

# Dijagnostika kvarova kotrljajućih ležajeva temeljena na strojnom učenju

---

**Antunović, Mario**

**Master's thesis / Diplomski rad**

**2022**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Rijeka, Faculty of Engineering / Sveučilište u Rijeci, Tehnički fakultet**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:190:691846>

*Rights / Prava:* [Attribution 4.0 International](#)/[Imenovanje 4.0 međunarodna](#)

*Download date / Datum preuzimanja:* **2024-12-25**



*Repository / Repozitorij:*

[Repository of the University of Rijeka, Faculty of Engineering](#)



SVEUČILIŠTE U RIJECI

TEHNIČKI FAKULTET

Diplomski sveučilišni studij strojarstva

Diplomski rad

**DIJAGNOSTIKA KVAROVA KOTRLJAJUĆIH LEŽAJEVA  
TEMELJENA NA STROJNOM UČENJU**

Rijeka, srpanj 2022.

Mario Antunović

0069080830

SVEUČILIŠTE U RIJECI

**TEHNIČKI FAKULTET**

Diplomski sveučilišni studij strojarstva

Diplomski rad

**DIJAGNOSTIKA KVAROVA KOTRLJAJUĆIH LEŽAJEVA  
TEMELJENA NA STROJNOM UČENJU**

Mentor: Prof. dr. sc. Sanjin Braut

Rijeka, srpanj 2022.

Mario Antunović

0069080830

Rijeka, 16. ožujka 2022.

Zavod: **Zavod za tehničku mehaniku**  
Predmet: **Vibracije**  
Grana: **2.11.01 opće strojarstvo (konstrukcije)**

## ZADATAK ZA DIPLOMSKI RAD

Pristupnik: **Mario Antunović (0069080830)**  
Studij: **Diplomski sveučilišni studij strojarstva**  
Modul: **Računarska mehanika i inženjerstvo**

Zadatak: **Dijagnostika kvarova kotrljajućih ležajeva temeljena na strojnom učenju /  
Machine learning-based fault diagnosis of rolling bearings**

### Opis zadatka:

U radu je potrebno opisati vrste strategija održavanja, obraditi osnove mjerenja vibracija te opisati dijagnostiku kvarova rotacijskih strojeva analizom vibracija. Potrebno je napraviti pregled tehnika koje se koriste u dijagnostici kvarova rotacijskih strojeva, poglavito kvarova na valjnim ležajevima. Nadalje potrebno je dati uvod u strojno učenje. Korištenjem javno dostupnih baza podataka sa rezultatima eksperimentalnih ispitivanja oštećenja valjnih ležajeva, potrebno je predložiti metodu dijagnostike oštećenja na valjnim ležajevima temeljenu na strojnom učenju.

Rad mora biti napisan prema Uputama za pisanje diplomskih / završnih radova koje su objavljene na mrežnim stranicama studija.

*Mario Antunović*

Zadatak uručen pristupniku: 21. ožujka 2022.

Mentor:

*SB Braut*

Prof. dr. sc. Sanjin Braut

Predsjednik povjerenstva za  
diplomski ispit:

*K. Lenić*

Prof. dr. sc. Kristian Lenić

# IZJAVA

Izjavljujem da sam ovaj diplomski rad pod nazivom „Dijagnostika kvarova kotrljajućih ležajeva temeljena na strojnom učenju“ izradio samostalno uz stručnu pomoć i pod nadzorom mentora prof. dr. sc. Sanjina Brauta te uz podršku doc. dr. sc. Marija Lovrića u Ascalia d.o.o.



# Sadržaj

<b>Sadržaj.....</b>	<b>III</b>
<b>1. Uvod.....</b>	<b>1</b>
<b>2. Principi održavanja.....</b>	<b>2</b>
<b>2.1. Klasifikacija strategija održavanja.....</b>	<b>3</b>
2.1.1. Korektivno održavanje.....	5
2.1.2. Preventivno održavanje.....	6
2.1.3. Prediktivno održavanje.....	7
<b>2.2. Pregled tehnika u dijagnostici kvarova rotacijskih strojeva.....</b>	<b>10</b>
2.2.1. Temperaturna analiza.....	10
2.2.2. Analiza ulja.....	11
2.2.3. Mjerenje potrošnje energije.....	11
2.2.4. Vibrodijagnostika.....	11
2.2.5. Analiza akustične emisije.....	12
<b>2.3. Osnove mjerenja vibracija i analiza vibracija.....</b>	<b>13</b>
<b>3. Strojno učenje.....</b>	<b>17</b>
<b>3.1. Oblici strojnog učenja.....</b>	<b>19</b>
3.1.1. Nadzirano učenje.....	19
3.1.2. Nenadzirano učenje.....	21
3.1.3. Polunadzirano učenje.....	21
3.1.4. Podržano učenje.....	21
<b>3.2. Algoritmi strojnog učenja.....</b>	<b>22</b>
3.2.1. Logistička regresija.....	23
3.2.2. Algoritam nasumičnih šuma.....	24
3.2.3. Strojevi s potpornim vektorima.....	25
3.2.4. Analiza glavnih komponenata.....	27

<b>4. Dijagnostika oštećenja na valjnim ležajevima.....</b>	<b>28</b>
<b>4.1. Ulazni skup podataka.....</b>	<b>29</b>
<b>4.2. Obrada podataka i izvlačenje značajki .....</b>	<b>32</b>
4.2.1. Vremenska domena .....	33
4.2.2. Frekvencijska domena.....	35
4.2.3. Vremensko-frekvencijska domena.....	40
<b>4.3. Treniranje modela .....</b>	<b>43</b>
<b>5. Rezultati.....</b>	<b>45</b>
5.1. Vizualizacija značajki.....	46
5.2. Rezultati klasifikacije .....	49
<b>6. Zaključak.....</b>	<b>55</b>
<b>LITERATURA .....</b>	<b>56</b>
<b>POPIS SLIKA.....</b>	<b>62</b>
<b>POPIS TABLICA .....</b>	<b>64</b>
<b>SAŽETAK.....</b>	<b>65</b>
<b>ABSTRACT .....</b>	<b>66</b>

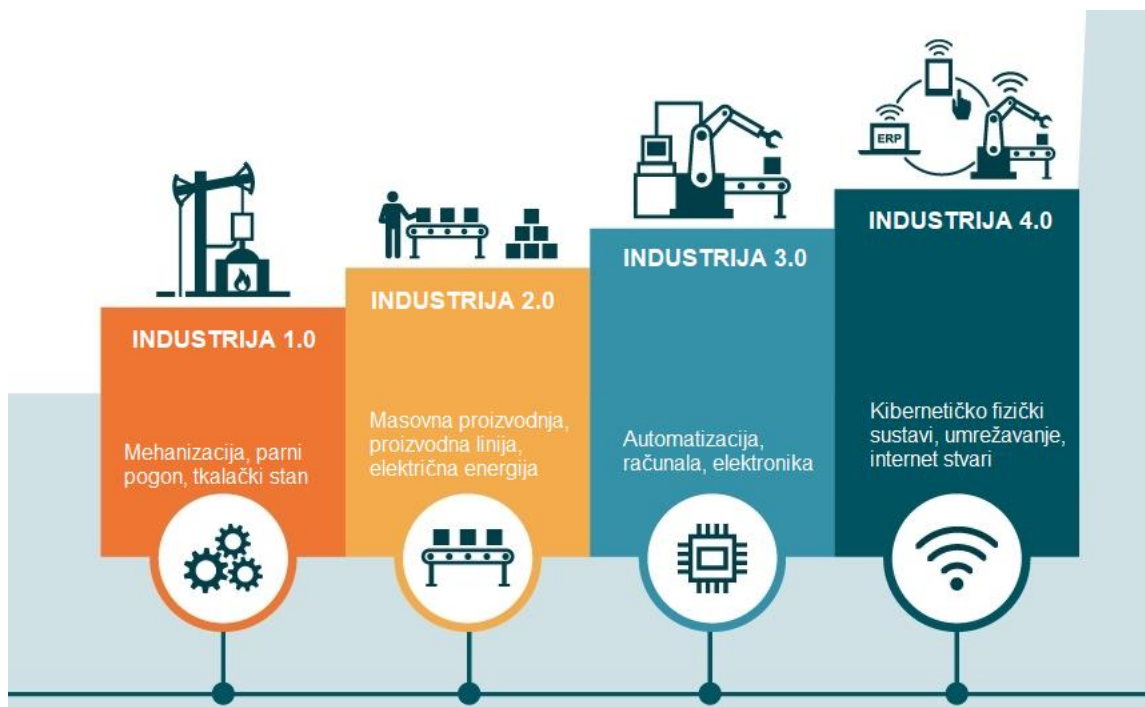


# 1. Uvod

Jedan od temelja uspjeha tvrtke u poslovnom svijetu jest mogućnost vođenja što učinkovitijih proizvodnih operacija. Brz tehnološki napredak tvrtkama u različitim industrijama uvelike olakšava modernizaciju, a samim time i povećanje učinkovitosti. Moderne digitalne tehnologije kao što su računalstvo u oblaku, strojno i duboko učenje, veliki podatci (engl. *Big Data*), Internet stvari (engl. *Internet of Things*) i slično, pokrenule su revoluciju potrošačkih tržišta i društvenih odnosa u protekla dva desetljeća, a samim time prisutna je i transformacija industrijskog sektora. Automatizacija i uvođenje pametnih digitalnih tehnologija u tradicionalne proizvodne procese doveli su do generacije inteligentnih proizvodnih metoda koje čine osnovu za stvaranje inteligentne tvornice, pri čemu se osim proizvodnog procesa razvija logistika i cjelokupni opskrbeni lanac. Inteligentna tvornica izdvaja se od ostalih zahvaljujući prilagodljivosti, optimalnom korištenju resursa te otpornosti na zastoje i smetnje. Robusnost samog proizvodnog procesa ključan je aspekt postojanosti i uspjeha tvrtke, u pogledu održavanja. Nepovoljno održavani resursi mogu uzrokovati nestabilnost te djelomično ili potpuno zaustaviti proizvodnju. Neispravan rad ili potpuni prekid rada strojeva mogu se pretvoriti u velik udarac na ekonomsku stabilnost tvrtke s obzirom na to da neplanirani zastoj rezultira mnogo većim gubitcima nego sam proces održavanja. U nastavku ovog rada bit će opisane strategije održavanja strojne opreme, pri čemu će naglasak biti na modernim tehnikama održavanja koje su obuhvaćene paradigmom inteligentne tvornice. Također, bit će dan pregled tehnika koje se koriste prilikom dijagnostike kvarova rotacijske opreme, poglavito valjnih ležajeva. Ovim je radom obuhvaćen koncept utvrđivanja zdravstvenog stanja i procjene rizika od kvara valjnih ležajeva temeljen na podacima koji su dobiveni eksperimentalnim ispitivanjem. Uvod u strojno učenje, uključuje treće poglavlje, nakon čega su detaljnije definirani korišteni algoritmi strojnog učenja u višerazrednom klasifikacijskom nadziranom učenju. U konačnici je uslijedila usporedba performansi najčešće korištenih metoda obrade signala i algoritama strojnog učenja za rješavanje ovakve vrste problema.

## 2. Principi održavanja

Razvijene industrije diljem svijeta trenutno prolaze kroz četvrtu industrijsku revoluciju (Slika 2.1.). Industrija 4.0 opće je prihvaćen izraz za integraciju između fizičkih i digitalnih sustava u proizvodnom kontekstu [1]. Iz organizacijske perspektive, pametna oprema i strojevi igraju bitnu ulogu u povećanju operativne učinkovitosti i djelotvornosti, a samim time i profitabilnosti. Navedeni principi se u velikoj mjeri zasnivaju na međusobnoj povezanosti, automatizaciji, strojnom učenju i podacima u realnom vremenu.



Slika 2.1. Razvoj industrije

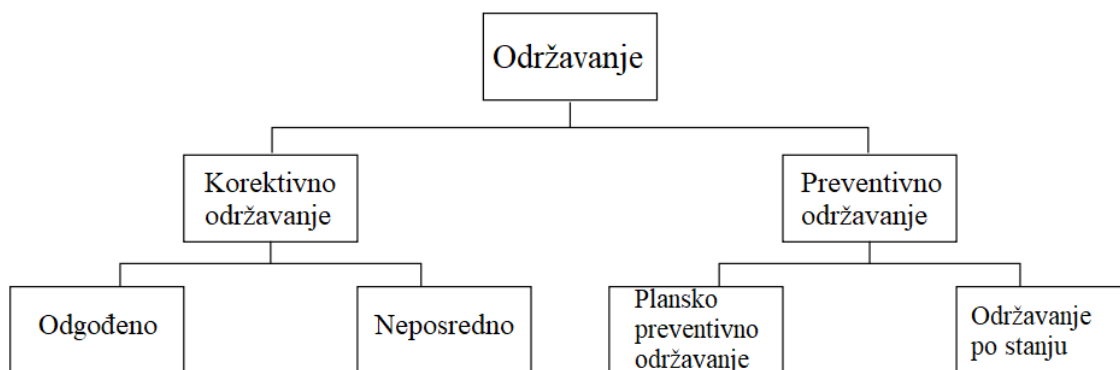
Uz ovu paradigmu razvile su se mnoge tehnologije koje se mogu grupirati u fizičke i digitalne tehnologije. Zahvaljujući naprednoj robotici, aditivnoj proizvodnji, sensorima i dronovima te umrežavanju i suvremenim komunikacijskim tehnologijama, odnosno razmjeni i analizi velikih količina podataka, proizvodni se procesi sada odvijaju u dva aspekta – u fizičkom svijetu, gdje se proizvod izrađuje te u digitalnoj sferi, gdje je proces proizvodnje optimiran priljevom vrijednih podataka. S pojavom industrije 4.0 koncept prognostičkog i zdravstvenog menadžmenta postao je nezaobilazna tendencija u okviru tzv. pametne proizvodnje, a uz to nudi i pouzdano rješenje za rukovanje zdravstvenim stanjem industrijske opreme [2].

## 2.1. Klasifikacija strategija održavanja

Zahtjevi za održavanjem opreme, čak i najrudimentarnijih alata, postoje od samog početka ljudske civilizacije. U to vrijeme ljudi bi opremu održavali samo kada bi postala neupotrebljiva. Ova vrsta održavanja naziva se korektivno održavanje. To je bila jedina strategija održavanja sve dok oko 1950. godine nije predložen novi koncept održavanja, pri kojem se vrši podmazivanje i zamjena određenih dijelova opreme prije nego što nastane kvar, s ciljem izbjegavanja zastoja i kašnjenja. Nova strategija nazvana je preventivno održavanje. Premda je vrlo korisna i pomaže u smanjenju zastoja, ova je strategija skupa alternativa s obzirom na to da se dijelovi mijenjaju na vremenskoj osnovi, odnosno mogli su trajati dulje.

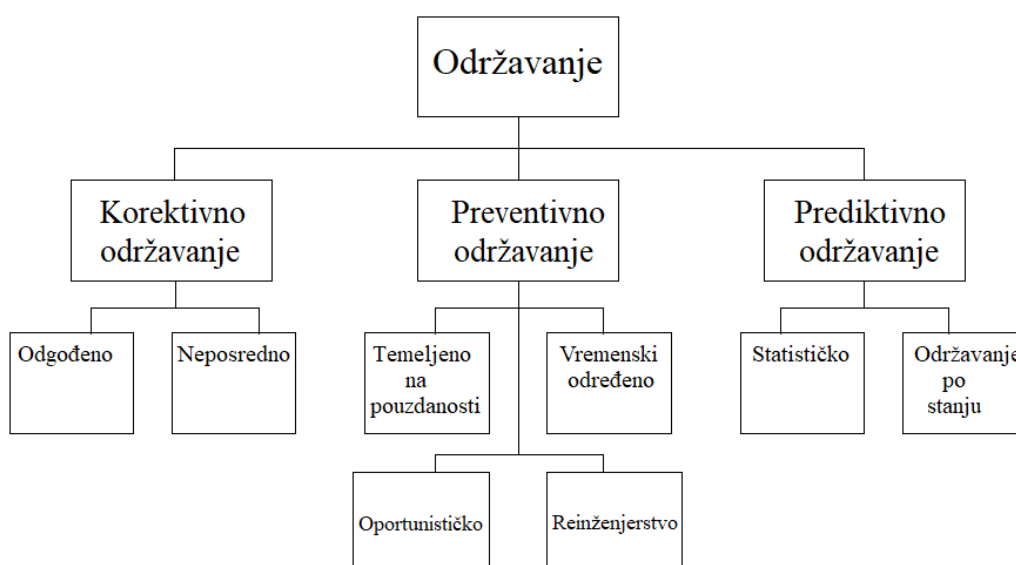
Istovremeno, s razvojem suvremene proizvodne tehnologije, proizvodni sustavi postaju sve složeniji i integriraniji. Različite komponente i podsustavi mogu raditi zajedno te ovisiti jedni o drugima za postizanje određenih ciljeva. Kada dođe do kvara, ključno je identificirati posljedice i uzroke što je prije moguće i poduzeti odgovarajuće mjere održavanja. Obično, kada dođe do kvara u sustavu, samo mali dio vremena zastoja može se potrošiti na otkrivanje temeljnog uzroka koji dovodi do kvara. Neočekivani neuspjeh može rezultirati nesrećom i financijskim gubitcima za tvrtku. Posljedično, sposobnost ranog predviđanja može spriječiti razvoj kvarova i u konačnici prerastanje u ozbiljne probleme. Posljednjih godina predložena je nova vrsta održavanja, nazvana prediktivno održavanje čija je ideja planirati održavanje prema dostupnim informacijama, koje mogu ukazati na trenutno stanje opreme, odnosno predvidjeti određene kvarove. Međutim, zbog tehnoloških ograničenja u posljednjih nekoliko desetljeća prisutne su poteškoće u prikupljanju i integraciji svih bitnih informacija. Danas, s trendom četvrte industrijske revolucije, uz izazove prediktivnog održavanja također je vidljiv i potencijal u novoj eri.

Postojeća istraživanja predstavila su različita tumačenja pojma „održavanje“ i njegove svrhe. Primarna uloga održavanja je smanjiti utjecaj kvara i proširiti dostupnost uz niže troškove te u konačnici poboljšati performanse [3]. Održavanje se odnosi na sve radnje koje pomažu u održavanju ispravnog stanja proizvodne opreme i strojeva [4]. Ovdje vidimo dva osnovna aspekta održavanja: korektivno održavanje i preventivno održavanje, koji se mogu podijeliti u nekoliko kategorija (Slika 2.2.).



Slika 2.2. Strategije održavanja

Najvažniji aspekt prediktivnog održavanja je korištenje metoda i modela za izradu prognoze daljnjeg razvoja stanja i preostalog vijeka trajanja [5]. Tradicionalno održavanje temeljeno na stanju preporučuje radnje održavanja na temelju informacija prikupljenih praćenjem stanja pri čemu je fokus na trenutnom stanju, u kontrastu s prediktivnim održavanjem čiji je najvažniji aspekt prognoza daljnjeg stanja. Održavanje se provodi nakon što analitički modeli utvrde da je prisutno dovoljno pokazatelja za točno predviđanje pojava kao što su propadanje komponenti ili određeni kvar. Korištenje podataka u stvarnom vremenu omogućuje operativnom timu da odredi prioritete i optimizira zakazivanje održavanja. Stoga je predstavljena nova klasifikacija strategija održavanja, prikazana na slici 2.3. koja odvaja prediktivno održavanje od preventivnog održavanja [6].



Slika 2.3. Moderna klasifikacija strategija održavanja (preuzeto iz [6])

### 2.1.1. Korektivno održavanje

Vrsta održavanja koja se provodi tek po pojavi kvara radi utvrđivanja i otklanjanja uzroka kvarova za neispravan sustav, a podrazumijeva sve radnje koje proizlaze zbog pojave kvara naziva se korektivno održavanje [7]. Korektivno održavanje prikladno je samo za nekritična područja kad su posljedice kvara neznatne, uz kratko vrijeme dijagnostike i popravka, te također ukoliko ne postoji visoki zahtjev za raspoloživošću opreme. Korištenje korektivnih metoda je razumno kada troškovi održavanja premašuju troškove popravka opreme, što podrazumijeva da zastoj stroja ne utječe na porast proizvodnih troškova [8]. Naime, u pogledu pametne tvornice sve ove preduvjete nije lako ispuniti. Strategiju korektivnog održavanja moguće je podijeliti u dvije klase:

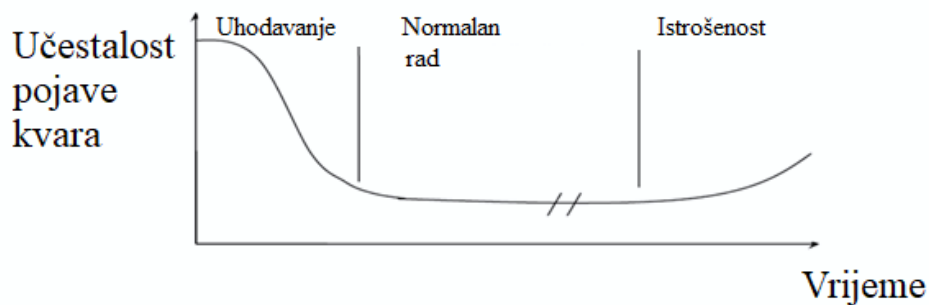
- neposredno korektivno održavanje
- odgođeno korektivno održavanje.

