň rizika	Predikovaný typ poruchy	Pravděpodobnost typu poruchy (%)
0	1	L1 - Řezací stanice	ANO	77.1	VYSOKÁ	Mechanická	35.4
1	12	L3 - Testovací stanice	ANO	71.9	VYSOKÁ	Elektrická	95.0
2	11	L3 - Montážní stanice	ANO	50.9	STŘEDNÍ	Mechanická	59.7

Obrázek 8-9: Tabulka výsledných predikcí typu poruchy (zdroj: autor)

Výstup modelu je navržen tak, aby byl snadno interpretovatelný i pro uživatele bez znalosti datové analytiky. Tabulka vezme jednotlivé stroje predikované první predikcí poruchy do 72h. To jsou zde stroje *s* ***ID 1, ID 12 a ID 11***. Znovu tedy zobrazuje ***ID***, ***název stroje***, ***predikci poruchy***, ***riziko poruchy*** *a* ***vizuální úroveň rizika***. K nim pak v posledních dvou sloupcích **„*Predikovaný typ poruchy*“** udá typ poruchy a dále **„*Pravděpodobnost typu poruchy (%)*“**, což je pravděpodobnost modelu, že vznikne právě tento typ poruchy. Znamená to opět, že predikce nedává typ poruchy jako jistý fakt, ale jako odhad založený na pravděpodobnosti z naučených dat.

V praxi to znamená, že zaměstnanec údržby z první predikce identifikuje rizikové stroje. Na jejím základě pak přejde k aktuální druhé predikci a dozví se, že hrozba na strojích je konkrétního typu. Pro stanici ***L1 – Řezací stanice***, která je ***nejrizikovější***, je ***pravděpodobný typ poruchy*** na základě nejnovější predikce ***mechanická***. V návaznosti může údržba přejít na stroj a hledat konkrétní příčiny, které by mohly vést k mechanickému typu poruchy. Dále predikce může ovlivňovat typ potřebných nástrojů, potřebný typ náhradních dílů či potřebné schopnosti zaměstnance

Vzhledem k povaze úlohy a potenciální podobnosti typů poruch dle změn hodnot se může stát, že i další typ poruchy bude mít nezanedbatelnou pravděpodobnost. Proto byla vytvořena i ***další tabulka, která zobrazí pravděpodobnost i ostatních typů poruch***. Je zobrazena na následujícím obrázku. Predikce dohromady dávají 100 % a je to tedy rozdělení pravděpodobnosti mezi jednotlivé typy. Kontrolou této tabulky a zjištěním, že jiný typ má také vyšší pravděpodobnost, pak může údržba získat informaci, že v druhé řadě by měla potenciálně zkontrolovat i jiný typ poruchy. To přidává zobrazeným výsledkům opět novou informační hodnotu.

Konkrétně u nejnovější predikce to znamená, že na stroji ***L1 – Řezací stanice je nejpravděpodobnější*** poruchou typ ***mechanický*** (35,4 %). Současně ale typ ***tlakové*** poruchy má ***taktéž vyšší pravděpodobnost*** (33,6 %), a proto je vhodné po kontrole možných mechanických příčin poruchy zkontrolovat i možné příčiny tlakové poruchy. Jde tedy o nejistou predikci. Nejvyšší hodnota je vždy zvýrazněna pro vizuální přehlednost.

	Stroj ID	Stroj název	Elektrická (%)	Mechanická (%)	Přehřátí (%)	Tlaková (%)
0	1	L1 - Řezací stanice	5.9	**35.4**	25.2	33.6
1	12	L3 - Testovací stanice	**95.0**	2.0	1.3	1.7
2	11	L3 - Montážní stanice	8.7	**59.7**	31.2	0.4

Obrázek 8-10: Tabulka pravděpodobností jednotlivých typů poruch (zdroj: autor)

Výsledky predikce ***typu poruchy mohou pomoci včas identifikovat, jaký typ poškození může údržba očekávat a podniknout včas potřebné kontroly*** konkrétnějších částí stroje, které by mohly tento typ způsobovat. Predikce tedy vytvoří konkrétnější zadání, kdy nebude vždy kontrolován celý stroj, ale hlavně prioritní části stroje. Po včasné identifikaci si může údržba rychleji připravit nástroje, náhradní díly či lidské zdroje a snažit se o snížení dopadu či úplnou eliminaci poruchy. To u klasického reaktivního typu údržby není možné a tyto úkony se dějí až po vzniku poruchy.