Neposredno korektivno održavanje će započeti odmah nakon što se otkrije kvar ili oštećenje. Kod odgođenog korektivnog održavanja održavanje se izvodi na planiran način, što znači da se otklanjanje kvara ne provodi nakon što dođe do kvara, već prema planu održavanja koji može smanjiti troškove ili vrijeme implementacije. Korektivno održavanje je održavanje uzrokovano kvarom, koje se poduzima nakon kvara ili kada je lociran očigledan kvar. Cilj korektivnog održavanja jest što je brže moguće vratiti stroj u stanje u kojem može obavljati traženu funkciju [9]. Nažalost, kao primitivno održavanje, korektivno održavanje ne uzima u obzir financijske gubitke uzrokovane nepredviđenim kvarovima. Ukoliko je ova strategija neprikladno primijenjena, pojavljuju se dodatni nedostaci korektivnog održavanja [9]:

- kvar komponente može proći nezamijećen, a posljedično voditi skupljem oštećenju
- potencijalna opasnost i havarije usred mehaničkog kvara
- potreban tim održavatelja na raspolaganju
- potrebno je osigurati rezervne dijelove te ih skladištiti kako bi se omogućio popravak u što kraćem roku.

### 2.1.2. Preventivno održavanje

Ova vrsta održavanja nastoji povećati pouzdanost i dostupnost opreme kroz minimiziranje broja kvarova i izbjegavanje zahtjeva za neplaniranim korektivnim održavanjem. Održavanje se planira prema karakteristikama pouzdanosti opreme, odnosno prema vremenu rada, starosti ili kalendarskom vremenu i provodi periodički kako bi se smanjila vjerojatnost kvara [7]. Međutim, zbog preventivnog djelovanja, do zamjene dijelova opreme može doći prerano. Primjer životnog vijeka proizvodne opreme vidljiv je iz krivulje kade (*engl. Bath tub curve*) koja prikazuje učestalost kvara opreme u ovisnosti o vremenu rada (Slika 2.4.). U prvoj fazi, tzv. fazi uhodavanja vidljiv je velik broj kvarova zbog pogrešaka u dizajnu, proizvodnji ili montaži. Druga faza opisuje koristan vijek trajanja opreme u kojoj se događaju slučajni kvarovi, koji mogu biti uzrokovani primjerice slučajnim vanjskim naprežanjima ili nepravilnim rukovanjem opremom. Posljednja faza opisuje kvarove koji nastaju zbog istrošenosti opreme. U ovoj fazi stopa kvarova raste jer starenje komponenti i istrošenost materijala ubrzava pojavu kvara.



Slika 2.4. Krivulja kade (preuzeto iz [10])

Bitno je naglasiti kako je preventivno održavanje prikladno samo za kvarove koji imaju jasnu karakteristiku trošenja, odnosno neprikladno je za slučajne kvarove, obzirom da ne može eliminirati njihovu pojavu. Osim toga, s povećanjem kvalitete i složenosti proizvoda, cijena čestih postupaka održavanja sve je viša. Preventivno održavanje povećava broj nepotrebnih postupaka održavanja, što može uvelike povećati troškove održavanja i uzrokovati slučajnu štetu na pojedinim komponentama opreme. Premda preventivno održavanje možda nije optimalna strategija održavanja, ono ipak ima nekoliko prednosti [9]:

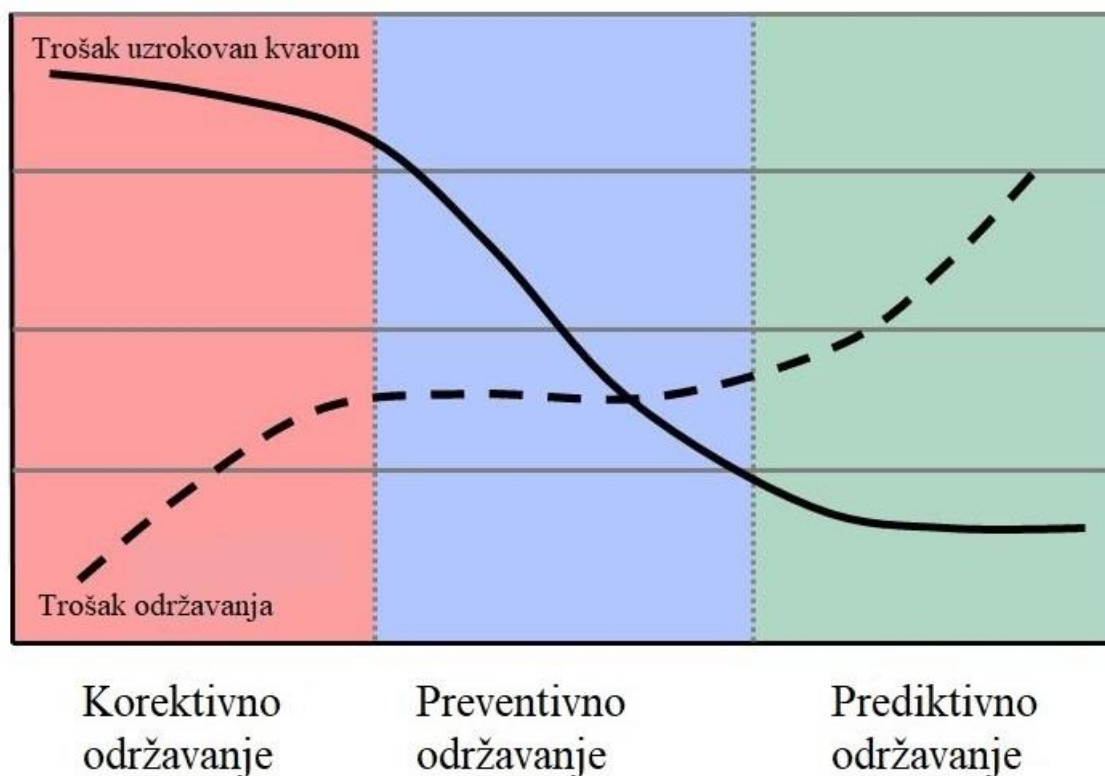
- ekonomičnost u troškovno-intenzivnim procesima
- fleksibilnost periodičnosti pregleda stanja opreme
- povećan životni vijek opreme
- smanjenje vjerojatnosti kvara opreme ili zastoja proizvodnih procesa.

### 2.1.3. Prediktivno održavanje

Prediktivno održavanje uključuje aktivnosti za provođenje odgovarajućih postupaka održavanja za maksimiziranje životnog vijeka opreme bez povećanja rizika od kvara, uzimajući u obzir predviđeni vijek trajanja opreme na temelju stvarnog stanja opreme [8]. U slučaju kada se ne radi o slučajnim kvarovima, proizvodna oprema svoju funkcionalnost ne gubi iznenada, već uslijed postupene degradacije i trošenja materijala. Tehničkom dijagnostikom, odnosno inteligentnom analizom podataka iz praćenja stanja i rada, prediktivno održavanje može smanjiti troškove smanjenjem broja nepotrebnih planiranih operacija održavanja temeljenih na vremenu rada ili starosti opreme [11]. U usporedbi sa strategijama korektivnog i preventivnog održavanja, prediktivno održavanje ima sljedeće prednosti [12]:

- prevencija neočekivanih katastrofalnih kvarova s potencijalno opasnim posljedicama
- oprema koja zahtijeva održavanje se gasi samo prije neminovnog kvara
- niski troškovi održavanja uslijed minimalnog broja postupaka održavanja
- eliminacija rizika od slučajnih oštećenja prilikom pregleda
- omogućuje pravovremenu nabavu rezervnih dijelova, smanjujući potrebe za skladištenjem silnih dijelova
- minimalne smetnje u postrojenju tijekom rada
- povećanje dostupnosti i pouzdanosti strojeva
- produljenje vijeka trajanja opreme
- fleksibilnost kod planiranja postupaka održavanja.

S druge strane, prediktivno održavanje slijede visoki početni troškovi ulaganja (Slika 2.5.) u dijagnostičku opremu i edukaciju operatera što manje razvijenim tvrtkama uvelike otežava implementaciju. Uz to postoji mogućnost pogrešne interpretacije podataka, što dovodi do lažnih zahtjeva za održavanjem.



*Slika 2.5. Odnos troškova i strategije održavanja*

Kao napredna strategija održavanja, prediktivno održavanje mjeri parametre o stanju opreme kako bi se izvršili odgovarajući zadatci za optimizaciju vijeka trajanja stroja i procesa bez povećanja rizika od kvara. Glavna zadaća održavanja po stanju je prikupljanje podataka s opremom koja je u pogonu, otkrivanje informacija iz prikupljenih podataka i utvrđivanje potencijalnih kvarova ili degradacije kroz povijesnu analizu slične opreme i znanja stečenog tijekom vremena. Ključna je točnost predviđanja kvarova ili degradacije.

Na temelju pristupa mjerenja simptoma kvarova, postoje dvije skupine prediktivnog održavanja [6]:

- prediktivno održavanje temeljeno na statističkim značajkama
- prediktivno održavanje temeljeno na trenutnom stanju.

Održavanje temeljeno na statističkim značajkama vrsta je tradicionalnog pristupa koja se temelji na primjeni statističke analize ili analize pouzdanosti kvara opreme. Pod prediktivnim održavanjem temeljenim na statističkim značajkama, cilj je postići minimalne ukupne troškove kroz fiksne statistički optimalne intervale održavanja za zamjenu ili remont opreme ili komponenti. Ključ ovog pristupa je primjena statističkih principa kako bi se odredilo kada će



se održavanje obavljati u budućnosti prema prikupljenim podacima o zastojsima opreme uzrokovanim kvarovima. Prediktivno održavanje temeljeno na stanju, s druge strane, uključuje primjenu praćenja stanja opreme i predviđanje kada, gdje i koje komponente mogu imati potencijalne kvarove prema trenutnom stanju, povijesnom stanju ili stanju u budućnosti. Ova vrsta održavanja uglavnom ovisi o uvjetima kontinuiranog ili periodičnog praćenja opreme kako bi se otkrili znakovi kvara i donijele odluke o održavanju. Glavni izazovi prediktivnog održavanja temeljenog na stanju u posljednjih nekoliko desetljeća je pristup potrebnim podacima i informacijama za praćenje stanja i predviđanje, te točnost detekcije kvarova ili predviđanja degradacije. Stoga je temeljni proces prediktivnog održavanja po stanju instalacija i/ili korištenje postojećih sustava za kontinuirano ili periodičko mjerenje i prikupljanje podataka, kao što su senzori i sustavi za prikupljanje i obradu podataka [9]. Srećom, s brzim razvojem tehnologije, a samim time i tehnika umjetne inteligencije u posljednjih nekoliko godina, napredak prediktivnog održavanja temeljenog na trenutnom stanju opreme uvelike je olakšan. Osobito u konceptu industrije 4.0, industrijska je oprema umrežena i autonomno razmjenjuje informacije, što znači da se mnogo industrijskih podataka može dobiti za potrebe održavanja po stanju. U praksi se koriste različite razine tehnologija te su prisutni različiti pristupi prilikom prikupljanja informacija za nadzor stanja, no svim pristupima je zajednička točka da se aktivnosti uglavnom izvode u radnom stanju opreme. Praćenje stanja moguće je izvoditi ručno ili automatski, a odnosi se na proces dobivanja, transformacije i interpretacije podataka o stanju opreme ili procesa [13]. Nadzor stanja opreme, ovisno o potrebi, može se provoditi kontinuirano ili periodički. U kontrastu s periodičkim mjerenjem, kontinuirani nadzor stanja je automatiziran, odnosno podatci se prikupljaju u stvarnom vremenu, smanjujući tako prostor za pogrešku. Uzorkovanje na kontinuirani način najčešće koristi senzore i instrumente povezane s opremom koji šalju signal u stvarnom vremenu, što se još naziva i online mjerenje [14]. Kontinuirano prikupljanje podataka omogućuje rješavanje potencijalnog problema čim je prisutno dovoljno pokazatelja za točno predviđanje kvara. Izvanmrežno (*engl. Offline*) mjerenje vrši se prema unaprijed definiranim vremenskim intervalima, a često ih izvode stručnjaci koji fizički izlaze na teren radi mjerenja [14]. Fizičke značajke opreme prate se pomoću različitih tipova senzora i instrumenata, mogu biti procesne varijable poput tlaka, protoka i temperature ili s druge strane - varijable opreme kao što su vibracije ili stanje sredstva podmazivanja. Nakon što su podatci prikupljeni, sljedeći korak je obrada i redukcija podataka do seta značajki koje mogu biti korištene kao pokazatelji zdravstvenog stanja.

## **2.2. Pregled tehnika u dijagnostici kvarova rotacijskih strojeva**

Trenutno je u upotrebi velik broj tehnika u dijagnostici kvarova rotacijskih strojeva čiji je razvoj tek u začetcima. Kako bi se olakšala dijagnostika i prognoza kvarova proizvodne opreme, u posljednjih nekoliko desetljeća intenzivno se istražuju područja kao što su:

- analiza temperature
- analiza ulja
- analiza potrošnje energije
- analiza vibracija
- analiza akustične emisije.

U praksi se navedene tehnike mogu koristiti za dijagnostiku i prognostiku jednakih tipova kvarova. Međutim, integracija svih ovih principa održavanja i dalje ostaje izazov do pojave industrije 4.0, koja kombinira prednosti optimizirane industrijske proizvodnje s mrežnim tehnologijama i značajno mijenja proizvodni proces, upravljanje održavanjem te strategije održavanja. U nastavku rada dan je pregled tehnika u dijagnostici i prognostici kvarova rotacijskih strojeva, koje je moguće implementirati tijekom provedbe strategije prediktivnog održavanja.

### **2.2.1. Temperaturna analiza**

Temperaturna analiza omogućuje otkrivanje prisutnosti potencijalnog kvara povezanog s promjenama temperature u opremi, stoga je jedna od najčešće korištenih metoda za planiranje postupaka održavanja. Svaki dio proizvodne opreme ima ograničenu radnu temperaturu koju najčešće nije problem izmjeriti, zato je temperaturna analiza pouzdan i jednostavan način za preliminarnu provjeru. Međutim, uslijed pojave problema temperatura raste postepeno, stoga ova vrsta analize nije dostatna za rano i precizno otkrivanje kvara. Osim toga, na izmjerenu temperaturu utječe i okolina pa se temperaturna analiza rijetko koristi kao primarni izvor informacija. Postoji nekoliko metoda praćenja temperature, no osnovna je podjela na pasivne beskontaktna i aktivne temperaturne senzore. Beskontaktnim mjerenjem moguće je dijagnosticirati probleme vezane za upravljačku elektroniku s obzirom na to da trošenje tih komponenata uzrokuje povećanje električnog otpora što u konačnici dovodi do povećanja emitirane topline. Kontaktnim se mjerenjem najčešće otkrivaju problemi povezani s podmazivanjem. Analiza temperature se primjenjuje na komponente kao što su ležajevi, rashladne tekućine, ulje za podmazivanje, motori i slično, pomoću instrumenata kao što je primjerice optički pirometar.

### 2.2.2. Analiza ulja

Istodobno s temperaturnom analizom ulja, moguće je provesti i analizu sastava ulja. Uslijed visokih temperatura, fizikalna svojstva ulja mogu se promijeniti, no isto tako sastav ulja mogu promijeniti krhotine koje se pojavljuju uslijed degradacije materijala na komponentama ležaja i zupčanika. Prema nekim istraživanjima, čak 80% problema s mjenjačkom kutijom moguće je pripisati ležajevima što dovodi do oštećenja zupčanika [15]. Automatski sustavi podmazivanja eliminiraju potrebu za čestim ručnim provjerama podmazivanja, osiguravajući siguran pristup podmazivanju stroja. U većini slučajeva, ulje cirkulira kroz zatvoreni sustav, pri čemu metalne krhotine ostaju u filteru. Količina i vrsta metalnih ostataka mogu ukazivati na zdravstveno stanje industrijske opreme.

### 2.2.3. Mjerenje potrošnje energije

Procjena potrošnje energije korisna je za ocjenjivanje energetske učinkovitosti. Karakteristike kao što su otpor, indukcija, pulsni i frekvencijski odziv koriste se prilikom utvrđivanja potencijalnih problema tako da se preventivne mjere mogu poduzeti prije pojave samog kvara. Obradni centri troše veliku količinu energije, a praćenje diferencijalne potrošnje električne energije može ukazati primjerice na potrošenost alata na motoru vretena.

### 2.2.4. Vibrodijagnostika

Najpoznatija i najčešće korištena tehnika za održavanje rotacijske opreme je vibrodijagnostika. Jedan od najranijih znakova degradacije rotacijske opreme je fluktuacija vibracija, koje mogu direktno ukazati na grešku. Vibracijski se signali mjere pomoću odabranih instrumenata, a izmjereni signali na složenim strojevima s mnogo komponenti, generiraju kombinaciju vibracija svake rotirajuće komponente. Rotacijski strojevi generiraju vibracije čak i dok su u ispravnom stanju. Primijenjena senzorska tehnologija odabire se uzimajući u obzir frekvencijski raspon i radne uvjete. Senzori za mjerenje položaja, brzine i ubrzanja koriste se za niske, srednje i visoke frekvencije, odnosno brzine vrtnje, respektivno [10]. U posljednjih nekoliko godina razvijeno je mnogo novih tipova senzora za praktično mjerenje signala vibracija. ISO standard [16] utvrđuje opće uvjete i postupke za mjerenje i ocjenu vibracija, ovisno o klasi rotirajućih strojeva. Ispitivanje prihvatljivosti vibracija, prema ovom standardu, provodi se pomoću sumarnih metrika (Slika 2.6.) kao što je korijen srednje kvadratne vrijednosti pomaka. Kriteriji ocjenjivanja temelje se na veličini vibracija, ali i o promjeni veličine vibracija. Ove su smjernice dizajnirane prvenstveno kako bi osigurale pouzdan i siguran rad stroja, bez potencijalnih opasnosti.

Granice	Klasa I	Klasa II	Klasa III	Klasa IV	Klasa V	Klasa VI	mm/s RMS
71	Red	Red	Red	Red	Red	Red	100
45							50
28	Red	Red	Red	Red	Orange	Orange	20
18							10
11	Red	Red	Orange	Orange	Yellow	Yellow	10
7,1							5
4,5	Orange	Orange	Yellow	Yellow	Green	Green	5
2,8							2
1,8	Yellow	Yellow	Green	Green	Green	Green	2
1,1							1
0,7	Green	Green	Green	Green	Green	Green	1
0,5							0,5
0,3	Green	Green	Green	Green	Green	Green	0,5
							0,3

Slika 2.6. ISO 20816-1:2016 (preuzeto iz [16])

### 2.2.5. Analiza akustične emisije

Fenomen akustične emisije temelji se na brzom oslobađanju energije u obliku tranzijentnih elastičnih valova unutar materijala putem dinamičkog procesa deformacije iz lokaliziranih izvora [17]. Plastična deformacija, napuknuće i rast pukotine, pucanje vlakana, trenje između površina, tipični su primjeri izvora ovog fenomena. Za razliku od vibracijske analize, analizom akustične emisije moguće je otkriti kvarove karakterizirane visokofrekventnim vibracijama u rasponu od 50 kHz do 1 MHz [18]. Veliki frekvencijski raspon i relativno visok omjer signala i šuma osnovne su prednosti ove tehnike. Nedostaci su veliki inicijalni troškovi te slabljenje signala tijekom širenja zbog čega je potrebno senzor smjestiti što bliže izvoru, što predstavlja praktično ograničenje u primjeni. Ova se metoda koristi za otkrivanje kvarova u komponentama kao što su ležajevi, mjenjačke kutije i vretena, no također i za otkrivanje problema s fluidima u cjevovodu. Najčešće mjereni parametri za dijagnozu su srednja kvadratna vrijednost amplituda, kurtozis, krest faktor i slično.

Provedena su mnoga istraživanja u kojima su uspoređene dijagnostičke i prognostičke sposobnosti prethodno navedenih metoda te je uočeno kako je samo akustična analiza osjetljivija u otkrivanju i praćenju grešaka od vibracijske analize [18].

### 2.3. Osnove mjerenja vibracija i analiza vibracija

Odabir strategije održavanja igra veliku ulogu u smanjenju troškova, minimiziranju učestalosti i trajanja zastoja, poboljšanju kvalitete, povećanju produktivnosti te osiguravanje pouzdanosti proizvodne opreme i kao rezultat postizanje zadanih ciljeva [7]. Postavljanje sustava održavanja uvelike ovisi o složenosti proizvodnih sustava, dostupnih resursa te o kritičnosti opreme. U pogledu održavanja, kritičnost opreme klasificira se u tri grupe [10]:

- kritična (najvažnija) oprema – visoka vrijednost, kvar dovodi do prekida proizvodnog procesa, potencijalna opasnost uslijed kvara
- važna oprema – kvar dovodi do prekida dijela procesa, potencijalna opasnost uslijed kvara, dostupnost opreme nije od ključne važnosti
- oprema za opću namjenu – nema utjecaja na sigurnost, nije kritična za daljnji rad postrojenja, niski troškovi održavanja.