8.8.3 Interpretace a výsledky – RUL

Pro třetí úlohu predikce ***Remaining Useful Life (RUL)*** vyšel jako nejlepší model ***XGBRegressor.*** Šlo o regresní úlohu a nikoli klasifikační jako u předchozích úloh. Vítězný model měl nejnižší hodnoty chyb MAE a RMSE, takže jeho predikce byly nejbližší skutečným hodnotám a nejvyšší hodnotu R^2, takže model pomocí vstupních proměnných nejlépe popisovaly variabilitu cílové proměnné.

Model fungoval samostatně bez návaznosti na ostatní predikce, ale přesto souvislost mezi predikcemi existuje. Vzhledem k tomu, že predikuje zbývající životnost do poruchy, tak ***nejnižší hodnoty RUL by měly být u strojů nejblíže k další poruše, tedy stroje, které byly první predikcí identifikovány jako rizikové do 72 hodin.***

Tabulka výsledných predikcí RUL pro všechny stroje je zobrazena na následujícím obrázku.

	Datum	Stroj ID	Stroj název	Predikované RUL (hodiny)	Predikované RUL (dny)	Úroveň rizika	Doporučení
0	2024-12-31	1	L1 - Řezací stanice	24.5	1.0	VYSOKÁ	Okamžitá kontrola
1	2024-12-31	12	L3 - Testovací stanice	49.2	2.0	VYSOKÁ	Okamžitá kontrola
2	2024-12-31	11	L3 - Montážní stanice	71.3	3.0	VYSOKÁ	Okamžitá kontrola
3	2024-12-31	3	L1 - Montážní stanice	123.9	5.2	STŘEDNÍ	Plán údržby
4	2024-12-31	7	L2 - Montážní stanice	132.7	5.5	STŘEDNÍ	Plán údržby
5	2024-12-31	9	L3 - Řezací stanice	316.7	13.2	NÍZKÁ	Běžná kontrola
6	2024-12-31	5	L2 - Řezací stanice	323.9	13.5	NÍZKÁ	Běžná kontrola
7	2024-12-31	6	L2 - Lis	363.7	15.2	NÍZKÁ	Běžná kontrola
8	2024-12-31	10	L3 - Lis	386.6	16.1	NÍZKÁ	Běžná kontrola
9	2024-12-31	8	L2 - Testovací stanice	448.8	18.7	NÍZKÁ	Běžná kontrola
10	2024-12-31	4	L1 - Testovací stanice	495.6	20.6	NÍZKÁ	Běžná kontrola
11	2024-12-31	2	L1 - Lis	695.4	29.0	NÍZKÁ	Běžná kontrola

Obrázek 8-11: Tabulka výsledných predikcí RUL (zdroj: autor)

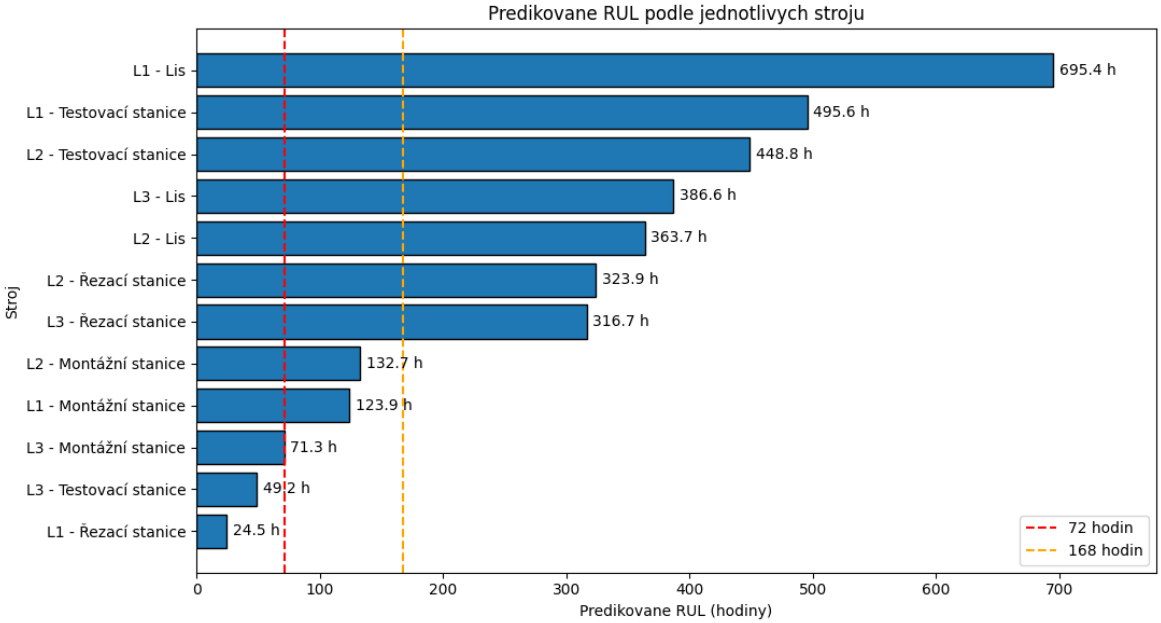
Výstup modelu je navržen tak, aby byl snadno interpretovatelný i pro uživatele bez znalosti datové analytiky. Tabulka obsahuje veškeré stroje společnosti a zobrazuje pro ně zbývající hodnotu RUL.