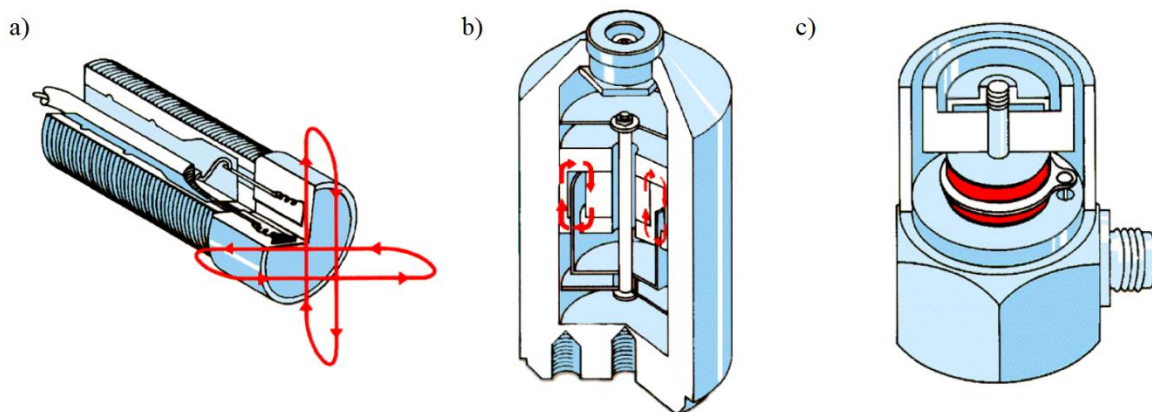
Obzirom na visok potencijal i perspektivnost, prediktivno održavanje po stanju pronalazi sve češću primjenu na mnogim područjima. Metodologija je izvedena iz fizikalnog fenomena: kada oprema počne otkazivati moguće je otkriti različite pokazatelje, poput promjene temperature, buke ili vibracija [19]. Danas, u doba brzog tehnološkog napretka, senzori su dovoljno napredni da učinkovito prikupe informaciju o degradaciji pojedinih komponenata. Kako identificirati, klasificirati i predvidjeti potencijalne kvarove uz nisku pogrešku, glavna su pitanja koja se pojavljuju uz paradigmu prediktivnog održavanja. Prikupljanje podataka je prvi korak u provedbi dijagnostike i prognoze i ključno je odabrati prikladnu tehniku te optimalan senzor za detekciju kvara. Proces prikupljanja podataka odnosi se na pretvorbu mehaničkog signala senzora u domenu ili više domena koje omogućavaju analizu stanja opreme [20]. U svrhu prikupljanja podataka razvili su se mnogi senzori kao što su mikrosenzori, ultrazvučni senzori, senzori vibracija i senzori akustične emisije. Odabir izvora podataka o stanju i odgovarajućih senzora važan je za učinkovitost procesa održavanja, uzimajući u obzir specifikacije senzora i isplativost implementacije. Zbog mnogih prednosti, kao što su pristupačnost i pouzdanost, analiza vibracija jedna je od najvažnijih dijagnostičkih metoda u industriji. Brz napredak tehnologije utjecao je i na razvoj senzora, što je prikupljanje podataka učinilo lakšim nego ikad prije. Analizom vibracija moguće utvrditi znakove potencijalnog kvara mjesecima prije same pojave kvara.

Tipične greške koje uzrokuju nepoželjne vibracije rotacijske opreme:

- debalans (neuravnoteženost) – rotirajuća komponenta sadrži područja veće mase koja stvaraju periodičku silu, a uzrokovane su greškama prilikom strojne obrade, nejednakom gustoćom materijala, nepravilnim balansiranjem, kavitacijom i dr.
- necentričnost – neusklađenost osi dviju rotirajućih komponenti stvaraju savojni moment prilikom rotacije, a najčešći uzroci su kriva montaža, neravni temelji, toplinsko širenje i dr.
- istrošene komponente kao što je primjerice kotrljajući ležaj ili oštećena mjenjačka kutija mogu dovesti do daljnjih oštećenja opreme te stvaraju potencijalnu opasnost
- nezategnutost mehaničkih dijelova – može uzrokovati vibracije čak i kod nerotirajuće opreme zbog nezategnutih vijaka ili korozije
- rezonancija – vibriranje rotacijske opreme će se snažno povećavati ukoliko periodička sila potiče stroj da vibrira brzinom kojom mu je najprirodnije.

Tri su vrste najčešće korištenih senzora za mjerenje vibracija (Slika 2.7.) ovisno o brzini vrtnje, odnosno o potrebnom frekvencijskom rasponu [10]:

- a) beskontaktni osjetnici pomaka
- b) kontaktni senzor brzine
- c) kontaktni senzor ubrzanja.

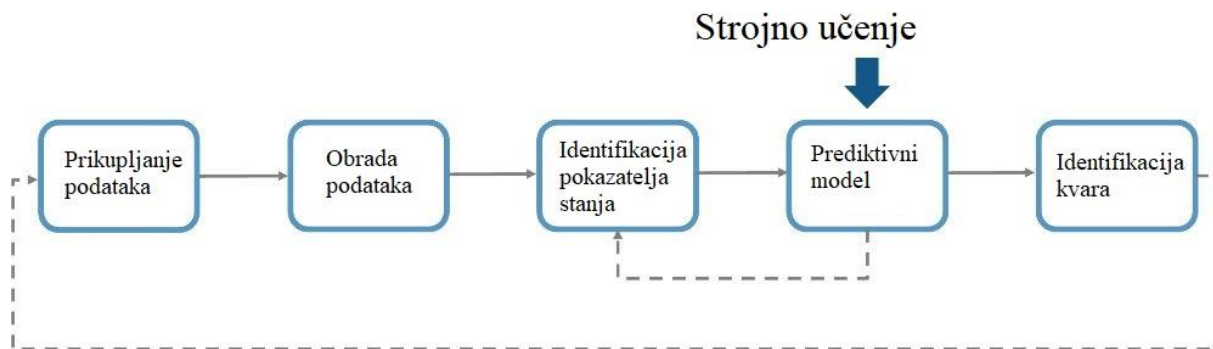


Slika 2.7. Instrumenti za mjerenje vibracija (preuzeto iz [10])

Beskontaktni senzor pomaka [mm] radi na principu vrtložnih struja (*engl. Eddy current*). Magnetsko polje proizvedeno izmjeničnom strujom oko zavojnice inducira vrtložne struje u bilo kojem električki vodljivom materijalu u blizini zavojnice. Ovakve je senzore potrebno kalibrirati, tj. prednamjestiti na traženi napon zračnog raspora te se pomak vodljivog materijala

registrira u obliku rasta napona. Nedostatak ovakvog mjernog instrumenta je u promjenjivim elektromagnetskim svojstvima osovine, koja može utjecati na točnost struje koja definira veličinu raspora. Kontaktni senzor vibracijske brzine [ $mm/s$ ] radi u skladu s elektrodinamičkim principom. Napon proporcionalan vibracijskoj brzini inducira se u zavojnici koja je ovješena u polju permanentnog magneta pomoću dvije helikoidne membranske opruge koje s oprugom tvore vibracijski sustav masa-opruga. Kontaktni senzor ubrzanja radi na piezoelektričnom principu koji se sastoji od mase i opruge te piezoelektričnog elementa smještenog na bazi senzora. Ukoliko se primjeni sila u smjeru polarizacije piezoelektričnog materijala, između površina poteći će naboj, odnosno električna struja. Razlika napona proporcionalna je narinutom opterećenju. Za rad ovakvog osjetnika nije potrebno napajanje obzirom da sam generira električnu struju, a također nema pokretnih dijelova koji bi bili podložni habanju. Kompaktnost i mala masa uz veliki frekvencijski spektar mjerenja osnovne su prednosti ove vrste senzora. Osnovni je nedostatak vrlo visoka impedancija izlaznog signala zbog čega je potrebno koristiti pretpojačalo koje pojačava signal. Integracijom signala ubrzanje se vrlo lako može pretvoriti u vibracijsku brzinu ili ako je potrebno u vibracijski pomak. Također je potrebno izvući informacije koje nose signali i prikupiti značajke koje opisuju potencijalnu grešku. Kako učinkovito i na odgovarajući način izdvojiti značajke iz prikupljenih podataka, identificirati i klasificirati kvarove s minimalnom pogreškom glavno je pitanje koje se veže uz koncept prognostičkog zdravstvenog menadžmenta. Mnoge učinkovite i korisne metode su predstavljene, istražene i primijenjene za izdvajanje značajki ili analizu signala u prediktivnom održavanju. Ove metode se uglavnom mogu podijeliti u tri tipa: vremenska domena, frekventna domena i vremensko-frekvencijska domena [21]. Budući da je izvlačenje značajki vrlo subjektivan problem, odabir metode ovisi o praktičnom problemu.

Središte paradigme prediktivnog održavanja je algoritam (Slika 2.8.) koji analizira prethodno prikupljene, a zatim i obrađene podatke u svrhu otkrivanja anomalija, dijagnosticiranja problema te predviđanja preostalog korisnog vijeka trajanja. Identifikacija pravih pokazatelja stanja je ključna za uspjeh korištenog algoritma i primijenjene strategije održavanja.



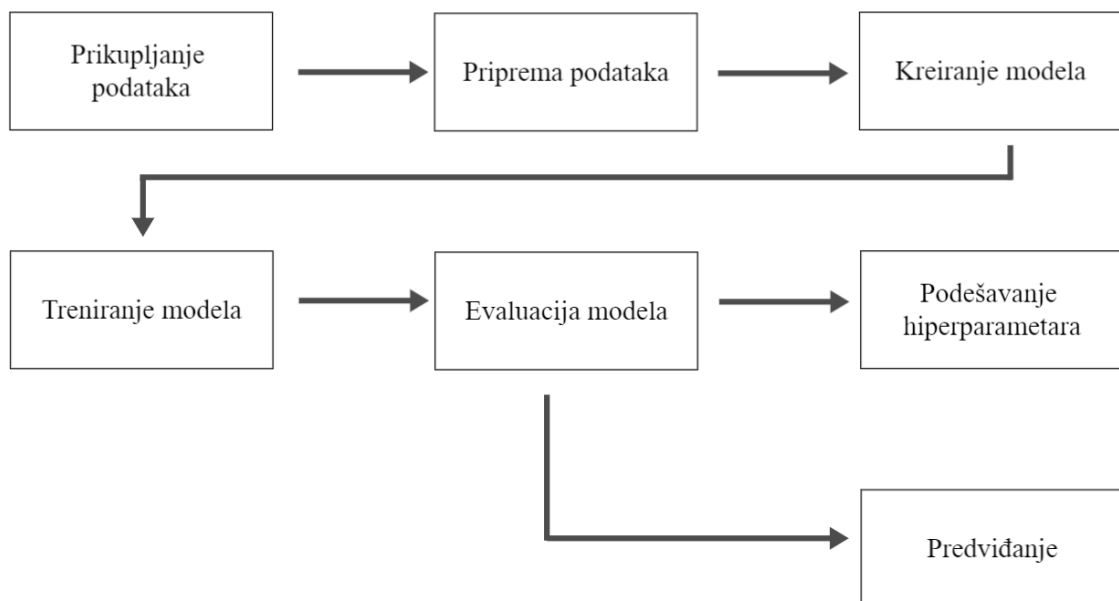
*Slika 2.8. Prediktivno održavanje temeljeno na stanju*



### 3. Strojno učenje

Unutar konteksta umjetne inteligencije (*engl. Artificial intelligence, AI*), strojno se učenje pokazuje kao jedan od najmoćnijih alata koji se primjenjuje za razvoj inteligentnih prediktivnih algoritama. Još važnije, potencijal strojnog učenja neprekidno raste. Tijekom posljednjih desetljeća, tehnike temeljene na strojnom učenju uspješno su primijenjene u područjima kao što je medicinska dijagnostika, prepoznavanje uzroka, računalni vid, zabava, financije i mnogim drugim [22]. Strojno učenje je grana umjetne inteligencije i računalne znanosti koja se bavi stvaranjem algoritama, temeljenog na prethodno prikupljenim podacima, kojima je moguće predvidjeti određene ishode [23]. Ovakav pristup nosi brojne prednosti, zbog sposobnosti rješavanja problema bez eksplicitnih naredbi, no također i zbog sposobnosti rukovanja multivarijantnim, visokodimenzionalnim setovima podataka. Visoko razvijeni algoritmi imaju sposobnost izdvajanja skrivenih odnosa među podacima koji potječu iz složenih, dinamičnih i kaotičnih okruženja.

Prilikom kreiranja modela (Slika 3.1.), ključan je korak točno definirati problem. Potrebno je odrediti što glavni cilj, odnosno što se predviđa te koje ulazne podatke je potrebno prikupiti kako bi se što bolje opisao problem te u konačnici dobio što bolji rezultat. Prikupljeni podatci moraju sadržavati dovoljno informacija kako bi bilo moguće naučiti odnos između ulaza i izlaza. Prije treniranja modela potrebno je transformirati podatke u oblik prihvatljiv modelu strojnog učenja.



Slika 3.1. Model strojnog učenja

Ključno je imati na umu da se strojno učenje može koristiti samo za predviđanje obrazaca koji su prisutni u podacima prilikom učenja. Slijedno tome, moguće je prepoznati samo ono što je već viđeno. Prilikom korištenja strojnog učenja, pretpostavka je kako će se budućnost ponašati kao prošlost, no to nije uvijek točno [24]. Matematičar Thomas Bayes utemeljio je bitne ideje u teoriji vjerojatnosti. Bayesova interpretacija vjerojatnosti je temelj strojnog učenja s obzirom na to da živimo u svijetu vjerojatnosti gdje svaki događaj ima neizvjesnost. Osnovna je ideja korištenje povijesnih podataka kako bi se kvantificirala neizvjesnost događaja [25]. Umjesto brojanja ponovljenih ispitivanja, ovim pristupom se vjerojatnosti temelje na dostupnim informacijama o ispitivanju. Budući da se proces donošenja odluka temelji na relevantnim značajkama, moguće je dodijeliti vjerojatnosti i rijetkim događajima. U praksi, kreiranje modela zahtjeva tri skupa podataka, odnosno potrebno je podijeliti izvorno prikupljene podatke na tri dijela [26]:

1. set za učenje (*engl. Train set*)
2. set za validaciju (*engl. Validation set*)
3. set za testiranje (*engl. Test set*).

Česta je praksa koristiti 70% izvornih podataka kao set za obuku te po 15% za preostala dva seta, no u doba velikih podataka kada se koriste veliki setovi podataka to ne mora biti pravilo. Set za učenje modela koristi se za kreiranje modela, dok se preostala dva skupa podataka ne smiju koristiti. Prilikom učenja modela, ne želimo da model radi dobro samo u predviđanju podataka koje je već vidio s obzirom da bi to bio trivijalan problem koji je moguće riješiti samo „pamćenjem“ ulaznih značajki. Set za validaciju koristi se za odabir algoritma te najboljih hiperparametara modela, a set za testiranje koristi se za procjenu točnosti prije puštanja u proizvodnju.

### 3.1. Oblici strojnog učenja

Postoji nekoliko oblika strojnog učenja; nadzirano, polunadzirano, nenadzirano te podržano učenje koje odabiremo prema definiranom problemu i željenim izlaznim podacima. Nadzirano i nenadzirano učenje (Slika 3.2.) su robusne metode koje se najčešće koriste. Polunadzirano i podržano učenje noviji su i složeniji oblici strojnog učenja, ali su pokazali impresivne rezultate u određenim problemima. Premda svaki od navedenih ima različite pristupe, svi slijede isti temeljni proces i teoriju [27].

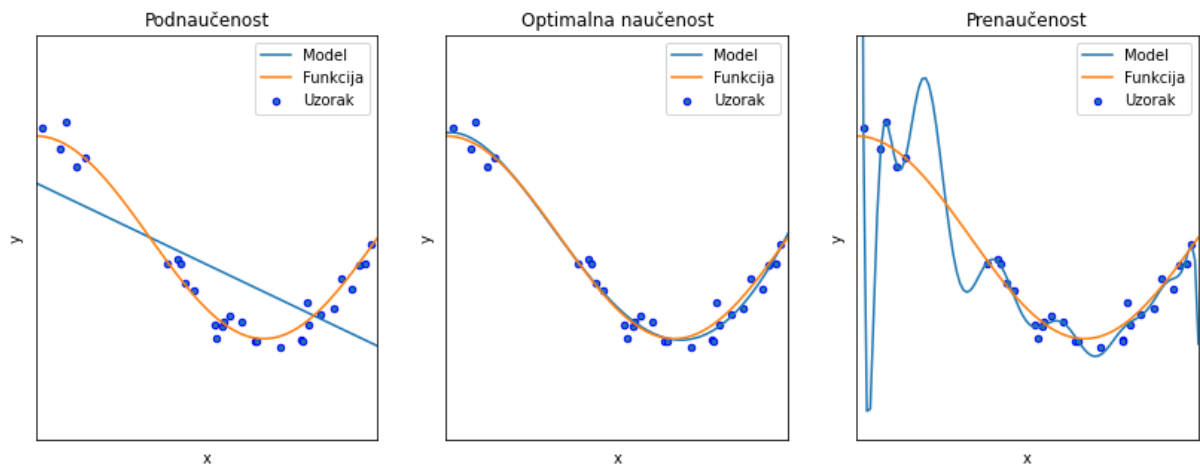


Slika 3.2. Oblici strojnog učenja

#### 3.1.1. Nadzirano učenje

Tijekom procesa učenja, temeljem podatkovnog seta za učenje, model uči kako se treba ponašati, odnosno koje izlazne vrijednosti treba dati za određene ulazne parametre, s obzirom na to da su izlazne vrijednosti poznate prije predviđanja. U nadziranom učenju skup podataka je zbirka označenih primjera  $\{(x_i, y_i)\}_{i=1}^N$ , pri čemu se svaki element  $x_i$  među  $N$  naziva vektorom značajki. Vektor značajki sadrži određenu vrijednost koja opisuje svaki primjer za  $j = 1, \dots, D$ , a naziva se značajka. Primjerice, ako svaki  $x = 1, \dots, N$ , opisuje čovjeka, tada će  $x^{(1)}$  predstavljati masu [kg],  $x^{(2)}$  predstavljat će primjerice visinu [m],  $x^{(N)}$  može predstavljati spol itd. Oznaka  $y_i$  može predstavljati element koji pripada određenom konačnom setu klasa  $\{1, 2, \dots, C\}$  kada se radi o klasifikacijskom problemu ili realan broj kada se radi o regresijskom problemu. Cilj algoritma nadziranog učenja je korištenje skupa podataka za izradu modela koji uzima vektor značajki  $x$  kao ulaz i daje informacije koje omogućuju dedukciju klasa za taj vektor značajki kada govorimo o klasifikacijskom problemu. Općenito, algoritam strojnog učenja može predvidjeti izlaz sustava s obzirom na nove ulazne vrijednosti.

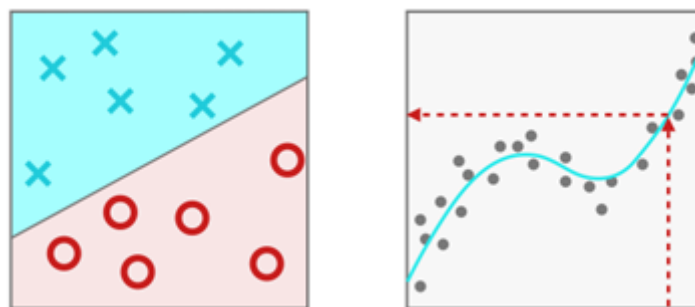
Potrebno je naglasiti kako se prilikom nadziranog učenja uvodi pristranost u učenje. Mogućnost prilagođavanja novim ulaznim podacima i točnog predviđanja ključan je generalizacijski dio strojnog učenja. Prilikom učenja je poželjno maksimizirati generalizaciju tako da nadzirani model definira stvarni odnos. Ako je model prenaučten (*engl. Over-trained*), uzrokuje prekomjerno prilagođavanje korištenim primjerima i model se ne može prilagoditi novim, nepoznatim ulaznim vrijednostima (Slika 3.3.).



Slika 3.3. Primjer nedovoljnog treniranja, dobrog treniranja i prekomjernog treniranja (preuzeto iz [26])

Nadziranim učenjem moguće je riješiti dvije vrste problema (Slika 3.4.):

- Regresijski problemi podrazumijevaju predviđanje ciljane vrijednosti koja je realan broj, odnosno kontinuirana varijabla.
- Klasifikacijski problemi podrazumijevaju dodjeljivanje klase (oznake) ulaznom setu značajki. Ukoliko postoje samo dvije kategorijske varijable radi se o binarnom klasifikacijskom problemu, a u slučaju kada ih je više radi se o višerazrednoj klasifikaciji.



Slika 3.4. Klasifikacija i regresija

### 3.1.2. Nenadzirano učenje

U nenadziranom učenju skup podataka je zbirka neoznačenih primjera  $\{(x_i)\}_{i=1}^N$ , pri čemu se svaki element  $x_i$  među  $N$  naziva vektorom značajki. Glavni cilj je otkriti unutarnje veze i obrasce među podacima. Ovakav pristup najčešće se koristi u svrhu grupiranja podataka ili za redukciju dimenzionalnosti kao transformaciju za rješavanje nekog praktičnog problema ili vizualizaciju podataka. Grupiranje (*engl. Clustering*) pokušava pronaći različite podgrupe unutar skupa podataka. Budući da je to učenje bez nadzora, odnosno ne postoji ograničenje skupom oznaka, potrebno je odabrati broj tzv. klastera. Smanjenje dimenzionalnosti ima za cilj pronaći najvažnije značajke kako bi se izvorni broj značajki smanjio. Nevažne značajke mogu zbuniti modele strojnog učenja te ih učiniti manje učinkovitim. Analiza glavnih komponenta je uobičajena tehnika kojom se identificiraju i koriste samo najvažnije značajke. Glavni nedostatak ovakvog pristupa je taj što nije moguće evaluirati performanse modela.

### 3.1.3. Polunadzirano učenje

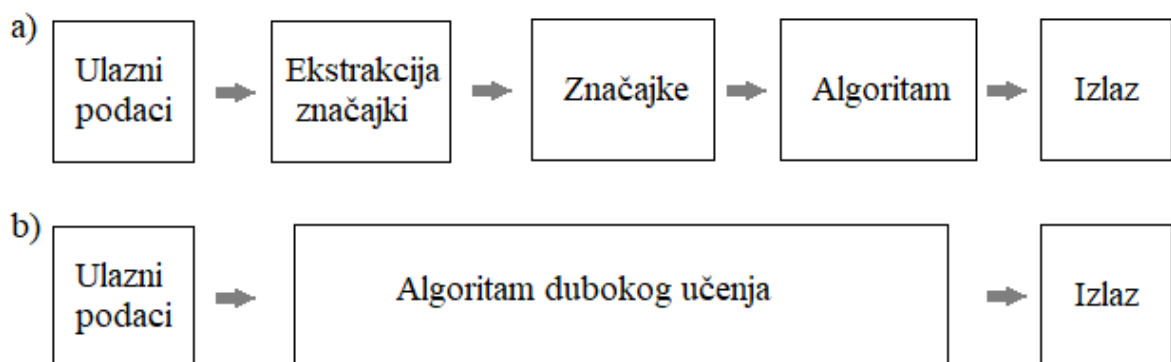
Korišteni skup podataka prilikom polunadziranog učenja sadrži i označene i neoznačene značajke. Često je količina neoznačenih primjera mnogo veća od broja označenih, s obzirom na to da je prikupljanje dovoljno pouzdano označenih podataka vrlo skupo. Cilj polunadziranog učenja isti je kao i cilj nenadziranog učenja.

### 3.1.4. Podržano učenje

Podržano učenje je mnogo složenije, ali u određenim problemima daje impresivne rezultate. Učenje je temeljeno na principu nagrađivanja, odnosno ovisno o njegovom ponašanju algoritam dobiva pozitivnu ili negativnu povratnu informaciju kojom se jača ili smanjuje neko ponašanje.

### 3.2. Algoritmi strojnog učenja

Ukoliko je rezultat predviđanja kategorička varijabla, radi se o klasifikacijskom problemu. Klasifikacijski algoritam na temelju ulaznih podataka određuje klasu primjera. Iz skupa podataka  $\{(x_i, y_i)\}_{i=1}^N$ , cilj klasifikacijskog algoritma strojnog učenja je naučiti koje vrijednosti značajki  $x_i$  pripadaju klasi  $y_i$ . Kada govorimo o utvrđivanju zdravstvenog stanja valjnih ležajeva, cilj je razdijeliti ispravne ležajeve od neispravnih. Tijekom procesa učenja u klasifikacijskom nadziranom učenju algoritam kreira opis svake od danih klasa, koristeći skup već klasificiranih primjera. Ukoliko je klasifikacijski problem jednostavan i linearan te ne koristi velik broj značajki, uputno je primijeniti jednostavnije algoritme kao što je primjerice logistička regresija (LogReg). S druge pak strane, kada se radi o složenijim problemima s većom količinom podataka, povećava se vjerojatnost slučajne korelacije među značajkama koje opisuju pojedine klase, stoga se jednostavne metode u takvim slučajevima pokazuju prilično inferiornima [28]. Napredne metode strojnog učenja koje su prikladnije za takvu vrstu problema su primjerice algoritam nasumičnih šuma (*engl. Random forest, RF*), strojevi s potpornim vektorima (*engl. Support vector machines, SVM*) te različitih inačica umjetnih neuronskih mreža (*engl. Neural networks, NN*). LogReg, RF i SVM su algoritmi strojnog učenja koji zahtijevaju izvlačenje značajki koje opisuju pojedine klase dok NN pripada podvrsti strojnog učenja pod nazivom duboko učenje koja ne zahtjeva izvlačenje značajki s obzirom na to da je taj proces automatiziran (Slika 3.5.).



Slika 3.5. Proces učenja za a) tradicionalne algoritme i b) algoritme dubokog učenja

Prilikom klasifikacije greške na valjnim ležajevima u ovom su radu korištena tri algoritma strojnog učenja: logistička regresija, nasumične šume te strojevi s potpornim vektorima. RF i NN smatraju se nelinearnim metodama, no oni mogu rješavati linearne i nelinearne klasifikacijske probleme.

### 3.2.1. Logistička regresija

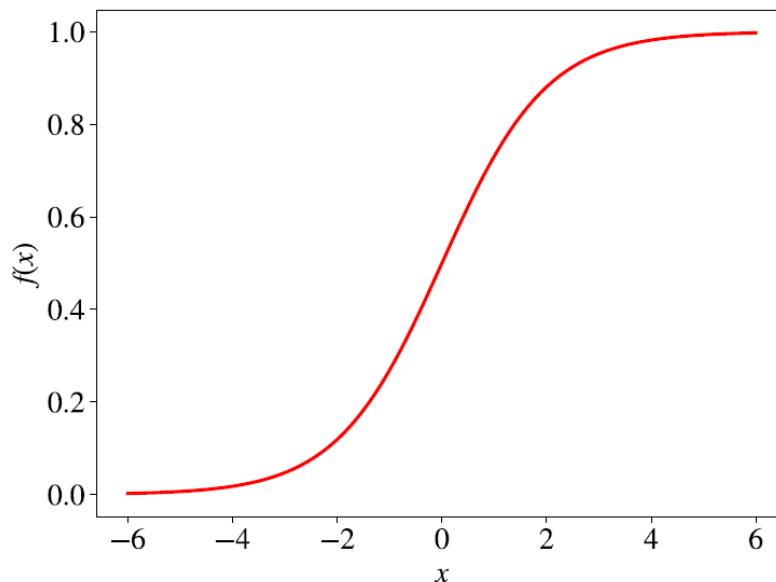
Premda se iz imena tako može naslutiti, logistička regresija nije regresijski algoritam. Ime potječe iz statistike zbog činjenice da je formulacija algoritma logističke regresije vrlo slična formulaciji linearne regresije [29]. Logistička regresija se pretežno primjenjuje kada su promatrane klase u višedimenzionalnom prostoru odijeljene linearno [28]. Najjednostavniji algoritam strojnog učenja je linearna regresija čiji se model može prikazati jednačbom:

$$f_{w,b}(x) = wx + b \quad (3.1)$$

pri čemu izlaznu vrijednost određuju vrijednosti značajke  $x_i$  te nj. uteženje parametrima  $w$  i  $b$ . Ukoliko se radi o binarnoj klasifikaciji, tj. skup podataka  $\{(x_i, y_i)\}_{i=1}^N$  sadrži dvije klase;  $y \in \{0, 1\}$ . Cilj logističke regresije je kreirati funkciju koja za bilo koji realan broj  $x_i$  daje vrijednost  $y_i$  koja može poprimiti vrijednost 0 ili 1.

Jedna od takvih funkcija (3.2) prikazana na slici 3.6., naziva se standardna logistička funkcija ili sigmoidna funkcija:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.2)$$



Slika 3.6. Standardna logistička funkcija

U ovom slučaju vrijednost koju daje model za ulazne vrijednost  $x$ , prosljeđuje se logističkoj funkciji koja poprima vrijednost 0 ukoliko je proslijeđena vrijednost bliža nuli, a u suprotnom vraća vrijednost 1. Logistička regresija je stoga probabilistički model, jer je izlazna

vrijednost vjerojatnost pripadanja nekoj od dvije klase za dane ulazne vrijednosti značajki  $x_i$ . U konačnici logistički model izgleda ovako:

$$f_{w,b}(x) = \frac{1}{1 + e^{-(wx+b)}} \quad (3.3)$$

Prilikom učenja modela logističke regresije potrebno je optimizirati hiperparametre modela  $w$  i  $b$  tako što se maksimizira vrijednost pripadanja značajki pojedinoj klasi prema izrazu (3.4) [26].

$$L_{w,b} = \prod_{i=1,\dots,N} f_{w,b}(x_i)^{y_i} (1 - f_{w,b}(x_i))^{(1-y_i)} \quad (3.4)$$

U praksi je prikladnije koristiti logaritam izraza 3.4 s obzirom na to da je potrebno osigurati da se maksimalna vrijednost pojavljuje u istoj točki kao i izvorna funkcija vjerojatnosti.

$$\text{Log}L_{w,b} = \ln(L_{w,b}(x)) \quad (3.5)$$

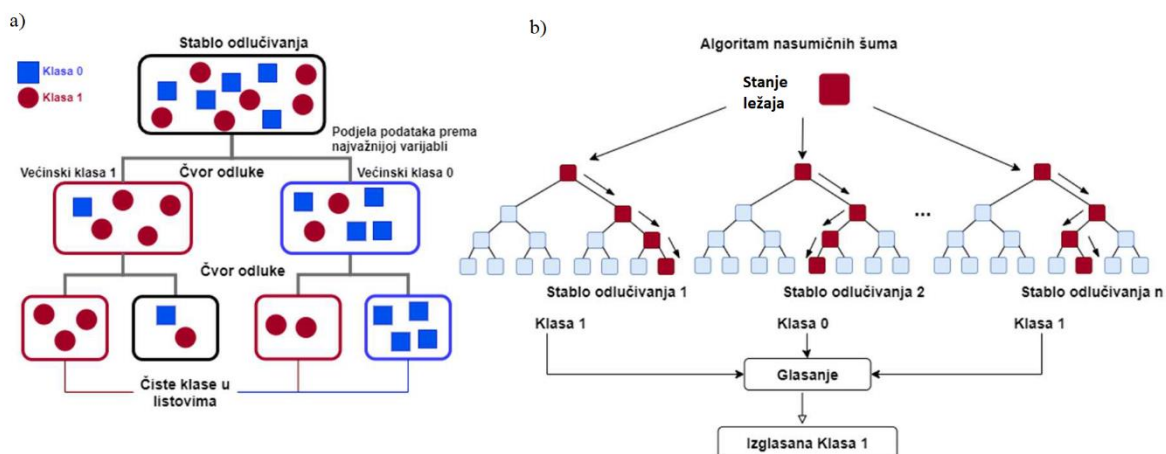
Optimizacijski postupak funkcije 3.5 zavisi o implementaciji algoritma, no najčešće se koristi tzv. gradijentni spust (*engl. Gradient descent*)[30]. Problem višerazredne logističke regresije rješava se pretvaranjem u binarnu klasifikaciju principom jedan protiv ostalih (*engl. One-vs-rest*), pri čemu se za svaku klasu konstruira i rješava binarni klasifikacijski problem.

### 3.2.2. Algoritam nasumičnih šuma

Algoritam nasumičnih šuma nadzirani je algoritam strojnog učenja koji se intenzivno koristi u klasifikacijske i regresijske svrhe zbog svoje fleksibilnosti i robusnosti. Temelji se na ideji koja, umjesto da pokušava naučiti jedan super-precizan model, trenira velik broj modela niske točnosti – stabala, koja su zatim sakupljena u šumu kako bi se dobio meta-model visoke točnosti [26]. RF model se obično sastoji od nekoliko stotina tzv. „slabih učenika“ (*engl. Weak learners*), odnosno stabala čije se odluke agregiraju kako bi se dobilo jedno konačno predviđanje.

Stablo odlučivanja (*engl. Decision tree*) je aciklički graf koji se koristi za donošenje odluka (Slika 3.7. a)). U svakom čvoru odluke ispituje se je li vrijednost specifične značajke ispod ili iznad određene granice, zatim se slijedi jedna od mogućih grana. Kada se dođe do tzv. listova, donosi se odluka o klasi kojoj primjer pripada.





Slika 3.7. Shematski prikaz binarne klasifikacije algoritmom nasumičnih šuma (preuzeto iz [28])

RF algoritam koristi modificirana stabla odlučivanja koja koriste različite podskupove značajki iz izvornog skupa. Algoritam tako osigurava da odluka svakog stabla ima nisku korelaciju s drugim stablima u svrhu maksimizacije uloge svakog stabla kao dijela glasačkog odbora. Ukoliko su jedna ili više značajki koje opisuju neku klasu vrlo jaki prediktori, te će značajke biti odabrane za čvor odluke u mnogim stablima što bi u konačnici rezultiralo mnogim koreliranim stablima te nasumične šume. Korelirani modeli ne mogu pomoći u poboljšanju točnosti predviđanja, štoviše korelirana stabla povećavaju vjerojatnost da će se loši modeli složiti prilikom odluke [31]. Prenaučenost se pojavljuje kada model pokušava objasniti male varijance u skupu podataka jer je set podataka za učenje samo mali uzorak u populaciji svih mogućih primjera fenomena koji pokušavamo modelirati. Za toliko dobru prediktivnost RF modela zaslužan je način uzorkovanja (tzv. *Bootstrap aggregation*), jer se korištenjem više podskupova izvornog skupa podataka smanjuje varijanca konačnog modela, što sprječava prenaučenos modela. Algoritam nasumičnih šuma pokazuje impresivne rezultate za velik broj problema, posebice nakon optimizacije hiperparametara modela kao što su broj stabala, dubina stabala, broj uključenih varijabli prilikom kreiranja modela te mnogi drugi.

### 3.2.3. Strojevi s potpornim vektorima

Strojevi s potpornim vektorima, kao i LogReg te RF je algoritam nadziranog strojnog učenja koji se danas koristi u mnogim područjima zbog svoje efikasnosti kada se radi o visokodimenzionalnim podacima [32]. Premda se ovaj algoritam koristi i za rješavanje regresijskih problema, u ovom će se slučaju opisati klasifikacijski problem. Iz ulaznog seta podataka svaki vektor značajki algoritam vidi kao točku i višedimenzionalnom prostoru. Naime, ovaj algoritam klasificira podatke pronalazeći hiperravninu koja najbolje razgraničava

klase iz danog seta podataka. Hiperravnina je generalizacija koncepta ravnine u višedimenzionalni prostor [33]. Kada govorimo o dvodimenzionalnom problemu s dvije klase, hiperravnina je linija koja dijeli dvije klase. Optimizacijski cilj prilikom treniranja modela je udaljenost između razdvajajuće hiperravnine i najbližeg uzorka podataka toj ravnini koji se još nazivaju i potporni vektori. Hiperravnina je zadana jednadžbom 3.6:

$$wx - b = 0 \quad (3.6)$$

Pri čemu je  $w$  vektor realnih vrijednosti jednake dimenzionalnosti kao i ulazni vektor  $x$ , a  $b$  je pomak od središta koordinatnog sustava.

Obzirom na to da govorimo o binarnom klasifikacijskom problemu, ulazni podatci moraju zadovoljavati jedan od dva uvjeta - 3.7 i 3.8:

$$wx_i - b \geq 1, \text{ za } y_i = +1, \text{ te} \quad (3.7)$$

$$wx_i - b \leq -1, \text{ za } y_i = -1 \quad (3.8)$$

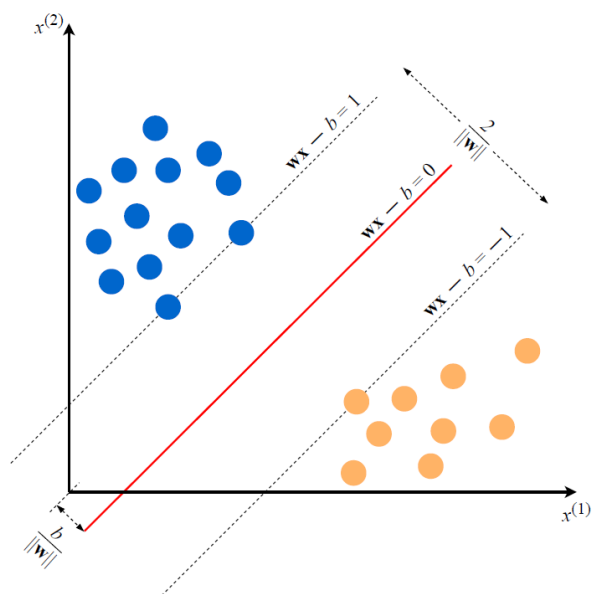
Cilj algoritma je optimizirati parametre kojima je zadana hiperravnina tako da su pozitivni i negativni uzorci razdvojeni najvećom marginom. Margina je najmanja udaljenost između dvije klase, te što je veća model će bolje generalizirati, odnosno bit će robusniji.

$$wx_i - b = 1 \quad (3.9)$$

$$wx_i - b = -1 \quad (3.10)$$

Ravnine opisane jednadžbama 3.9 i 3.10 definiraju dvije paralelne hiperravnine (Slika 3.8.), čija je međusobna udaljenost jednaka:

$$\rho = \frac{2}{\|w\|} \quad (3.11)$$



Slika 3.8. SVM model za dvodimenzionalne vektore značajki (preuzeto iz [26])

Iz izraza 3.11 vidljivo je kako se minimiziranjem euklidske norme  $\|w\|$  postiže maksimizacija širine margine  $\rho$ . Kada se radi o problemima koji nisu linearno razdvojni, SVM algoritam preslikava ulazne podatke u novi prostor više dimenzionalnosti, pri čemu vjerojatnost za linearno razdvajanje raste.

#### 3.2.4. Analiza glavnih komponenta

Analiza glavnih komponenta (*engl. Principal component analysis, PCA*) je jedan od najčešće korištenih nenadziranih algoritama strojnog učenja, a služi za smanjenje dimenzionalnosti seta podataka [34]. Temelji se na maksimizaciji varijance u manjedimenzionalnom projiciranom prostoru. Izvorni podatci  $(X \sim (n, p))$  umnožak su dviju matrica  $(T \sim (n, k))$  i  $(P \sim (p, k))$ :

$$X = TP^T + E \quad (3.12)$$

Pri tomuge je  $(E \sim (n, p))$  rezidualna matrica, a  $n$ ,  $p$  i  $k$  su uzorci, značajke te komponente, respektivno. Parametri su namještaju tako da glavne komponente obuhvate što više varijance iz izvornog seta podataka.

$$\{T, P\} = \operatorname{argmax}_{T, P} (\|X - TP^T\|_2^2) \quad (3.13)$$

Kombinacija vektora  $T$  i  $P$  nazivaju se glavne komponente, a korisne su za mnogo zadataka kao što je primjerice vizualizacija velikih setova podataka.

## 4. Dijagnostika oštećenja na valjnim ležajevima

Dijagnostika kvarova rotacijskih strojeva je od velikog praktičnog značaja u industriji kako bi se izbjegli ekonomski gubitci i povećala dostupnost strojeva [35]. Valjni ležajevi su precizno izrađeni dijelovi koji moraju podnijeti velika opterećenja, iznenadne udare i složena radna okruženja [36]. Tijekom rada ležaja, sila koja se prenosi s drugih dijelova na ležaj je nestabilna i mijenja se s vremenom, odnosno svaki ležaj proizvodi vibracije. Kontaktna sila ispravnog ležaja je vremenski kontinuirana u normalnom stanju. Međutim, ukoliko je ležaj istrošen, generirat će periodične impulzivne vibracijske signale koji se razlikuju od onih u normalnom radu. Ti se vibracijski signali mogu mjeriti sensorima i analizirati pomoću tehnika obrade signala za izdvajanje značajki. Za obradu vibracijskog signala razvile su se metode, koje izdvajaju značajke iz vremenske, frekvencijske i vremensko-frekvencijske domene, ali i kombinacijom fizikalnih i matematičkih zakonitosti što se naziva hibridno modeliranje [37], pomoću kojih se može točnije identificirati tip kvara. Trenutno, metode dijagnostike kvarova ležaja obično se dijele u tri kategorije:

- a. konvencionalne metode: sumarne metrike, analiza struje motora i dr.
- b. metode strojnog učenja: algoritam nasumičnih šuma, algoritam potpornih vektora i dr.
- c. metode dubokog učenja: umjetne neuronske mreže, konvolucijske neuronske mreže, rekurentne neuronske mreže i dr.

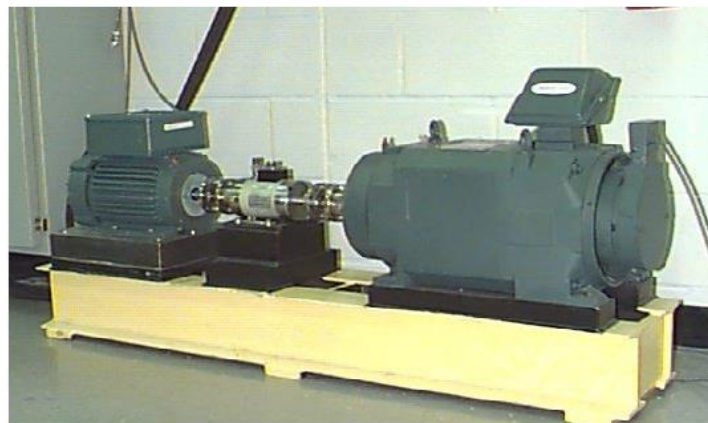
Među ovim metodama, duboko učenje pokazalo je najznačajniju učinkovitost uz visoku stopu točnosti klasifikacije. Međutim, većina istraživanja usredotočena je na poboljšanje mrežnih struktura i stope klasifikacije s manjim naglaskom na učinke obrade signala za pretvorbu izvornih u ulazne podatke jer za razliku od algoritama strojnog učenja, algoritmi dubokog učenja ne zahtijevaju izvlačenje značajki, odnosno algoritam sam stvara umjetne veze između ulaza i izlaza kako bi omogućio klasifikaciju. Ovaj diplomski rad obuhvaća obradu signala pomoću valične i Fourierove transformacije u svrhu izvlačenja značajki koje se mogu koristiti kao ulazni podatci za učenje algoritama strojnog učenja.

#### 4.1. Ulazni skup podataka

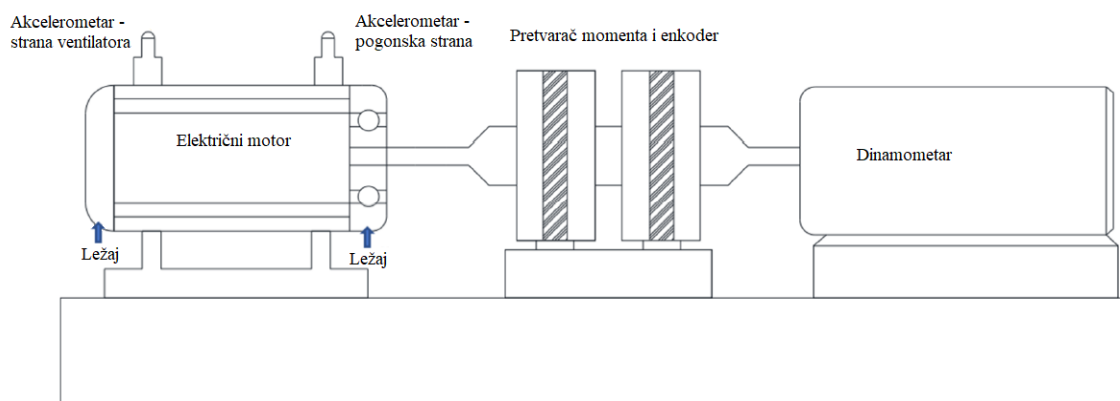
Podatci su temelj za sve algoritme umjetne inteligencije. Kvaliteta seta podataka je ključna za razvoj učinkovitog modela strojnog ili dubokog učenja za dijagnostiku kvara. S obzirom na to da je prirodna degradacija ležaja postupan proces koji može potrajati godinama, većina eksperimenata provodi se na ležajevima čije su greške izazvane umjetno ili s ubrzanim metodama ispitivanja vijeka trajanja [38]. Kao temeljni skup podataka, za provjeru učinkovitosti izvlačenja značajki i klasifikacije pojedinih algoritama, u ovom radu korišteni su vibracijski podatci dostupni na web stranici Sveučilišta Case Western Reserve (CWRU) [39] [40].