Obsahuje sloupec **„*Datum*“** pro kontrolu, zda jsou všechny stroje predikovány k aktuálnímu dni (v generovaných datech je 31.12.2024 poslední záznam). Dále obsahuje opět ***ID a název*** stroje. Následně zobrazuje predikovanou hodnotu RUL ve dvou časových jednotkách ve sloupcích **„*Predikované RUL (hodiny)*“** a **„*Predikované RUL (dny)*“**, aby měl uživatel představu konkrétní predikované hodnoty, ale i vyjádření ve dnech pro lepší dlouhodobou představu. Dále tabulka obsahuje sloupec **„*Úroveň rizika*“**, který je opět záměrně přidaný vizuální sloupec pro okamžitou identifikaci nízkého RUL a daných strojů. Dále je pak přidán sloupec **„*Doporučení*“**, které se snaží doporučit, jaká akce by měla být podniknuta na základě výše RUL (okamžitá kontrola / plán údržby / běžná kontrola).

Rizikové stroje s poruchou do 72 hodin jsou dle aktuální predikce opravdu ty s nejnižším RUL (první 3 horní stroje tabulky). Stroje s ID 1, ID 12 a ID 11 jsou stroje identifikované i první predikcí a stroj ID 1 s nejnižším RUL je i strojem s největší pravděpodobností poruchy. To mírně potvrzuje, že modely správně pracují na stejném souboru dat a vracejí spolu korespondující predikce, které si odpovídají.

V praxi pak při kontrole výsledné predikce dokáže zaměstnanec identifikovat stroje, které mají vysoké RUL, a měly by se tedy delší dobu vyhnout poruše, a které mají nízké RUL, tedy jejich životnost se zkracuje a měla by jim být věnována zvýšená pozornost. Pro tyto stroje pak mohou být včas plánována potřebná opatření proti očekávané poruše, kdy zmírníme její dopad, či se jí úplně vyhneme a stroj bude revitalizován (jeho RUL se zvýší).

Pro lepší přehlednost a interpretaci predikcí byl pak vytvořen ***graf predikovaného RUL dle strojů***, který je zobrazen na dalším obrázku. Ten jasně zobrazuje jednotlivé RUL pro každý stroj v hodinách a lze tak jasně porovnávat stav jednotlivých strojů. Obsahuje i hranice, které znázorňují dobu 72 hodin a 168 hodin pro lepší představu výše RUL v čase.



Obrázek 8-12: Graf predikovaného RUL dle strojů (zdroj: autor)

Posledním bodem je, že dle vyhodnocení modelu, měl model tendenci predikce RUL podhodnocovat, takže ***predikované RUL je nižší než skutečné hodnoty***. To je v tomto případě v pořádku, protože je lepší, když údržba začne na stroj reagovat dříve a pokusí se zlepšit stav stroje před poruchou, než kdyby byly predikce naopak nadhodnocené a porucha by přišla nečekaně dříve, než bylo predikováno.