Na slici 4.1. prikazana je eksperimentalna stanica s koje su prikupljani podatci. Stanica se sastoji od motora, dva akcelerometra, pretvarača momenta i enkodera te dinamometra. Analiza je vršena na dva ležaja koja podupiru vratilo motora, od kojih je jedan na pogonskoj strani motora, a drugi se ležaj nalazi kod ventilatora.

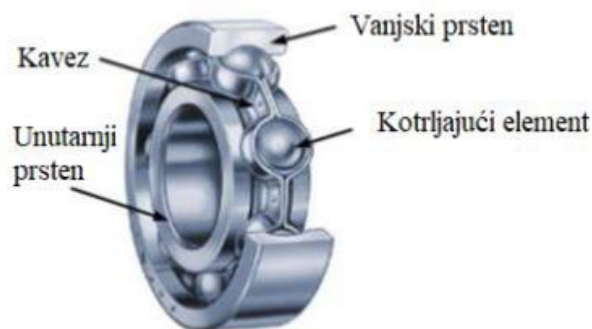
a)



b)

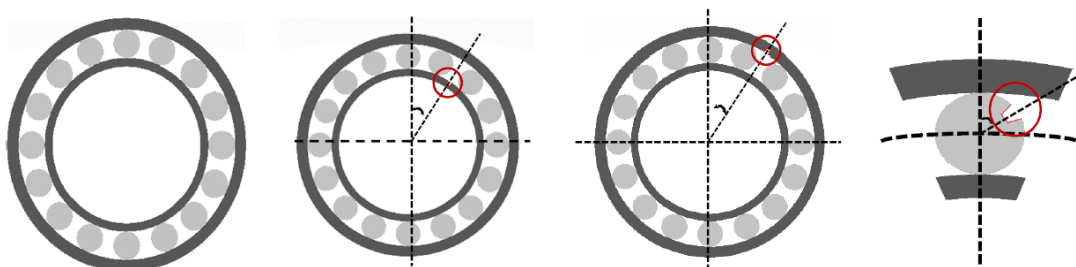


Slika 4.1. Eksperimentalna stanica



Slika 4.2. Kuglični ležaj

Uporabom elektroerozijske obrade (*engl. Electrical discharge machining, EDM*) na ispravne ležajeve nanese su greške u jednoj točki na jednom od 3 dijela kugličnog ležaja – unutarnji prsten, vanjski prsten i kotrljajući element (Slika 4.2. i 4.3.).



Slika 4.3. Vrste nanesenih grešaka (preuzeto iz [41])

Greške na ležajeve nanese su s tri veličine promjera: 0.007 inch (0.178 mm), 0.014 inch (0.356 mm) i 0.021 inch (0.533 mm). Uz ispravan ležaj na koji nije nanese greška to ukupno čini 10 različitih ležajeva (tablica 4.1.).

Tablica 4.1 Imena klasa prilikom klasifikacije

Dubina greške [inch]	Greška	Korišteno ime klase
0.007	Unutarnji prsten	IR_007
	Kotrljajući element	Ball_007
	Vanjski prsten	OR_007
0.014	Unutarnji prsten	IR_014
	Kotrljajući element	Ball_014
	Vanjski prsten	OR_014
0.021	Unutarnji prsten	IR_021
	Kotrljajući element	Ball_021
	Vanjski prsten	IR_021
0	Ispravan	Ispravan

Specifikacije ležaja na pogonskoj strani motora:

- naziv: 6205-2RS JEM
- proizvođač: SKF
- vrsta ležaja: kuglični ležaj s dubokim utorima
- promjer unutarnjeg prstena: 0.9843 inch (25 mm)
- promjer vanjskog prstena: 2.0472 inch (52 mm)
- debljina: 0.5906 inch (15 mm)
- promjer kuglice: 0.3126 inch (7.94 mm)
- promjer između kugličnih elemenata: 1.537 inch (39.04 mm).

Specifikacije ležaja na strani ventilatora:

- naziv: 6203-2RS JEM
- proizvođač: SKF
- vrsta ležaja: kuglični ležaj s dubokim utorima
- promjer unutarnjeg prstena: 0.6693 inch (17 mm)
- promjer vanjskog prstena: 1.5748 inch (40 mm)
- debljina: 0.4724 inch (12 mm)
- promjer kuglice: 0.2656 inch (6.75 mm)
- promjer između kugličnih elemenata: 1.122 inch (28.5 mm).

Za provjeru valjanosti metoda izvlačenja značajki i performansi modela strojnog učenja korišteni su podatci o vibracijama proizvedeni pri opterećenju od 1 konjske snage pri brzini vrtnje od 1772 okretaja po minuti (RPM). Brzina prikupljanja podataka iznosila je 48000 uzoraka po sekundi (*sample/s*). Za prikupljanje vibracijskih signala korišten je sustav za akviziciju podataka sa 16 kanala te prikupljeni podatci za jednu os, a zatim su obrađeni u softverskom paketu MATLAB.

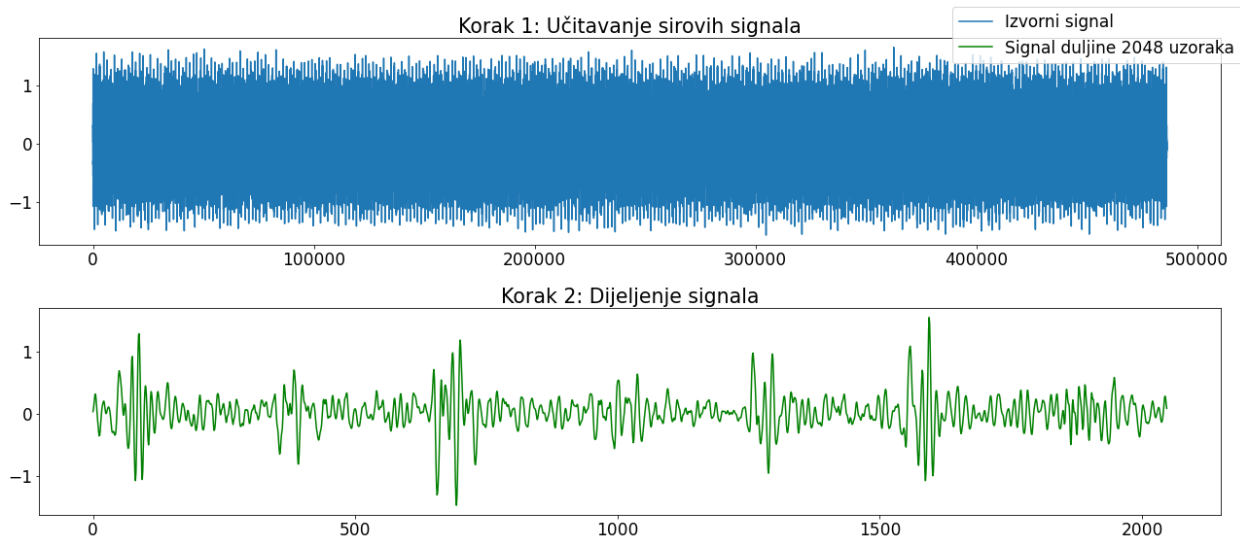
## 4.2. Obrada podataka i izvlačenje značajki

Za učitavanje, obradu podataka, smanjivanje dimenzionalnosti te treniranje algoritama strojnog učenja koristit će se Python - objektno orijentirani programski jezik opće namjene. Prilikom korištenja ovog programskog jezika moguće je izabrati između nekoliko integriranih razvojnih okruženja. Za učitavanje, oblikovanje i označavanje vibracijskog signala prema pripadajućem stanju ležaja koristit će se Jupyter Notebook [42]. Razvila ga je neprofitna tvrtka Project Jupyter s ciljem olakšavanja pristupa u područje znanosti o podacima (*engl. Data science*). Kao baza Jupyter Notebook-a uzet je IPython Notebook. Razvojno okruženje koje će se koristiti za izvlačenje značajki iz signala i treniranje modela strojnog učenja naziva se PyCharm [43], a često se koristi u računalnoj znanosti za služenje Pythonom. PyCharm je razvila češka tvrtka pod nazivom JetBrains 2010. godine. Python koristi biblioteke, odnosno module specijalizirane za različite zadatke. U nastavku su navedene biblioteke korištene u ovom radu:

- *Pandas* – učitavanje i manipulacija tabličnim podacima
- *Numpy* – zbirka matematičkih funkcija visoke razine namijenjena za rad na vektorima
- *Matplotlib i Seaborn* – kreiranje različitih grafičkih prikaza
- *Scikit-Learn* – niz alata za strojno učenje i statističko modeliranje uključujući klasifikaciju, regresiju, grupiranje, smanjivanje dimenzionalnosti i slično
- *SciPy* – omogućuje niz funkcija korištenih i znanstvenom i tehničkom računalstvu, kao što je optimizacija, statistika i obrada signala.

Prvi korak u kreiranju modela strojnog učenja je učitavanje ulaznih podataka koje će algoritam koristiti prilikom učenja, testiranja i validacije modela. Izvorni podatci podrazumijevaju snimljene signale duljine oko 480 000 uzoraka, za 10 stanja ležaja, a svaka datoteka sadrži podatke za ležaj na pogonskoj strani te na strani ventilatora elektromotora. Snimljeni signali podijelit će se na 230 signala duljine 2048 uzoraka (Slika 4.4.).





*Slika 4.4. Priprema podataka*

Kao rezultat oblikovanja sirovih podataka dobiva se 4600 signala duljine 2048 uzoraka. Osim segmentacije signala, izvršit će se i podjela na tri skupa podataka. Kao set za učenje modela koristi se 70% podataka, odnosno u ovom slučaju 3220 signala. Ostatak se dijeli na validacijski set te na testni set. Validacijski set stoga sadrži 690 signala, a služi za nepristranu procjenu točnosti modela, odnosno za odabir algoritma i najboljih hiperparametara modela. Set za testiranje jednak je validacijskom setu, a služi za provjeru točnosti modela, odnosno za ocjenu generalizacije. Nadalje je potrebno izvući značajke kako bi se dobili ulazni podatci koje algoritam može primiti. Kvalitetno izvlačenje značajki je ključ sustava za dijagnostiku kvarova rotacijskih strojeva, a najčešće se izvodi kombinacijom metoda u vremenskoj, frekvencijskoj i vremensko-frekvencijskoj domeni [44].

#### 4.2.1. Vremenska domena

Vremenske značajke definiraju različite statističke deskriptore izvučene iz vremenske domene [45]. U ovom će se radu koristiti niz statističkih značajki iz vremenske domene kao što su srednja vrijednost, standardna devijacija, srednji kvadrat i drugi, kako bi se okarakterizirali različiti signali. Osim toga, obzirom na nestacionarnu prirodu vibracijskog signala neispravnog ležaja koristit će se naprednije statističke značajke kao što su kurtosis i skewness.

U nastavku su navedene korištene značajke uz kratak opis.

1. Maksimum – maksimalna vrijednost amplitude signala

$$\max(x_i) \quad (4.1)$$

2. Minimum - minimalna vrijednost amplitude signala

$$\min(x_i) \quad (4.2)$$

3. Raspon – razlika maksimalne i minimalne amplitude signala

$$\max(x_i) - \min(x_i) \quad (4.3)$$

4. Srednja vrijednost – suma svih amplituda dijeljena sa brojem uzoraka

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (4.4)$$

5. Standardna devijacija – prosječno srednje kvadratno odstupanje amplituda od srednje amplitude

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}} \quad (4.5)$$

6. Korijen srednje kvadratne vrijednosti – kvadratni korijen zbroja kvadrata dijeljen sa brojem uzoraka

$$RMS = \sqrt{\frac{\sum x_i^2}{N}} \quad (4.6)$$

7. Faktor forme – omjer korijena srednje kvadratne vrijednosti i srednje vrijednosti

$$\frac{RMS}{\mu} \quad (4.7)$$

8. Kurtosis – kvantifikacija vršnih vrijednosti

$$\frac{\frac{\sum_{i=1}^N (x_i - \mu)^4}{N}}{\sigma^4} - 3 \quad (4.8)$$

9. Skewness – stupanj izobličenja signala od normalne distribucije

$$\frac{\frac{\sum_{i=1}^N (x_i - \mu)^3}{N}}{\sigma^3} \quad (4.9)$$

10. Crest faktor – mjera udara oštećenog elementa ležaja, odnosno omjer vršnih i efektivnih vrijednosti

$$\frac{\max(x_i)}{RMS} \quad (4.10)$$

Statističke značajke vremenske domene opisuju veličinu greške, no ne daju informaciju o točnom mjestu greške. Stoga će se u nastavku izvući značajke iz frekvencijske domene.

4.2.2. Frekvencijska domena

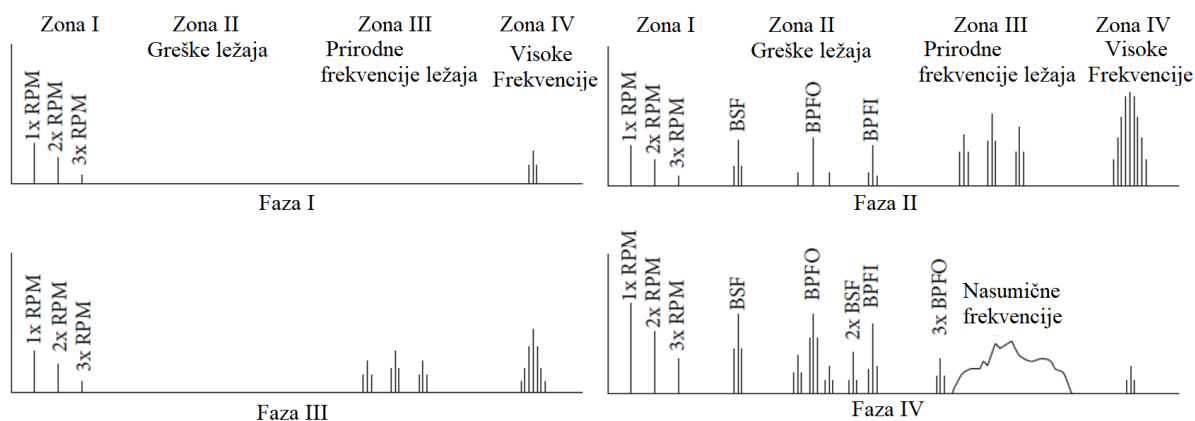
Spektralna analiza, odnosno analiza signala u frekvencijskoj domeni je često korišten pristup rješavanju ovakvih problema. Prilikom analize frekvencijskog spektra signali se iz vremenske domene pretvaraju u signale frekvencijske domene korištenjem Fourierove transformacije [46]. Kako bi izvlačenje značajki bilo što efikasnije, najprije će se na signale primijeniti filter, zatim će se pomoću Hilbertove transformacije izvući omotnica (engl. *Envelope*) signala, te se u konačnici dobiva frekvencijski spektar omotnice primjenom brze Fourierove transformacije.

Interakcija oštećenog elementa ležaja s drugim elementom ležaja stvara kratkotrajne impulse, koji dovode do povećanja amplitude u frekvencijskoj domeni. Na osnovi geometrijskih karakteristika ležaja moguće je izračunati teorijske karakteristične frekvencije na kojima će doći do povećanja amplitude što u konačnici može ukazati na problem s ležajem. U praksi, uslijed činjenice da se pri rotaciji valjnog ležaja osim kotrljanja pojavljuje i klizanje, greške u montaži i mjerenju, stoga se snimljene karakteristične frekvencije oštećenja ležaja mogu u određenoj mjeri razlikovati u odnosu na teoretske frekvencije.

Karakteristične frekvencije kvarova ležaja mogu se grupirati u četiri različite zone [47]:

- I. zona rotacije vratila
- II. zona greške ležaja
- III. zona prirodne rezonancije ležaja
- IV. visoka frekvencijska zona.

Prva zona sadrži harmonike uzrokovane rotacijom vratila. Ispravan ležaj će imati određenu energiju povezanom sa stanjem vratila, kao što je primjerice neuravnoteženost ili neusklađenost osi. Početna faza degradacije ležaja definirana je povećanom energijom u zoni I te zoni IV. Zona III noseći prirodne frekvencije imat će povećanu energiju ukoliko dođe do sljedeće faze degradacije ležaja. U posljednjoj fazi degradacije ležaja, degradacija se uočava i u zoni II. Sve četiri faze degradacije ležaja prikazane na slici 4.5.



Slika 4.5. Faze degradacije ležaja (preuzeto iz [47])

Oznake karakterističnih frekvencija oštećenja, prema skraćenicama izvedenim iz engleskog jezika prikazane su u nastavku:

$$BPFO = \frac{n f_r}{2} \left\{ 1 + \frac{d}{D} \cos \alpha \right\} \quad (4.11)$$

$$BPFI = \frac{n f_r}{2} \left\{ 1 - \frac{d}{D} \cos \alpha \right\} \quad (4.12)$$

$$FTF = \frac{f_r}{2} \left\{ 1 - \frac{d}{D} \cos \alpha \right\} \quad (4.13)$$

$$BSF = \frac{D f_r}{d} \left\{ 1 - \left( \frac{d}{D} \cos \alpha \right)^2 \right\} \quad (4.14)$$

Pri čemu je  $n$  = broj valjnih elemenata,  $f_r$  = frekvencije rotacije,  $D$  = promjer kaveza,  $d$  = promjer valjnog elementa,  $\alpha$  = kontaktni kut.

Najprije je na sirove podatke primijenjen band-pass filter kojim su isključene najniže i najviše frekvencije kako bi ostali podatci oko rezonantne frekvencije [48]. Nadalje je primijenjena Hilbertova transformacija. Analiza omotnice signala temeljena na Hilbertovoj transformaciji široko se koristi u dijagnostici kvarova ležajeva. Omotnica signala  $E(t)$  dobiva se uzimanjem amplitude analitičkog signala koji je zapis čija je realna komponenta izvorni signal  $s(t)$ , a imaginarna komponenta Hilbertova transformacija izvornog signala  $h(t)$ . Hilbertova transformacija definira se prema sljedećem izrazu:

$$h(t) = H \{s(t)\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau = \frac{1}{\pi t} s(t) \quad (4.15)$$

Stoga se  $h(t)$  dobiva kao konvolucija funkcije  $\left(\frac{1}{\pi t}\right)$  i izvornog signala  $s(t)$ , jer je Fourierova transformacija funkcije  $\left(\frac{1}{\pi t}\right)$ :

$$F\left(\frac{1}{\pi t}\right) = \begin{cases} -j, & \text{ako je } f > 0 \\ j, & \text{ako je } f < 0 \end{cases} \quad (4.16)$$

Izvorni signal i njegova Hilbertova transformacija tada čine analitički signal:

$$z(t) = s(t) + jh(t) \quad (4.17)$$

Omotnica signala  $E(t)$  dobiva se kao:

$$E(t) = |z(t)| = |s(t) + jh(t)| = \sqrt{h^2(t) + h^2(t)} \quad (4.18)$$

Tada se može dobiti spektar omotnice signala primjenom Fourierove transformacije na dobivenom signalu omotnice. Spektralna analiza korištenjem Fourierove transformacije koristi se za dekonstrukciju signala na njegove pojedinačne komponente sinusne funkcije. S obzirom na to da imamo diskretan skup brojeva (dobiven uzorkovanjem funkcije) koristit će se diskretna Fourierova transformacija (DFT). Budući da svaki od signala sadrži 2048 uzoraka,

izračunavanje konvencionalnog DFT-a bilo bi računski preskupo, stoga će se koristiti optimizirani algoritam, odnosno Brza Fourierova transformacija koja služi za brže i učinkovitije izračunavanje DFT-a. Brza Fourierova transformacija  $X_k$  izračunava diskretnu transformaciju signala u frekvencijsku domenu te njen inverz, odnosno koristi se za pretvaranje digitalnog signala duljine  $N$  iz vremenske u frekvencijsku domenu prema izrazu:

$$X_k = \sum_{n=0}^{N-1} E(t) e^{-i2\pi nk/N} \quad (4.19)$$

Nakon transformacije signala u frekvencijsku domenu moguće je prijeći na izvlačenje značajki iz frekvencijske domene. Najprije su zapisane vrijednosti maksimalnih amplituda u područjima teoretskih karakterističnih frekvencija (4.11 – 4.14) na način da se za svaku traži maksimalna amplituda u rasponu od  $\pm 30$  Hz te se tako dobivaju četiri značajke za svaki signal. Osim prethodno navedenih značajki na temelju signala u frekvencijskoj domeni izračunate su statističke značajke navedene u nastavku.

1. Maksimum – maksimalna vrijednost amplitude signala

$$\max(X_k) \quad (4.20)$$

2. Minimum - minimalna vrijednost amplitude signala

$$\min(X_k) \quad (4.21)$$

3. Raspon – razlika maksimalne i minimalne amplitude signala

$$\max(X_k) - \min(X_k) \quad (4.22)$$

4. Srednja vrijednost – suma svih amplituda dijeljena sa brojem uzoraka

$$\mu = \frac{1}{N} \sum_{i=1}^N X_k \quad (4.23)$$

5. Standardna devijacija – prosječno srednje kvadratno odstupanje amplituda od srednje amplitude

$$\sigma = \sqrt{\frac{\sum (X_k - \mu)^2}{N}} \quad (4.24)$$

6. Korijen srednje kvadratne vrijednosti – kvadratni korijen zbroja kvadrata dijeljen sa brojem uzoraka

$$RMS = \sqrt{\frac{\sum X_k^2}{N}} \quad (4.25)$$

7. Faktor forme – omjer korijena srednje kvadratne vrijednosti i srednje vrijednosti

$$\frac{RMS}{\mu} \quad (4.26)$$

8. Kurtosis – kvantifikacija vršnih vrijednosti

$$\frac{\frac{\sum_{i=1}^N (X_k - \mu)^4}{N}}{\sigma^4} - 3 \quad (4.27)$$

9. Skewness – stupanj izobličenja signala od normalne distribucije

$$\frac{\frac{\sum_{i=1}^N (X_k - \mu)^3}{N}}{\sigma^3} \quad (4.28)$$

10. Crest faktor – mjera udara oštećenog elementa ležaja, odnosno omjer vršnih i efektivnih vrijednosti

$$\frac{\max (X_k)}{RMS} \quad (4.29)$$

11. Spektralno središte – frekvencijski pojas u kojem je koncentrirana većina energije

$$\frac{\sum_{i=2}^N X'_k X_k}{2\pi \sum_{i=1}^N X_k^2} \quad (4.30)$$

12. Korijen srednje kvadratne frekvencije – ukupna razina energije u danom rasponu frekvencija

$$\sqrt{\frac{\sum_{i=2}^N (X'_k)^2}{4\pi^2 \sum_{i=1}^N X_k^2}} \quad (4.31)$$

13. Korijen varijance frekvencija - korijen razlike srednje kvadratne frekvencije i kvadrata spektralnog središta

$$\sqrt{\frac{\sum_{i=2}^N (X'_k)^2}{4\pi^2 \sum_{i=1}^N X_k^2} - \left(\frac{\sum_{i=2}^N X'_k X_k}{2\pi \sum_{i=1}^N X_k^2}\right)^2} \quad (4.32)$$

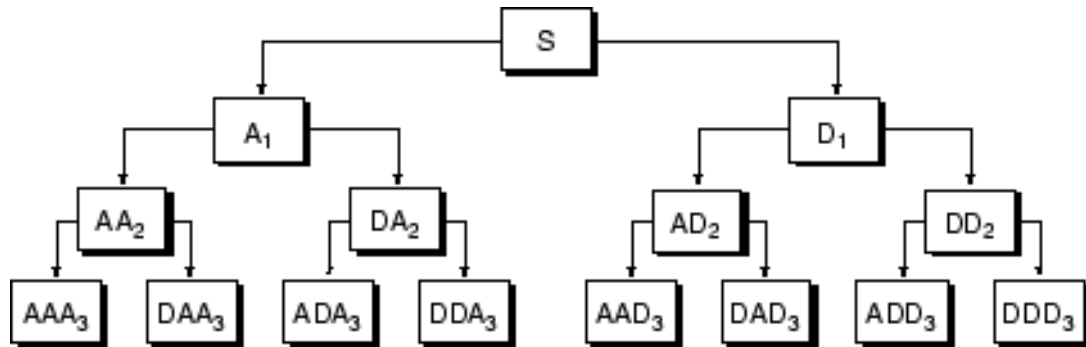
#### 4.2.3. Vremensko-frekvencijska domena

S obzirom na nestacionarnost signala dobivenih mjerenjem vibracija uzrokovanih greškom na ležaju, potrebno je koristiti odgovarajuće tehnike obrade signala kako bi se izvukle značajke koje sadrže korisne informacije o stanju elemenata ležaja. Općenito, stacionarnost podrazumijeva promjenjivost statističkih značajki u vremenu [49]. Potrebno je dakle identificirati nestacionarnosti kako bi ih bilo moguće dovesti u vezu s uzrokom. Za nestacionarne signale provedena su brojna istraživanja koja koriste kratkotrajnu Fourierovu transformaciju, valićnu transformaciju (*engl. Wavelet transformation*) te Wigner-Villeovu distribuciju. Ovim radom obuhvatit će se korištenje energetske i entropijske značajke pomoću valićne transformacije prilikom izvlačenja značajki za klasifikaciju vibracijskih signala iz vremensko-frekvencijske domene. Značajke koje se temelje na valićnoj transformaciji korisne su za otkrivanje naglih promjena izmjenjenog vibracijskog signala. Valićna paketna transformacija (*engl. Wavelet packet transform, WPT*) je generalizacija klasične kontinuirane valićne transformacije. Valićna transformacija prikazuje signal kao sumu određenih valnih oblika, odnosno malih valova – po čemu je transformacija dobila i ime. Prilikom paketne transformacije, reprezentacija signala se dobiva korištenjem digitalnih filtara. Prilikom prolaska kroz filtre, signal se razlaže na niskofrekvencijski pojas A1 i visokofrekvencijski pojas D1. Signal A1 sadrži frekvencijske komponente  $0 - \frac{f_s}{2}$  Hz i naziva se signal aproksimacije signala,



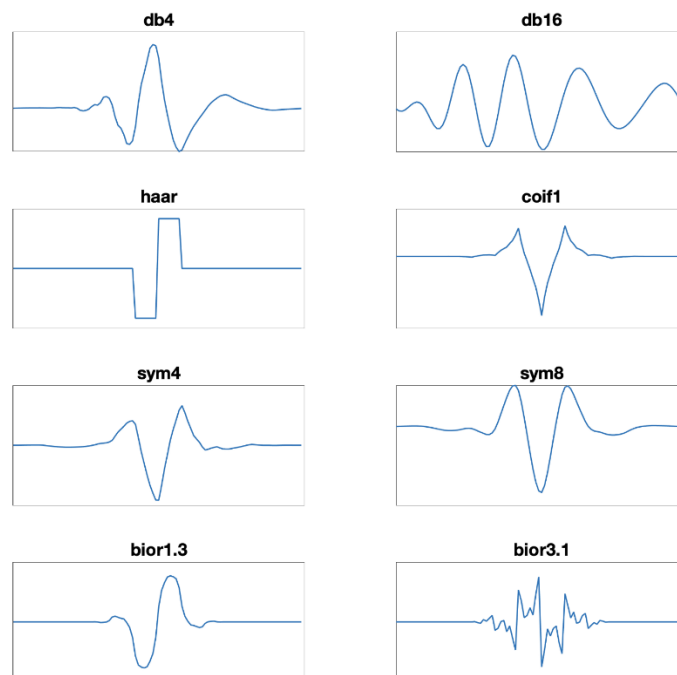
dok signal  $D_1$  sadrži spektar od  $\frac{f_s}{2} - f_s$  te se naziva signal detalja. Razlaganje je moguće nastaviti dok se ne dođe do željene razine, a posljedica razlaganja je struktura u obliku stabla.

Na slici 4.6. prikazana je valična paketna dekompozicija do treće razine.



Slika 4.6. Struktura stabla kod upotrebe WPT do treće razine

Prilikom valične dekompozicije moguće je odabrati neku od niza funkcija na kojima se temelji transformacija, od kojih je svaka korisna za različite primjene. U ovom radu koristit će se knjižnica PyWavelets u kojoj je implementirana valična paketna transformacija pri čemu je moguće odabrati između nekoliko funkcija na kojima se temelji valična transformacija od kojih su neke prikazane na slici 4.7. U ovom diplomskom radu korištena je temeljna funkcija `sym8` iz skupine funkcija pod nazivom `Symlets`.



Slika 4.7. Temeljne funkcije – valiči

Nakon valićne transformacije do treće razine dobiva se osam signala dobivenih iz izvornog.

Nadalje će se iz svakog od osam dobivenih signala izračunati vrijednosti energije prema:

$$E_i = \int_{-\infty}^{\infty} x_j^i(t) dt \quad (4.33)$$

Ukupnu energiju signala moguće je izračunati kao:

$$E = \sum_{i=1}^{2^j} E_i \quad (4.34)$$

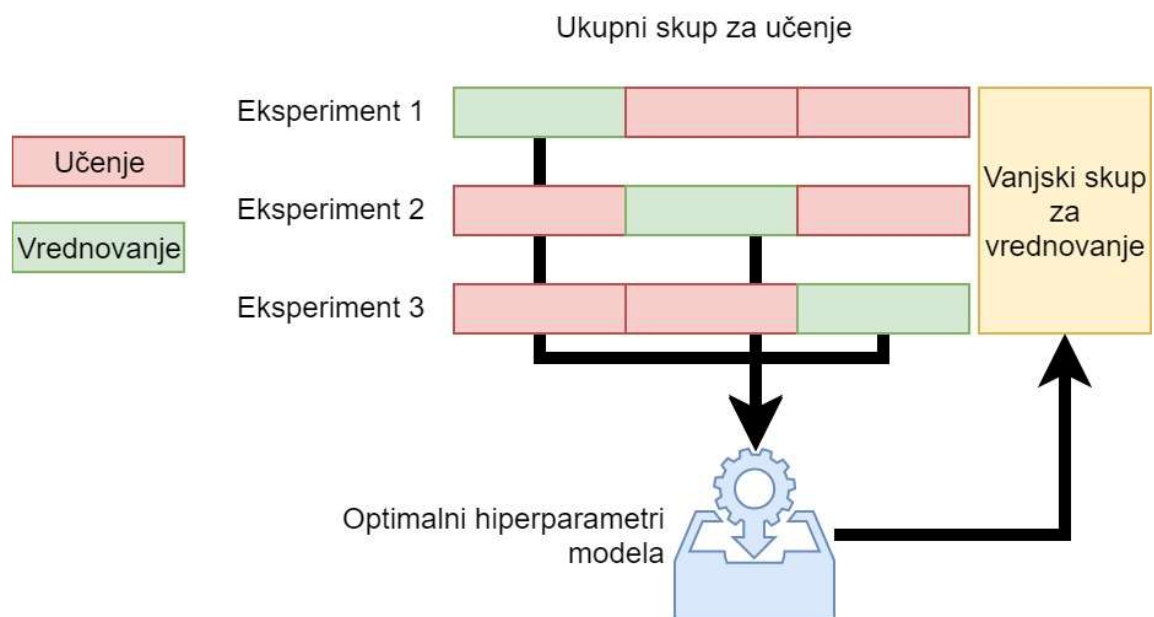
Također, moguće je izračunati i vrijednosti entropije. Koncept entropije često se koristi kao mjera neuređenosti sustava [50]. Koristeći izraz koji je predložio Shannon izračunat će se i vrijednosti entropije za svaki od osam dobivenih signala:

$$S = - \sum \frac{E_i}{E} \ln \frac{E_i}{E} \quad (4.35)$$

### 4.3. Treniranje modela

Cilj algoritma strojnog učenja svodi se na uspješno previđanje na novim, nepoznatim podacima, a njihova sposobnost za takve operacije naziva se generalizacija [51]. Kvaliteta predviđanja na vanjskom skupu može se kontrolirati pažljivim optimiranjem hiperparametara modela. Prilikom kreiranja instance modela, početnom se modelu dodjeljuju zadane vrijednosti hiperparametara te je potrebno zadati vrijednosti hiperparametara koje je potrebno kontrolirati. Hiperparametri modela u strojnom učenju imaju snažan učinak na njegovu izvedbu, stoga je nužno prilagoditi hiperparametre modela i pronaći optimum u parametarskom prostoru s kojima će model dati najbolju generalizaciju. Odabrani hiperparametri trebaju dati najbolje rezultate metričke procjene modela izmjerene na skupu za provjeru valjanosti, odnosno pogrešku modela. Ovisno o broju hiperparametara i složenosti modela, ručno pretraživanje može biti izrazito dugotrajno s obzirom na to da se model svaki put iznova trenira i validira s novim skupom hiperparametara. Jedan od načina ispitivanja prostora hiperparametara je mrežno pretraživanje (engl. *Grid search*). Metoda prolazi kroz sve kombinacije definiranog prostora pretraživanja pri čemu se kombinacija s najboljim rezultatom smatra optimalnim rješenjem. Mana ovog jednostavnog optimizacijskog algoritma je niska brzina izvođenja optimizacije.

Kako bi se izbjegla mogućnost da točnost modela ovisi o načinu odabira seta za učenje, prilikom mrežnog pretraživanja prostora hiperparametara provodi se i unakrsna validacija (Slika 4.8.). Unakrsna validacija je postupak kojim se skup za učenje modela dijeli na  $k$  jednakih dijelova, pri čemu se jedan koristi za validaciju, a  $k-1$  dijelova služi za treniranje modela. Proces se ponavlja  $k$  puta pri čemu se koristi drugi skup za validaciju.

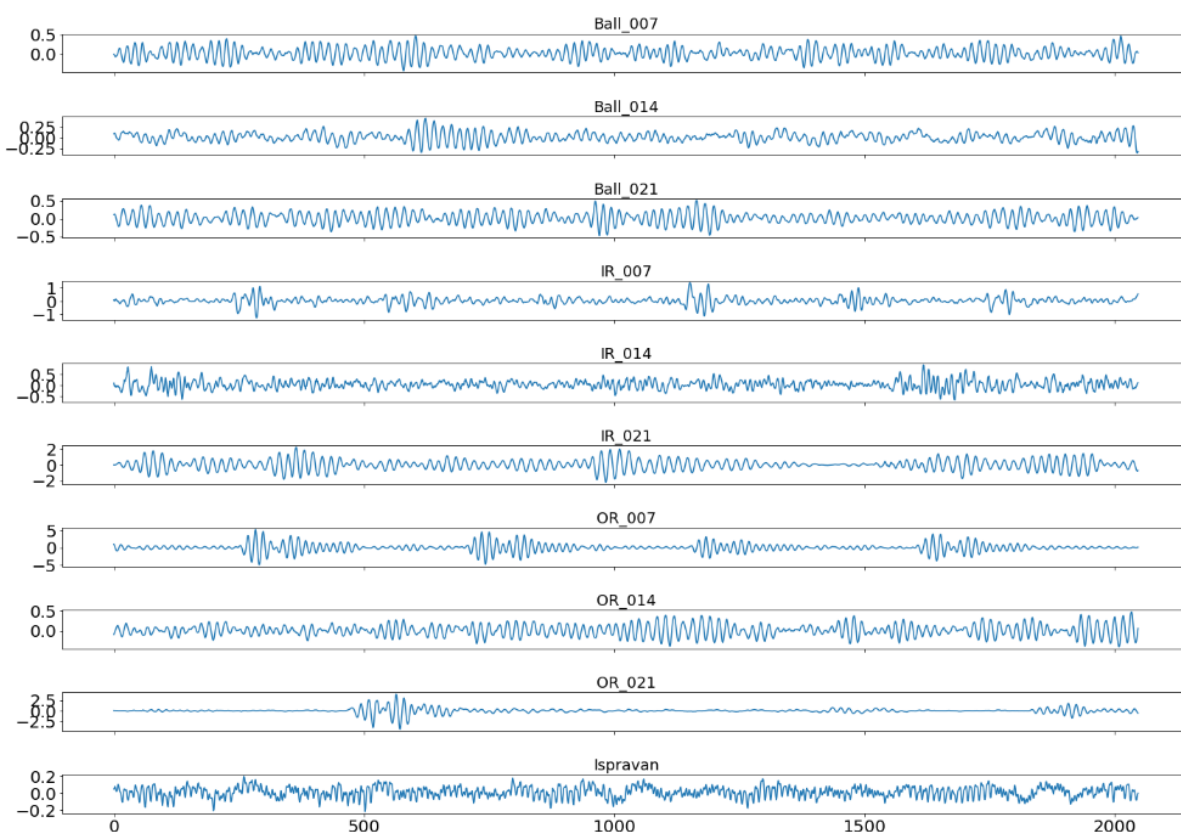


Slika 4.8. Trostruka unakrsna validacija (preuzeto iz [28])

Rezultat točnosti je prosječna točnost svih predikcija. Prilikom treniranja algoritama logističke regresije, strojeva s potpornim vektorima i nasumičnih šuma korištena je deseterostruka validacija. Rezultat optimizacije hiperparametarskog prostora je model treniran na svim značajkama. Kako bi se ubrzao algoritam i spriječila prenaučenos modela, također se provodi i odabir najvažnijih varijabli. Permutacijska važnost varijabli je metoda kojom se mjeri važnost varijable na način da se izračunava utjecaj promjene, odnosno permutacije svake od značajki na rezultat klasifikacija [52]. Ukoliko nasumična izmjena neke varijable značajno mijenja rezultate modela, toj se značajci pridaje visoka važnost. Varijabla je manje važna ukoliko se permutacijom njezinih vrijednosti rezultat mijenja minimalno ili ostaje nepromijenjen. Isti postupak se provodi više puta, kako bi se povećao interval pouzdanosti promjene rezultata uslijed permutacije vrijednosti pojedinog deskriptora. Zbog svoje univerzalne primjenjivosti u mnogim složenim modelima i jednostavnog pristupa, ova je metoda vrlo popularna [28]. Smanjenjem broja varijabli, odnosno uklanjanjem suvišnih podataka, u konačnici se postižu bolje performanse, odnosno smanjuje se računalno vrijeme te se olakšava razumijevanje modela. Nakon odabira najvažnijih varijabli još se jednom provodi mrežno pretraživanje optimalnih hiperparametara i unakrsna validacija kako bi se model prilagodio novom skupu značajki.

## 5. Rezultati

Na ukupno 4600 signala dobivenih mjerenjem vibracija primijenjene su tri tehnike izvlačenja značajki kako bi se omogućilo treniranje modela strojnog učenja. Slika 5.1. prikazuje 10 signala od kojih svaki sadrži 2048 uzoraka te opisuje različito zdravstveno stanje ležaja. Prva tri prikazana signala prikazuju tri veličine oštećenja kotrljajućeg elementa (0.007 inch, 0.014 inch te 0.021 inch), nadalje prikazani su signali s jednakim oštećenjima na unutarnjem prstenu ležaja te na vanjskom prstenu ležaja, respektivno. Posljednji signal prikazuje ispravan ležaj.



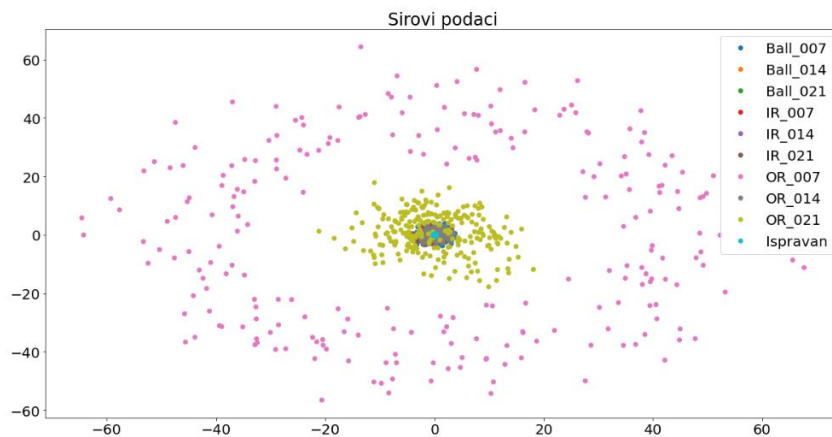
*Slika 5.1. Vizualizacija signala*

U nastavku ovoga rada usporedit će se utjecaj različitih načina izvlačenja značajki na performanse algoritma Logističke regresije, Nasumičnih šuma te Strojeva s potpornim vektorima prilikom nadziranog klasifikacijskog problema. Pokazat će se doprinos pojedinih grupa značajki najboljim rezultatima, ali isto tako i doprinos pojedinih algoritama. Kako bi se vizualizirao način na koji model uči i u konačnici predviđa rezultate, koristit će se algoritam nenadziranog učenja – Analiza glavnih komponenta. U konačnici će se opisati model i značajke čija kombinacija daje najbolje rezultate.

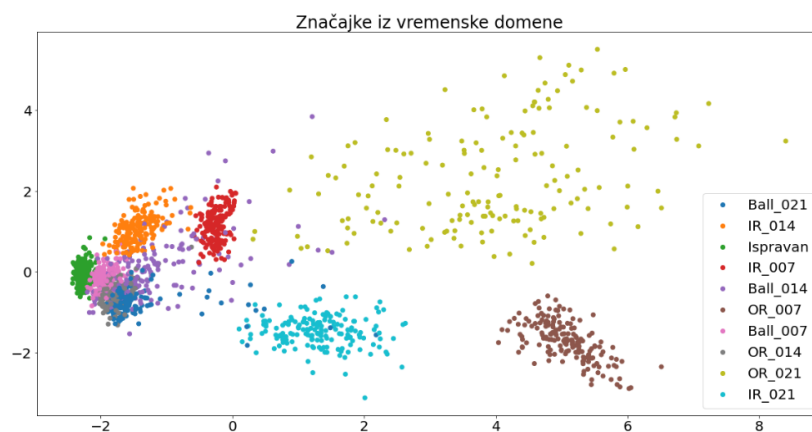
## 5.1. Vizualizacija značajki

Prilikom provedbe analize maksimalan broj glavnih komponenta jednak je broju deskriptora koji opisuju signal te su tako obuhvaćene sve informacije koje nosi set podataka na koji je primijenjena metoda redukcije dimenzionalnosti. Objašnjena varijanca je statistička mjera koja govori kolika se varijanca u skupu podataka može pripisati svakoj od glavnih komponentata (svojstvenih vektora), odnosno količina varijabilnosti u skupu podataka koja se može pripisati svakoj pojedinačnoj glavnoj komponenti [53]. Koncept objašnjene varijance je koristan u procjeni važnosti svake komponente – što je veća objašnjena varijanca, to je komponenta važnija. Slika 5.2. prikazuje rezultate analize glavnih komponentata. Vizualizacija se provodi odabirom najvažnijih komponentata, čiji broj ovisi o dimenzionalnosti prostora vizualizacije. U ovom slučaju vizualizacija je provedena u dvodimenzionalnom prostoru.