Výsledky predikce tak jasně udávají výši životnosti do další poruchy pro každý jednotlivý stroj. Údržba tak může včas provést přípravy na přicházející poruchy strojů s nízkým RUL a včas se snažit snížit dopad poruchy, eliminovat poruchy a zvýšit RUL stroje do poruchy. To může pomoci snížit ztráty z nečekaných prostojů, správně nastavit využívání zdrojů na očekávané opravy a optimalizovat celkový chod oddělení údržby. Model tak poskytuje nejen odhad zbývající životnosti, ale i podporu rozhodování při plánování údržby.

8.9 Závěr návrhu a tvorby prediktivního řešení



- Pro úlohu ***predikce poruchy do 72 hodin*** byl nejlepší model ***laděný Random Forest***, který dle analýzy *feature importance* využíval pro své predikce relevantní vstupní proměnné, které jsou vhodné pro predikci poruch. Jeho výsledky byly zobrazeny v přehledné tabulce, která zobrazuje konkrétní stroj, predikci, zda je riziko poruchy do 72h, pravděpodobnost vzniku poruchy na základě učení a vizuální úroveň rizika, která umožňuje okamžitou identifikaci nejrizikovějších strojů. Oddělení údržby je z vytvořených predikcí schopné určit rizikové stroje a připravit se na možnou poruchu, či ji případně eliminovat.
- Pro úlohu ***typu poruchy*** byl nejlepším modelem ***laděný Histogram Gradient Boosting.*** Tato úloha byla vázána s úlohou první, kdy typ poruchy byl predikován jen pro stroje, které byly identifikovány jako rizikové do 72 hodin. Výsledky predikcí byly zobrazeny v přehledné tabulce, která opět zobrazovala rizikové stroje, jejich výsledky predikce z první úlohy, predikovaný typ poruchy pro stroj a pravděpodobnost vzniku tohoto typu poruchy. Dále byla popsána druhá tabulka, která zobrazuje pro každý stroj rozdělení pravděpodobnosti vzniku jednotlivých typů, jelikož ne vždy může vzniknout jednoznačná predikce a další typ poruchy může mít také vyšší pravděpodobnost. To je způsobeno možnou podobností vývoje vstupních proměnných u více typů poruchy. Údržba díky predikci může cíleněji hledat příčiny poruch na stroji a lépe se připravit na konkrétní typ poruchy.
- Pro třetí úlohu predikce ***RUL*** byl nejlepším modelem ***XGBRegressor***, který dosáhl nejnižších chyb predikcí. Výsledné predikce byly zobrazeny v tabulce, která obsahovala identifikaci strojů, predikované RUL v hodinách a ve dnech, úroveň rizika na základě výše RUL a doporučení dalšího postupu. Platí, že stroje s nejmenším RUL jsou stroje, které jsou i rizikové pro poruchu do 72h. Údržba se tak díky predikci dokáže včas připravit na poruchu stroje, než skutečně vznikne a podporovat své rozhodování a plánování údržby.
- Společně tyto predikce tvoří ucelený základ pro zavedení ***prediktivní údržby***, podporu rozhodování a optimalizaci procesů i fungování oddělení údržby a celé společnosti.

Zdroje

Bharati, P. V., Kumar, J. S. V. S., Anumula, S. K., Krishna, P. V., & Malla, S. (2025). *IoT and Predictive Maintenance in Industrial Engineering: A Data-Driven Approach* (arXiv:2511.04923). arXiv. https://doi.org/10.48550/arXiv.2511.04923

Cummins, L., Sommers, A., Ramezani, S. B., Mittal, S., Jabour, J., Seale, M., & Rahimi, S. (2024). *Explainable Predictive Maintenance: A Survey of Current Methods, Challenges and Opportunities* (arXiv:2401.07871). arXiv. https://doi.org/10.48550/arXiv.2401.07871

Česká agentura pro standardizaci. (2018). *ČSN EN 13306 (010660)*. https://www.technicke-normy-csn.cz/csn-en-13306-010660-158418.html

DIAGNOSTIKASTROJU.cz. (2025). *TRIBODIAGNOSTIKA - ESOS*. diagnostikastroju.cz. https://www.diagnostikastroju.cz/tribodiagnostika/

Guo, S. (2021, únor 15). Data-Driven Predictive Maintenance In a Nutshell. *Towards Data Science*. https://towardsdatascience.com/data-driven-predictive-maintenance-in-a-nutshell-ccc65a13b998/

Hill, R., & Berry, S. (2021). *Guide to Industrial Analytics: Solving Data Science Problems for Manufacturing and the Internet of things*. Springer.