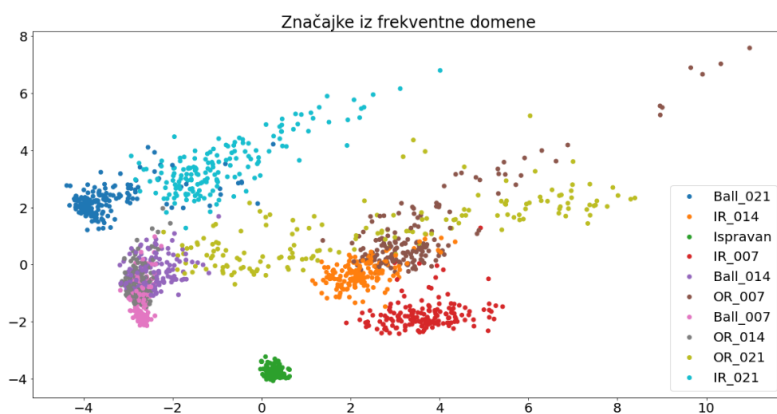
a)



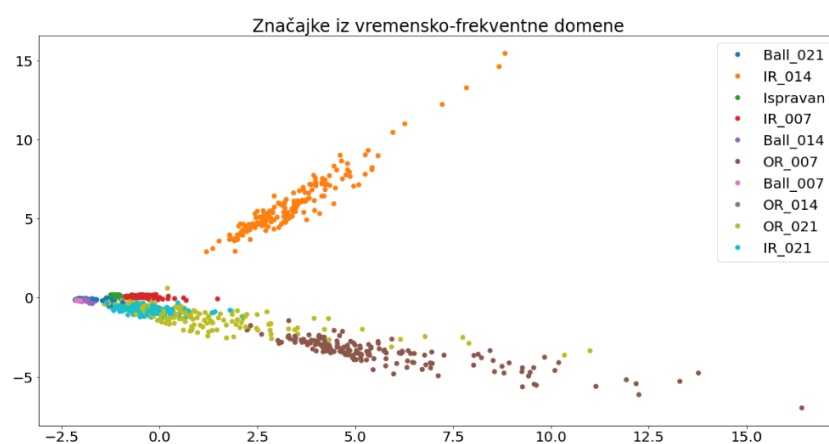
b)



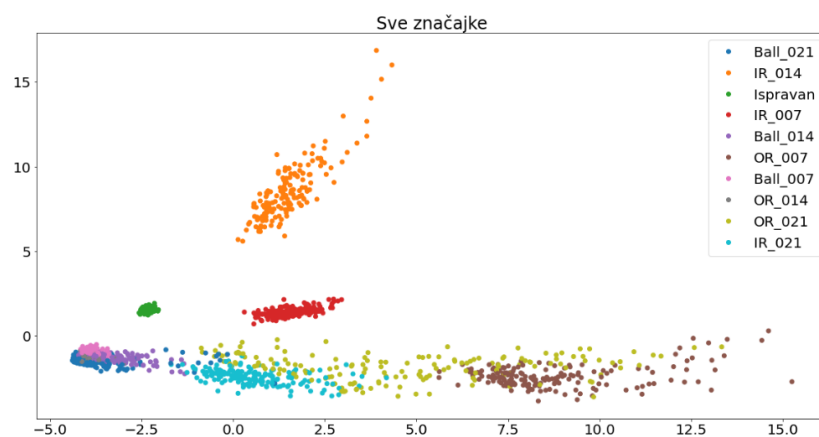
c)



d)



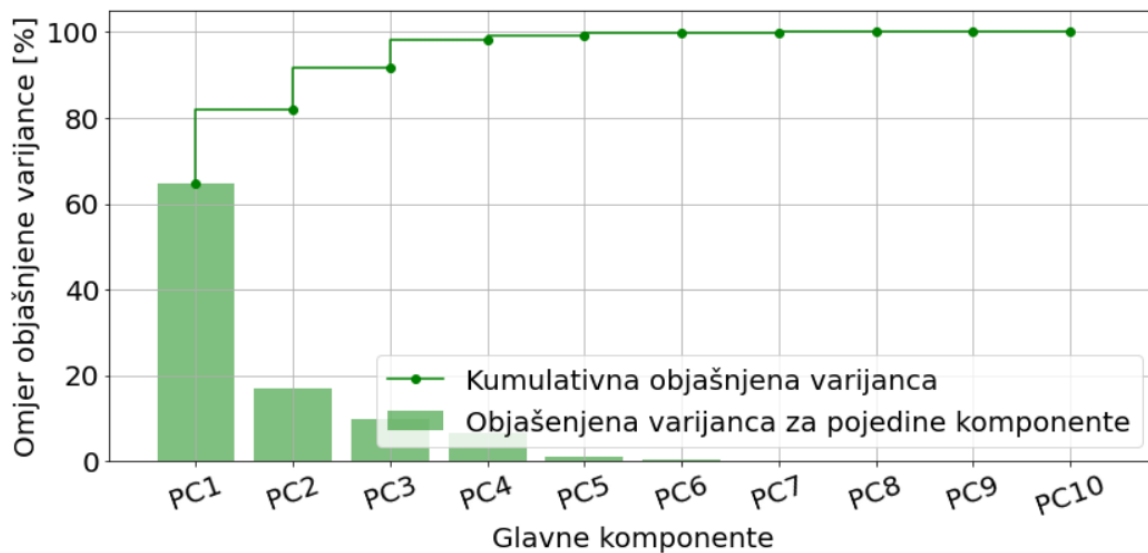
e)



Slika 5.2. Vizualizacija značajki PCA analizom

Analiza glavnih komponentata je najprije izvršena na izvornim podacima (signalima duljine 2048 uzoraka) te se kao rezultat dobiva 2048 glavnih komponentata koje u potpunosti opisuju izvorne signale. Odabirom dviju najvažnijih komponentata za vizualizaciju izvornih signala ukupno objašnjena varijanca iznosi 12%. Na slici 5.2. a) prikazane su dvije glavne komponente izvornih signala. Vidljivo je kako izvorne signale nije moguće opisati s dvije

značajke uz jasnu podjelu, a s obzirom na to da algoritmi strojnog učenja ostvaruju bolje performanse s manje značajki, u nastavku će se prikazati rezultati analize glavnih komponenata nakon izvlačenja značajki. Analizom glavnih komponenata na deset statističkih značajki dobivenih iz vremenske domene prikazana na slici 5.2. b). Ukupno objašnjena varijanca iznosi 82% (Slika 5.3.).



Slika 5.3. Objašnjena varijanca za vremenske značajke

Prilikom vizualizacije signala moguće je zamijetiti kako je teško razlučiti između veličina grešaka na kotrljajućem elementu. Objašnjena varijanca za dvije najvažnije komponente iz frekvencijske domene iznosi 73%. Izvlačenjem značajki iz frekvencijske domene ispravan ležaj označen zelenom bojom (Slika 5.2. c)) sada je potpuno odvojen od ostalih, što je vrlo bitno u kontekstu zdravstvenog menadžmenta. Vizualizacijom značajki iz vremensko-frekvencijske domene objašnjeno je 81% podataka. U konačnici na slici 5.2. e) vizualizirane su sve značajke, odnosno ukupno 44 značajke, sa objašnjenom varijancom u iznosu od 65%. Signali su gotovo jasno odvojeni, no i dalje nije moguće jasno razlučiti između veličina grešaka na kotrljajućem elementu.



## 5.2. Rezultati klasifikacije

U nastavku rada usporedit će se performanse pojedinih algoritama s tri grupe značajki. Sposobnost generalizacije algoritma ispitivat će se na testnom setu podataka s kojim se model još nije susreo. Točnost (engl. *Accuracy*) klasifikacije definirana je kao udio točno klasificiranih primjera u skupu svih primjera, putem izraza:

$$\text{točnost} = \frac{\text{broj točno klasificiranih signala}}{\text{ukupan broj signala}} \quad (5.1)$$

Testni set podataka, jednako kao i validacijski, sadrži 690 signala na kojima će se provjeriti rezultati klasifikacije. Naime, u potrazi za najboljom grupom značajki te algoritma koji ima najveću točnost, ispitane su sve kombinacije algoritama i grupa značajki. Radi jednostavnosti u nastavku će se koristiti oznake za grupe značajki navedene u tablici 5.1.

Tablica 5.1. Oznake grupa značajki

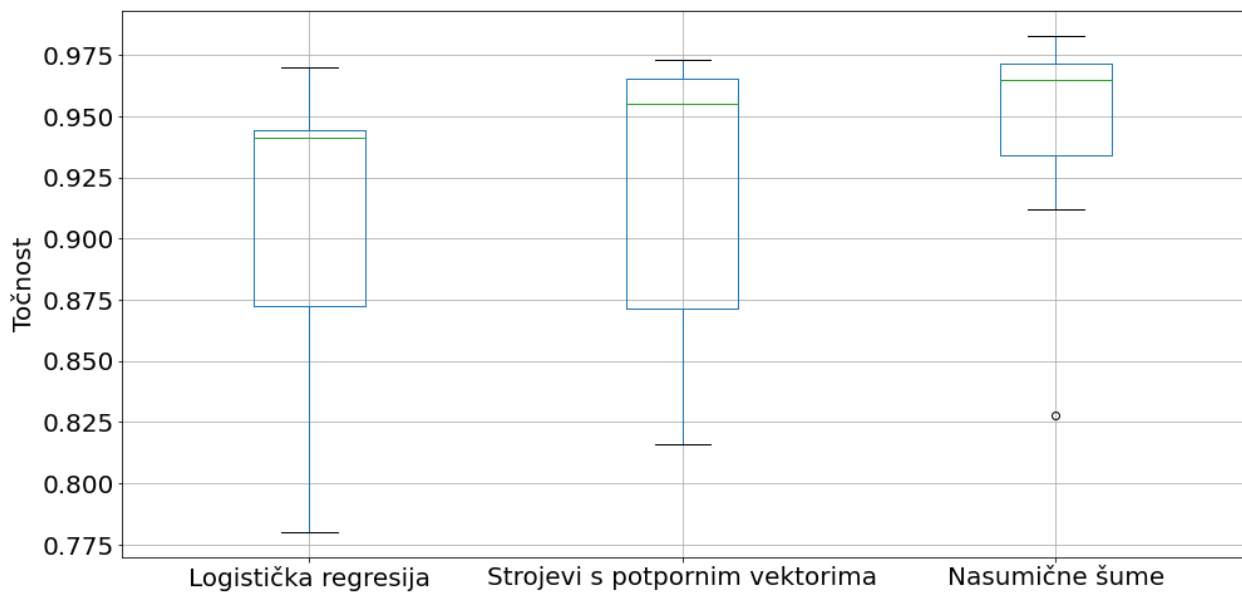
Grupa značajki	Oznaka
Značajke vremenske domene	P1
Značajke frekvencijske domene	P2
Značajke vremensko-frekvencijske domene	P3

Tablica 5.2. prikazuje rezultate predviđanja, prema izrazu (5.1), pojedinih kombinacija seta značajki i korištenog algoritma na validacijskom setu podataka koji sadrži 690 signala. Vidljivo je kako je najbolje rezultate dao Algoritam nasumičnih šuma za set značajki koji obuhvaća sve tri metode izvlačenja značajki.

Tablica 5.2. Točnost klasifikacije prema izrazu 5.1.

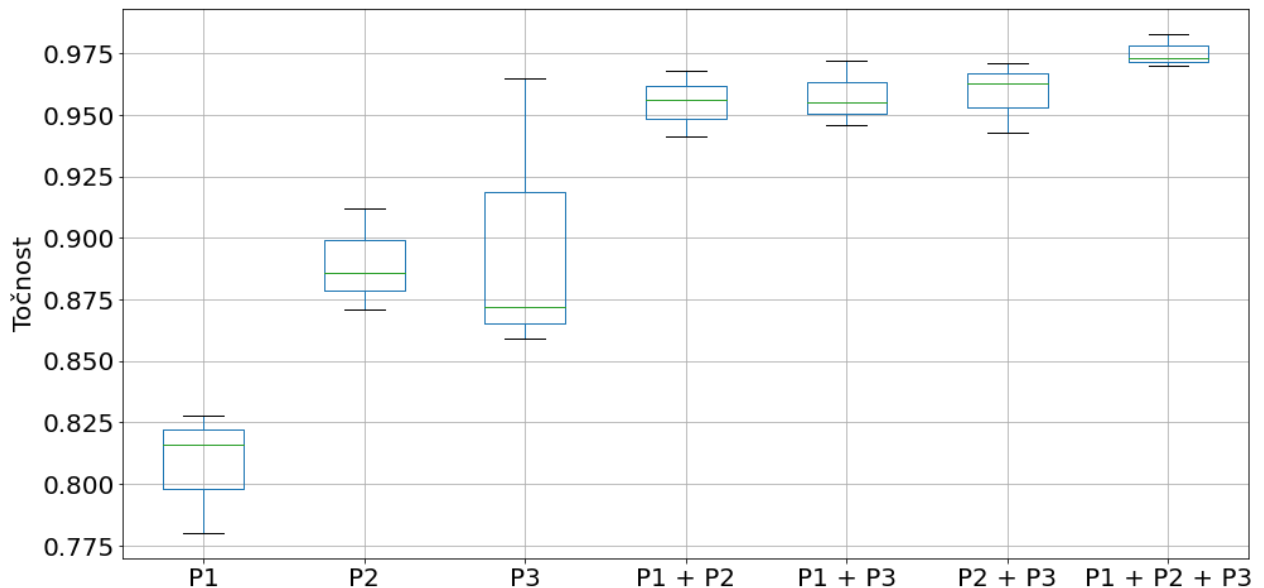
	<b>Algoritam nasumičnih šuma [%]</b>	<b>Strojevi s potpornim vektorima [%]</b>	<b>Logistička regresija [%]</b>
<b>P1</b>	82.8	81.6	78.0
<b>P2</b>	91.2	87.1	88.6
<b>P3</b>	96.5	87.2	85.9
<b>P1 + P2</b>	95.6	96.8	94.1
<b>P1 + P3</b>	97.2	95.5	94.6
<b>P2 + P3</b>	97.1	96.3	94.3
<b>P1 + P2 + P3</b>	98.8	97.3	97.0

Rezultati prikazani u tablici 5.2. također su prikazani na slikama 5.4. i 5.5. Naime, premda je model Nasumičnih šuma pokazao najbolje rezultate, model Strojeva s potpornim vektorima također je pokazao vrlo dobre rezultate. Logistička regresija pokazala je rezultate vrlo slične Strojevima s potpornim vektorima, no prilikom treniranja modela s vremenskim značajkama nije uspjela postići visoku točnost kao s ostalim grupama značajki.



Slika 5.4. Doprinos algoritma

Očekivano, značajke iz vremenske domene nisu dovoljne za potpuno točnu klasifikaciju grešaka na ležaju, no sadrže informacije koje smanjuju devijaciju od najboljeg rješenja prilikom korištenja sve tri grupe značajki. U nastavku će se detaljnije opisati rezultati najboljeg modela.



Slika 5.5. Doprinos seta značajki

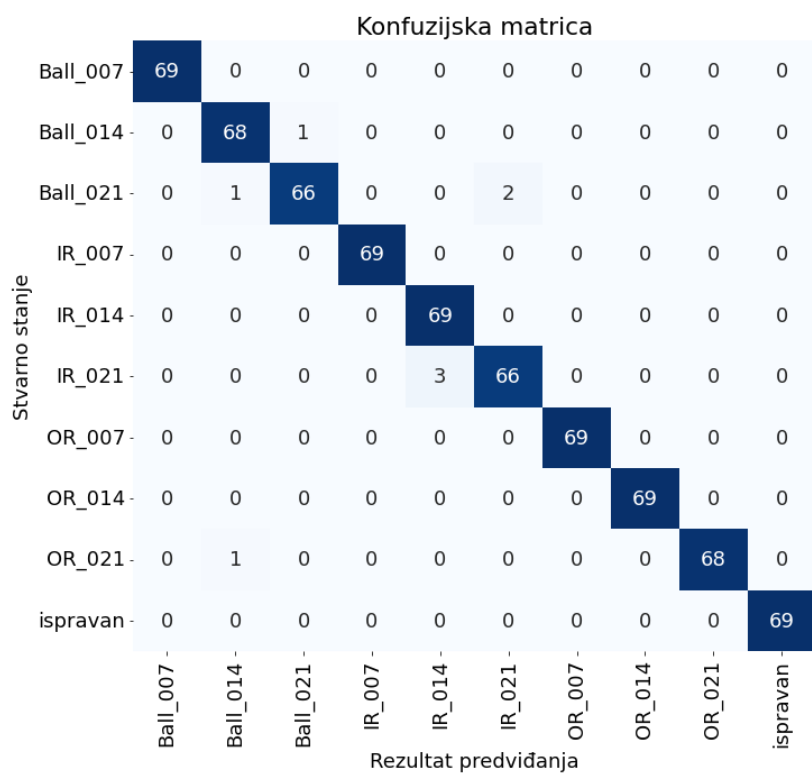
Kako bi se potvrdila točnost modela Nasumičnih šuma, prevedeno je 10 neovisnih predviđanja na nasumično poduzorkovanim skupovima unutar testnog seta. Rezultati su prikazani u tablici 5.3.

Tablica 5.3. Rezultati predviđanja modela

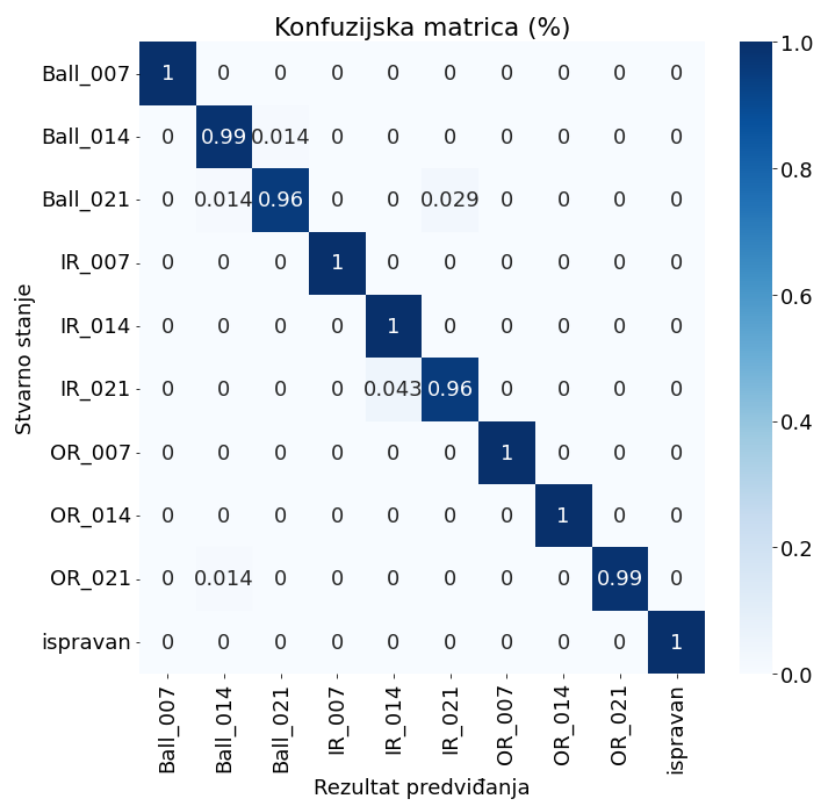
	Točnost modela (5.1.) [%]
Srednja vrijednost točnosti algoritma	97.97
Standardna devijacija	1.47
Maksimalna točnost	100
Minimalna točnost	95.65

Temeljem klasifikacije signala na cijelom validacijskom setu korištenjem najboljeg modela, moguće je prikazati konfuzijsku matricu, odnosno matricu zabune modela. Konfuzijska matrica se koristi za određivanje izvedbe modela klasifikacije za dani skup testnih podataka. Validacijski set sadrži 10 klasa koje je potrebno klasificirati, pri čemu svaka klasa sadrži 69 signala. Konfuzijska matrica prikazana je na slici 5.6.

a)

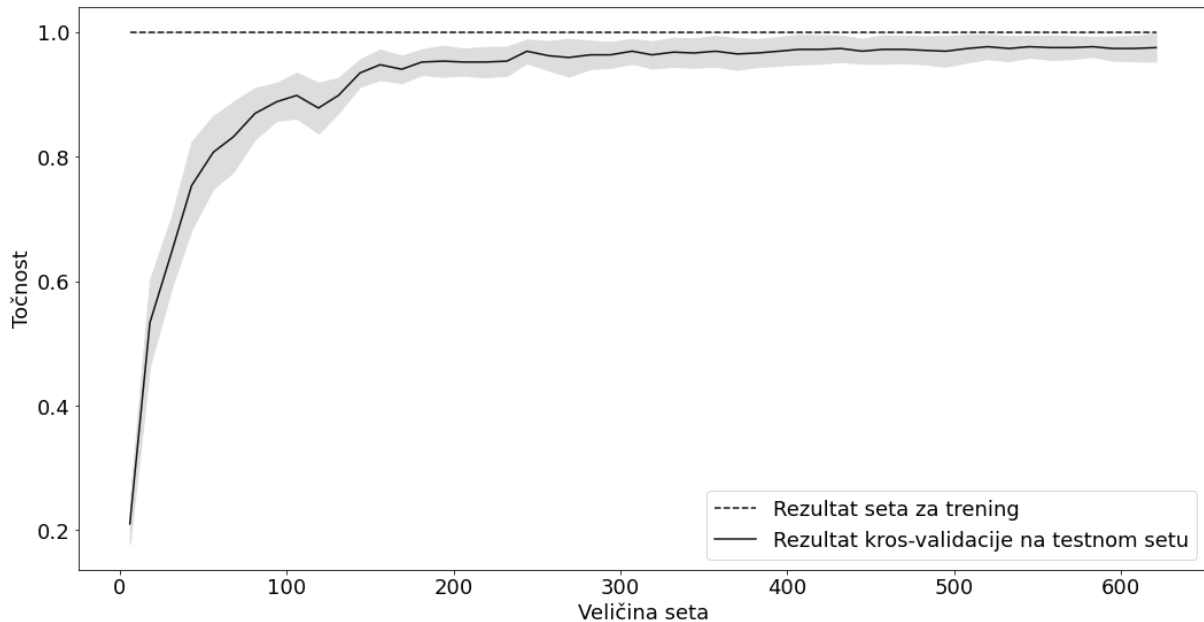


b)



Slika 5.6. Konfuzijska matrica

Od ukupno 690 signala, model nije točno klasificirao 8 signala, no vrlo je bitno zamijetiti kako je algoritam sve ispravne ležajeve označio ispravnim, što je u kontekstu zdravstvenog menadžmenta vrlo bitno. Dodatna provjera obučenog modela, može se izvršiti provjerom tijeka učenja (Slika 5.7.). Bitno je vidjeti pati li model od podnaučenosti ili prenaučivosti.