Hussien Gomaa, A. (2024). Digital Twins for Improving Proactive Maintenance Management. *ResearchGate*, *2024*(9). https://doi.org/10.11648/j.es.20240903.12

IBM.com. (2023, květen 4). *What is Condition-based Maintenance?* | *IBM*. Ibm.Com. https://www.ibm.com/think/topics/condition-based-maintenance

Iung, B., Levrat, E., Marquez, A. C., & Erbe, H. (2007). E-MAINTENANCE: PRINCIPLES, REVIEW AND CONCEPTUAL FRAMEWORK. *IFAC Proceedings Volumes, 1st IFAC Conference on Cost Effective Automation in Networked Product Development and Manufacturing*, *40*(19), 18–29. https://doi.org/10.3182/20071002-MX-4-3906.00005

Kaabar, S. (2025, červenec 17). Histogram-Based Gradient Boosting Regression Tree for Time Series Forecasting [Substack newsletter]. *Engineering Alpha*. https://abouttrading.substack.com/p/histogram-based-gradient-boosting

Kasalová, A. (2016). *Proaktivní systém údržby strojů* [DIPLOMOVÁ PRÁCE, Vysoké učení technické v Brně - Fakulta strojního inženýrství: Ústav výrobních strojů, systémů a robotiky. Vedoucí práce: doc. Ing. Miloš Hammer, CSc.]. https://dspace.vut.cz/items/3416dc12-d14c-413c-b6ab-4938a8c438c0

Kasourova, E. (2026, červen 30). *RMSE Explained: A Guide to Regression Prediction Accuracy*. https://www.datacamp.com/tutorial/rmse

Kavlakoglu, E. (2021, říjen 20). *What Is Random Forest?* | *IBM*. IBM.Com. https://www.ibm.com/think/topics/random-forest

Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling* (Third edition). John Wiley & Sons, Inc.

Legát, V., Pexa, M., & Aleš, Z. (2018). *Zajištěnost údržby v koncepci Průmysl 4.0: Materiály ze 72. semináře Odborné skupiny pro spolehlivost, konaného dne 11.9.2018 v Praze* (1. vydání). Česká společnost pro jakost.

MaintainX.com. (2025, září 11). *Maintenance KPIs: The Most Important Metrics to Track in 2025*. Maintainx.Com. https://www.getmaintainx.com/blog/beginners-guide-maintenance-kpis

MBI tým. (2023). *IT a anatomie firmy (Prediktivní analytika)* (Roč. 2023). https://databusiness.cz/wp-content/uploads/page/2142/AF_II_05_05_Prediktivni_Analytika-2.pdf

Murel, J., & Kavlakoglu, E. (2024, leden 20). *What is a feature engineering? | IBM*. IBM.Com. https://www.ibm.com/think/topics/feature-engineering

NI.com. (2025). *What is Acoustic Monitoring? – Condition Monitoring*. Nationalintelligence.Com. https://www.ni.com/en/solutions/condition-monitoring/what-is-acoustic-monitoring.html

Novotný, 0, Basl, J., Gála, L., & Umlauf, M. (2024). *IT a anatomie výrobní firmy (Strojírenská analytika)* (Roč. 2024). https://databusiness.cz/wp-content/uploads/page/2375/AF_III_02_01_02_Stroj_Analytika-6.pdf

Novotný, 0, Buchalcevová, A., Bruckner, T., Staanovská, I., Šedivá, Z., & Pour, J. (2025). *IT a anatomie firmy (Podniková analytika)* (Roč. 2025). Professional Publishing. https://databusiness.cz/wp-content/uploads/page/2129/AF_II_05__Podnikova_Analytika.pdf

Novotný, O., Sládek, P., Nidl, M., & Chlapek, D. (2025). *IT a Anatomie výrobní firmy (Strojírenská firma)* (Roč. 2025). Professional Publishing. https://databusiness.cz/wp-content/uploads/page/2170/AF_III_02_01__Strojirenska_Firma-1.pdf