Slika 5.7. Krivulja učenja

Ukoliko linije konvergiraju jedna prema drugoj, model je najvjerojatnije dobar. Podnaučenost modela moguće je vidjeti do veličine seta od otprilike 100 signala, no već nakon 200 danih signala može se reći da je učenje konvergiralo. Ako se kros validacijska krivulja počne smanjivati, riječ je o prenaučivosti. Temeljem prethodno navedenog, krivulje na slici 5.7. jasno prikazuju kako model dobro generalizira. Tijekom učenja modela mrežnim se pretraživanjem pronalaze najbolji hiperparametri, a također permutacijskom važnosti odabiru se značajke koje u najvećoj mjeri utječu na ishod predviđanja. Najbolji model dobiven je značajkama navedenim u nastavku:

- vremenska domena: srednja vrijednost, faktor forme, standardna devijacija, crest faktor, kurtozis, maksimalna vrijednost
- frekvencijska domena: spektralno središte, kurtozis, skewness, faktor forme, BPFO, BSF, BPFI

- vremensko-frekvencijska domena: DAA – entropija, ADA – entropija, DAD – entropija, ADA – energija, AAA – energija, AAD – entropija, DAA – energija, ADD – energija.

Za model nasumičnih šuma parametri odabrani mrežnim pretraživanjem prikazani su u tablici 5.3:

*Tablica 5.3. Odabrani hiperparametri*

	Parametarski prostor	Odabrani parametar
Broj stabala	100, 150, 200, 250, 300	250
Dubina drveta	5, 6, 7, 8, 9, 10	9
Minimalni broj uzoraka u listu	2, 3, 4, 5	3
Maksimalan broj značajki	auto, log2	log2

## 6. Zaključak

S pojavom koncepta industrije 4.0 koncept prognostičkog i zdravstvenog menadžmenta postao je nezaobilazna tendencija u okviru prediktivnog održavanja. Kao napredna strategija održavanja, prediktivno održavanje mjeri parametre o stanju opreme kako bi se izvršili odgovarajući zadatci za optimizaciju vijeka trajanja stroja i procesa bez povećanja rizika od kvara. Glavna zadaća je prikupljanje podataka s opremom koja je u pogonu, otkrivanje informacija iz prikupljenih podataka i utvrđivanje potencijalnih kvarova ili degradacije kroz povijesnu analizu slične opreme i znanja stečenog tijekom vremena, a ključna je točnost predviđanja kvarova ili degradacije. Cilj ovoga rada bio je usmjeren na dizajn i izgradnju modela sposobnog za izvršavanje zadataka u sklopu paradigme prediktivnog održavanja. Za ovu vrstu problema odabran je pristup vođen podacima, koji se temelji na algoritmima strojnog učenja. Osjetnici ubrzanja prikupljali su podatke sa eksperimentalne stanice u radu, čiji su ležajevi dovedeni u neispravno stanje upotrebom elektroerozijske obrade. Nakon testiranja različitih kombinacija za izdvajanje i odabir značajki, odabrana je najbolja metoda obrade izvornih podataka prema modelu strojnog učenja koji će se koristiti za problem višerazrednog klasifikacijskog problema. Razvijeni model pokazao se sposobnim klasificirati i detektirati anomalije, što omogućava dijagnostičke funkcije. Najbolji rezultat klasifikacije dobiven je pomoću algoritma Nasumičnih šuma na odabranim varijablama pomoću permutacijske važnosti na značajkama iz vremenske, frekvencijske i vremensko-frekvencijske domene. Kako bi se potvrdila točnost modela Nasumičnih šuma prevedeno je 10 neovisnih predviđanja na nasumično poduzorkovanim skupovima unutar testnog seta, gdje je dobivena prosječna vrijednost točnosti jednaka 0.98. Premda su modeli doista pokazali izvrsnu točnost u razlikovanju stanja, problem je što moraju trenirati na skupu podataka svih grešaka koje treba identificirati. Konačni je cilj prediktivnog održavanja usmjeren na razvoj sustava koji ima mogućnost samostalne procjene zdravstvenog stanja opreme, odnosno koji je sposoban samostalno donositi odluke o aktivnostima održavanja koje je potrebno provesti. Stoga bi daljnji smjer istraživanja mogao bi biti usmjeren na progresivno učenje koje polazi od prepoznavanja anomalija poznavanjem samo normalnog rada stroja i koje omogućuje tvrtkama da povećavaju skup poznatih grešaka kako se pojavljuju na strojevima. Integracijom algoritama za klasifikaciju i otkrivanje anomalija, može se održati dobra točnost modela klasifikacije bez gubitka prednosti koje uvodi algoritam za otkrivanje anomalija, kao što je sposobnost predviđanja sa malom količinom neoznačenih podataka. To bi omogućilo primjenu prediktivnog održavanja i na strojevima koji nemaju reprezentativan skup podataka.

## LITERATURA

- [1] B. Mrugalska and M. K. Wyrwicka, “Towards Lean Production in Industry 4.0,” *Procedia Engineering*, vol. 182, pp. 466–473, Jan. 2017, doi: 10.1016/J.PROENG.2017.03.135.
- [2] Z. M. Çinar, A. A. Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, “Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0,” *Sustainability 2020, Vol. 12, Page 8211*, vol. 12, no. 19, p. 8211, Oct. 2020, doi: 10.3390/SU12198211.
- [3] B. Bajic, A. Rikalovic, N. Suzic, and V. Piuri, “Industry 4.0 Implementation Challenges and Opportunities: A Managerial Perspective,” *IEEE Systems Journal*, vol. 15, no. 1, pp. 546–559, Mar. 2021, doi: 10.1109/JSYST.2020.3023041.
- [4] A. Sahli, R. Evans, and A. Manohar, “Predictive Maintenance in Industry 4.0: Current Themes,” *Procedia CIRP*, vol. 104, pp. 1948–1953, 2021, doi: 10.1016/J.PROCIR.2021.11.329.
- [5] S. Fernandes, M. Antunes, A. R. Santiago, J. P. Barraca, D. Gomes, and R. L. Aguiar, “Forecasting Appliances Failures: A Machine-Learning Approach to Predictive Maintenance”, doi: 10.3390/info11040208.
- [6] Z. Li, K. Wang, and Y. He, “Industry 4.0 - Potentials for Predictive Maintenance,” 2016, doi: 10.2991/IWAMA-16.2016.8.
- [7] M. Bashiri, H. Badri, and T. H. Hejazi, “Selecting optimum maintenance strategy by fuzzy interactive linear assignment method,” *Applied Mathematical Modelling*, vol. 35, no. 1, pp. 152–164, Jan. 2011, doi: 10.1016/J.APM.2010.05.014.
- [8] P. Gackowiec, “General overview of maintenance strategies – concepts and approaches,” *Multidisciplinary Aspects of Production Engineering*, vol. 2, no. 1, pp. 126–139, Sep. 2019, doi: 10.2478/MAPE-2019-0013.
- [9] D. Kolar, “Model rane procjene kvarova rotacijske opreme primjenom dubokog strojnog učenja | Repozitorij Fakulteta strojarstva i brodogradnje Sveučilišta u Zagrebu.” <https://repozitorij.fsb.unizg.hr/islandora/object/fsb%3A5538> (accessed Apr. 04, 2022).



- [10] Sanjin Braut, “Eksperimentalno ispitivanje u mehanici konstrukcija i strojeva - predavanja.”
- [11] M. Lovrić *et al.*, “Parasitic resistance as a predictor of faulty anodes in electro galvanizing : a comparison of machine learning , physical and hybrid models,” *Advanced Modeling and Simulation in Engineering Sciences*, vol. 7, no. 46, 2020, doi: 10.1186/s40323-020-00184-z.
- [12] C. M. Harris and A. G. Piersol, “Harris’ shock and vibration handbook,” 2002.
- [13] “Advances in Machine Learning and Data Analytics,” *Lecture Notes in Electrical Engineering*, vol. 48 LNEE, pp. 3–29, Jan. 2021, doi: 10.1016/B978-0-323-85510-5.00001-6.
- [14] G. Ravnestad, S. S. Panesar, D. Kayrbekova, and T. Markeset, “Improving periodic preventive maintenance strategies using condition monitoring data,” *IFIP Advances in Information and Communication Technology*, vol. 384 AICT, pp. 260–267, 2012, doi: 10.1007/978-3-642-33980-6\_30.
- [15] R. Dupuis, “Application of Oil Debris Monitoring For Wind Turbine Gearbox Prognostics and Health Management,” *Annual Conference of the PHM Society*, vol. 2, no. 1, 2010, doi: 10.36001/PHMCONF.2010.V2I1.1867.
- [16] “ISO - ISO 20816-1:2016 - Mechanical vibration — Measurement and evaluation of machine vibration — Part 1: General guidelines.” <https://www.iso.org/standard/63180.html> (accessed Apr. 10, 2022).
- [17] V. Saxena, “Assessment of Gearbox Fault Detection Using Vibration Signal Analysis and Acoustic Emission Technique,” *IOSR Journal of Mechanical and Civil Engineering*, vol. 7, no. 4, pp. 52–60, 2013, doi: 10.9790/1684-0745260.
- [18] A. M. Al-Ghamd and D. Mba, “A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size,” *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1537–1571, Oct. 2006, doi: 10.1016/J.YMSSP.2004.10.013.
- [19] H. M. Hashemian and W. C. Bean, “State-of-the-art predictive maintenance techniques,” *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 10, pp. 3480–3492, Oct. 2011, doi: 10.1109/TIM.2009.2036347.

- [20] M. H. Mohd Ghazali and W. Rahiman, "Vibration Analysis for Machine Monitoring and Diagnosis: A Systematic Review," *Shock and Vibration*, vol. 2021, 2021, doi: 10.1155/2021/9469318.
- [21] A. Yunusa-Kaltungo and R. Cao, "Towards Developing an Automated Faults Characterisation Framework for Rotating Machines. Part 1: Rotor-Related Faults," *Energies 2020, Vol. 13, Page 1394*, vol. 13, no. 6, p. 1394, Mar. 2020, doi: 10.3390/EN13061394.
- [22] A. Mayr *et al.*, "Machine Learning in Production – Potentials, Challenges and Exemplary Applications," *Procedia CIRP*, vol. 86, pp. 49–54, Jan. 2019, doi: 10.1016/J.PROCIR.2020.01.035.
- [23] Z. M. Çinar, A. A. Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, "Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0," *Sustainability 2020, Vol. 12, Page 8211*, vol. 12, no. 19, p. 8211, Oct. 2020, doi: 10.3390/SU12198211.
- [24] "How To Develop a Machine Learning Model From Scratch | by Victor Roman | Towards Data Science." <https://towardsdatascience.com/machine-learning-general-process-8f1b510bd8af> (accessed Apr. 30, 2022).
- [25] "Bayes' Theorem Definition." <https://www.investopedia.com/terms/b/bayes-theorem.asp> (accessed Apr. 30, 2022).
- [26] "The Hundred-Page Machine Learning Book by Andriy Burkov." <http://themlbook.com/> (accessed Apr. 30, 2022).
- [27] "Machine Learning | An Introduction | by Gavin Edwards | Towards Data Science." <https://towardsdatascience.com/machine-learning-an-introduction-23b84d51e6d0> (accessed Apr. 30, 2022).
- [28] M. Lovrić, "Razvoj i primjena modela za procjenu ekotoksikoloških rizika bioaktivnih kemijskih spojeva," University of Zagreb, 2021. [Online]. Available: <https://urn.nsk.hr/urn:nbn:hr:217:730117>
- [29] "Commonly Used Machine Learning Algorithms | Data Science." <https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/> (accessed May 20, 2022).

- [30] “Introduction to Logistic Regression | by Ayush Pant | Towards Data Science.” <https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148> (accessed May 23, 2022).
- [31] “Understanding Random Forest. How the Algorithm Works and Why it Is... | by Tony Yiu | Towards Data Science.” <https://towardsdatascience.com/understanding-random-forest-58381e0602d2> (accessed May 23, 2022).
- [32] “SVM Machine Learning Tutorial – What is the Support Vector Machine Algorithm, Explained with Code Examples.” <https://www.freecodecamp.org/news/svm-machine-learning-tutorial-what-is-the-support-vector-machine-algorithm-explained-with-code-examples/> (accessed May 23, 2022).
- [33] “SUPPORT VECTOR MACHINES(SVM). Introduction: All you need to know... | by Ajay Yadav | Towards Data Science.” <https://towardsdatascience.com/support-vector-machines-svm-c9ef22815589> (accessed May 23, 2022).
- [34] M. Lovrić *et al.*, “Should We Embed in Chemistry? A Comparison of Unsupervised Transfer Learning with PCA, UMAP, and VAE on Molecular Fingerprints,” *Pharmaceuticals*, vol. 14, no. 8, 2021, doi: 10.3390/ph14080758.
- [35] Y. Lei, *Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery*.
- [36] X. Zhang, B. Zhao, and Y. Lin, “Machine Learning Based Bearing Fault Diagnosis Using the Case Western Reserve University Data: A Review,” *IEEE Access*, vol. 9, pp. 155598–155608, 2021, doi: 10.1109/ACCESS.2021.3128669.
- [37] J. G. Hoffer *et al.*, “Theory-inspired machine learning—towards a synergy between knowledge and data,” *Welding in the World* 66:7, vol. 66, no. 7, pp. 1291–1304, Apr. 2022, doi: 10.1007/S40194-022-01270-Z.
- [38] Y. Cai, “BEARING FAULT DIAGNOSIS USING DEEP LEARNING NEURAL NETWORKS WITH INPUT PROCESSING,” Dec. 2021, doi: 10.25394/PGS.17162480.V1.

- [39] “Case Western Reserve University: One of the nation’s best.” <https://case.edu/> (accessed Jun. 15, 2022).
- [40] “Bearing Data Center | Case School of Engineering | Case Western Reserve University.” <https://engineering.case.edu/bearingdatacenter> (accessed Feb. 05, 2022).
- [41] S. L. Lin, “Application of Machine Learning to a Medium Gaussian Support Vector Machine in the Diagnosis of Motor Bearing Faults,” *Electronics 2021, Vol. 10, Page 2266*, vol. 10, no. 18, p. 2266, Sep. 2021, doi: 10.3390/ELECTRONICS10182266.
- [42] “Project Jupyter | Home.” <https://jupyter.org/> (accessed Jun. 15, 2022).
- [43] “PyCharm: the Python IDE for Professional Developers by JetBrains.” <https://www.jetbrains.com/pycharm/> (accessed Jul. 02, 2022).
- [44] T. Thelaidjia, A. Moussaoui, and S. Chenikher, “Feature extraction and optimized support vector machine for severity fault diagnosis in ball bearing,” *Engineering Solid Mechanics*, vol. 4, no. 4, pp. 167–176, 2016, doi: 10.5267/J.ESM.2016.6.004.
- [45] A. Kafeel *et al.*, “An Expert System for Rotating Machine Fault Detection Using Vibration Signal Analysis,” *Sensors 2021, Vol. 21, Page 7587*, vol. 21, no. 22, p. 7587, Nov. 2021, doi: 10.3390/S21227587.
- [46] P. H. Jain and S. P. Bhosle, “Study of effects of radial load on vibration of bearing using time-Domain statistical parameters,” *IOP Conference Series: Materials Science and Engineering*, vol. 1070, no. 1, p. 012130, Feb. 2021, doi: 10.1088/1757-899X/1070/1/012130.
- [47] L. Eren, “Bearing fault detection by one-dimensional convolutional neural networks,” *Mathematical Problems in Engineering*, vol. 2017, 2017, doi: 10.1155/2017/8617315.
- [48] L. Guo, J. Chen, and X. Li, “Rolling Bearing Fault Classification Based on Envelope Spectrum and Support Vector Machine;,” <http://dx.doi.org/10.1177/1077546308095224>, vol. 15, no. 9, pp. 1349–1363, Jul. 2009, doi: 10.1177/1077546308095224.

- [49] “PREDNOSTI ANALIZE NESTACIONARNIH HIDROLOŠKIH VREMENSKIH SERIJA METODOM VALIĆNE TRANSFORMACIJE - CROSBİ.” <https://www.bib.irb.hr/810056> (accessed Jun. 27, 2022).
- [50] M. Varanis and R. Pederiva, “Wavelet Packet Energy-Entropy Feature Extraction and Principal Component Analysis for Signal Classification,” Aug. 2015, doi: 10.5540/03.2015.003.01.0471.
- [51] C. Albon, “Machine learning with Python cookbook : practical solutions from preprocessing to deep learning”.
- [52] “4.2. Permutation feature importance — scikit-learn 1.1.1 documentation.” [https://scikit-learn.org/stable/modules/permutation\\_importance.html](https://scikit-learn.org/stable/modules/permutation_importance.html) (accessed Jul. 07, 2022).
- [53] “PCA Explained Variance Concepts with Python Example - Data Analytics.” <https://vitalflux.com/pca-explained-variance-concept-python-example/> (accessed Jul. 03, 2022).

## POPIS SLIKA

Slika 2.1. Razvoj industrije .....	2
Slika 2.2. Strategije održavanja.....	4
Slika 2.3. Moderna klasifikacija strategija održavanja (preuzeto iz [6]) .....	4
Slika 2.4. Krivulja kade (preuzeto iz [10]).....	6
Slika 2.5. Odnos troškova i strategije održavanja .....	8
Slika 2.6. ISO 20816-1:2016 (preuzeto iz [15]).....	12
Slika 2.7. Instrumenti za mjerenje vibracija (preuzeto iz [10]).....	14
Slika 2.8. Prediktivno održavanje temeljeno na stanju .....	16
Slika 3.1. Model strojnog učenja.....	17
Slika 3.2. Oblici strojnog učenja .....	19
Slika 3.3. Primjer nedovoljnog treniranja, dobrog treniranja i prekomjernog treniranja (preuzeto iz [25]).....	20
Slika 3.4. Klasifikacija i regresija .....	20
Slika 3.5. Proces učenja za a) tradicionalne algoritme i b) algoritme dubokog učenja.....	22
Slika 3.6. Standardna logistička funkcija .....	23
Slika 3.7. Shematski prikaz binarne klasifikacije algoritmom nasumičnih šuma (preuzeto iz [27]).....	25
Slika 3.8. SVM model za dvodimenzionalne vektore značajki (preuzeto iz [25]).....	27
Slika 4.1. Eksperimentalna stanica.....	29
Slika 4.2. Kuglični ležaj .....	30
Slika 4.3. Vrste nanesenih grešaka (preuzeto iz [38]).....	30
Slika 4.4. Priprema podataka.....	33
Slika 4.5. Faze degradacije ležaja (preuzeto iz [44]) .....	36
Slika 4.6. Struktura stabla kod upotrebe WPT do treće razine .....	41
Slika 4.7. Temeljne funkcije – valići.....	41
Slika 4.8. Trostruka unakrsna validacija (preuzeto iz [27]) .....	43
Slika 5.1. Vizualizacija signala .....	45
Slika 5.2. Vizualizacija značajki PCA analizom.....	47
Slika 5.3. Objasnjena varijanca za vremenske značajke .....	48
Slika 5.4. Doprinos algoritma .....	50
Slika 5.5. Doprinos seta značajki .....	51
Slika 5.6. Konfuzijska matrica .....	52

Slika 5.7. Krivulja učenja ..... 53

## **POPIS TABLICA**

Tablica 4.1 Imena klasa prilikom klasifikacije .....	30
Tablica 5.1. Oznake grupa značajki .....	49
Tablica 5.2. Točnost klasifikacije prema izrazu 5.1.....	50
Tablica 5.3. Odabrani hiperparametri .....	54



## SAŽETAK

Jedna od najznačajnijih primjena tehnologije u kontekstu Industrije 4.0 je razvoj sustava koji je sposoban samostalno procijeniti zdravstveno stanje opreme i donositi odluke o aktivnostima održavanja koje je potrebno provesti. Ovaj diplomski rad bavi se uvodom u ciljeve, poteškoće i poznate probleme, ali i rješenja koja nudi literatura vezano za osnovne koncepte prediktivnog održavanja i zdravstvenog menadžmenta. U radu je opisan dizajn i implementacija modela koji može odrediti zdravstveno stanje ležaja. Ovo rješenje temelji se na pristupu vođenim podacima, koristeći podatke prikupljane osjetnikom vibracija te algoritme strojnog učenja za određivanje stanja ležaja. Posebna pažnja je posvećena obradi podataka te izvlačenju i odabiru relevantnih značajki.

**Ključne riječi:** Industrija 4.0, prediktivno održavanje, valjni ležajevi, Algoritam Nasumičnih Šuma, Analiza glavnih komponenata

## **ABSTRACT**

One of the most significant applications of technology in the context of Industry 4.0 is the development of a system capable of independently assessing the health of equipment and making decisions about maintenance activities that need to be carried out. This thesis deals with an introduction to goals, difficulties and known problems, as well as solutions offered by the literature related to the basic concepts of predictive maintenance and health management. The paper describes the design and implementation of a model that can determine the state of health of the bearing. This solution is based on data-driven approach, using data collected by an accelerometer and machine learning algorithms to determine the condition of the bearing. Special attention is paid to data preprocessing and extraction and selection of relevant features.

**Keywords:** Industry 4.0, predictive maintenance, Random Forest algorithm, rolling bearings, Principal component analysis