Pechová, K. (2025). *Optimalizace údržby výrobních strojů pomocí prediktivní analytiky a vizualizace strojových dat v reálném čase* [DIPLOMOVÁ PRÁCE, Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky, Vedoucí práce: doc. Ing. Jan Pour, CSc.]. https://vskp.vse.cz/96411_optimalizace-udrzby-vyrobnich-stroju-pomoci-prediktivni-analytiky-a-vizualizace-strojovych-dat-v-realnem-case?author=Pechov%C3%A1&type=Diplomov%C3%A1+pr%C3%A1ce&year=2025&page=1

Pošta, J., Dvořák, M., & Veselý, P. (2002). *Degradace strojních součástí: Monografie* (Vyd. 1). Česká zemědělská univerzita, Technická fakulta. https://degradace.tf.czu.cz/Mngr_ram.htm

Puškár, J. (2024). *Koncepce prediktivní a proaktivní údržby* [BAKALÁŘSKÁ PRÁCE, Česká zemědělská univerzita v Praze - Technická fakulta: Katedra jakosti a spolehlivosti strojů. Vedoucí práce: prof. Ing. Martin Pexa, Ph.D.]. https://theses.cz/id/akf8bi/

Pykes, K. (2025, leden 15). *Data Preprocessing: A Complete Guide with Python Examples*. DATACAMP.Com. https://www.datacamp.com/blog/data-preprocessing

Rada, S. (2022). *Návrh systému údržby v podniku zemědělské prvovýroby* [DIPLOMOVÁ PRÁCE, Česká zemědělská univerzita v Praze - Technická fakulta: Katedra jakosti a spolehlivosti strojů. Vedoucí práce: prof. Ing. Vladimír Jurča, CSc.]. Theses.cz. https://theses.cz/id/6h3lnn/?lang=cs

SAP.com. (2025). *Co je to prediktivní údržba? Kompletní přehled*. SAP. https://www.sap.com/cz/products/scm/apm/what-is-predictive-maintenance.html

scikit-learn.org. (2026a). *DummyRegressor*. Scikit-Learn. https://scikit-learn/stable/modules/generated/sklearn.dummy.DummyRegressor.html

scikit-learn.org. (2026b). *F1_score*. Scikit-Learn. https://scikit-learn/stable/modules/generated/sklearn.metrics.f1_score.html

scikit-learn.org. (2026c). *HistGradientBoostingRegressor*. Scikit-Learn. https://scikit-learn/stable/modules/generated/sklearn.ensemble.HistGradientBoostingRegressor.html

scikit-learn.org. (2026d). *RandomForestRegressor*. Scikit-Learn. https://scikit-learn/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html

Sharda, R., Delen, D., & Turban, E. (2020). *Analytics, data science, & artificial intelligence: Systems for decision support* (Eleventh edition). Pearson.

Siddiqui, L. (2026, červen 8). *Coefficient of Determination: What R-Squared Tells Us*. DATACAMP.Com. https://www.datacamp.com/tutorial/coefficient-of-determination

Stryker, C., Lee, F., Bergmann, D., & Scapicchio, M. (2024, květen 9). *What is XGBoost? | IBM*. IBM.Com. https://www.ibm.com/think/topics/xgboost

Štaffenová, K., Rakyta, M., Fusko, M., & Štepo, M. (2021). New trends and concepts of maintenance management in technical service. *Průmyslové Inženýrství 2021: Sborník Příspěvků*, 220–228. https://doi.org/10.24132/PI.2021.07927.220-228

towardsdatascience.com. (2025, listopad 26). Towards Data Science. https://towardsdatascience.com/

Waples, J. (2025, srpen 8). *Mean Absolute Error Explained: Measuring Model Accuracy*. DATACAMP.Com. https://www.datacamp.com/tutorial/mean-absolute-error

Weinpold, M. (2024, březen 21). *IT A ANATOMIE FIRMY* [Web Vysoké školy ekonomické vPraze]. MBI-AF.cz. https://databusiness.cz/it-a-anatomie-firmy/

zeiss.cz. (2025). *Analýza vibrací: Prediktivní údržba*. zeiss.cz. https://www.zeiss.cz/metrologie/explore/temata/analyza-vibraci.html