

*Metode umetne inteligence za modeliranje mehanizmov
tremorjev*

Vida Groznik

DOKTORSKA DISERTACIJA

PREDANA

FAKULTETI ZA RAČUNALNIŠTVO IN INFORMATIKO

KOT DEL IZPOLNJEVANJA POGOJEV ZA PRIDOBITEV NAZIVA

DOKTOR ZNANOSTI

S PODROČJA

RAČUNALNIŠTVA IN INFORMATIKE



Ljubljana, 2018

distributed under FRI license

IZJAVA

Izjavljam, da sem avtorica dela in da slednje ne vsebuje materiala, ki bi ga kdorkoli predhodno že objavil ali oddal v obravnavo za pridobitev naziva na univerzi ali na drugem visokošolskem zavodu, razen v primerih, kjer so navedeni viri.

— Vida Groznik —

junij 2018

ODDAJO SO ODOBRLI

akad. dr. Ivan Bratko

redni profesor za področje računalništva in informatike

MENTOR IN ČLAN OCENJEVALNE KOMISIJE

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko

dr. Franc Solina

redni profesor za področje računalništva in informatike

PRESEDNIK OCENJEVALNE KOMISIJE

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko

dr. Zvezdan Pirtošek

redni profesor za področje medicine

ZUNANJI ČLAN OCENJEVALNE KOMISIJE

Univerza v Ljubljani, Medicinska fakulteta

PREDHODNA OBJAVA

Izjavljam, da so bili rezultati obravnavane raziskave predhodno objavljeni v recenzirani reviji ali javno predstavljeni v naslednjih primerih:

- [1] V. Groznik, M. Guid, A. Sadikov, M. Možina, D. Georgiev, V. Kragelj, S. Ribarič, Z. Pirtošek, and I. Bratko. Elicitation of neurological knowledge with ABML. In M. Peleg, N. Lavrač, and C. Combi, editors, *Artificial Intelligence in Medicine*, volume 6747 of *Lecture Notes in Computer Science*, pages 14–23, Berlin Heidelberg, 2011. Springer.
- [2] M. Guid, M. Možina, V. Groznik, D. Georgiev, A. Sadikov, Z. Pirtošek, and I. Bratko. ABML knowledge refinement loop : a case study. In L. Chen, et al., editors, *Foundations of intelligent systems*, volume 7661 of *Lecture notes in computer science, Lecture notes in artificial intelligence*, pages 41–50, Heidelberg, 2012. Springer.
- [3] D. Georgiev, V. Groznik, A. Sadikov, M. Možina, M. Guid, V. Kragelj, I. Bratko, S. Ribarič, Z. Pirtošek. Digitalised spirometry and clinical examination based decision support system for differentiating between tremors. In *European Journal of Neurology*, volume 19, pages 298, 2012.
- [4] V. Groznik, M. Guid, A. Sadikov, M. Možina, D. Georgiev, V. Kragelj, S. Ribarič, Z. Pirtošek, and I. Bratko. Elicitation of neurological knowledge with argument-based machine learning. *Artificial Intelligence in Medicine*, 57:133–144, 2013. doi: [10.1016/j.artmed.2012.08.003](https://doi.org/10.1016/j.artmed.2012.08.003)
- [5] A. Sadikov, V. Groznik, J. Žabkar, M. Možina, D. Georgiev, Z. Pirtošek, and I. Bratko. ParkinsonCheck smart phone app.. In T. Schaub, G. Friedrich, B. O’Sullivan, editors, *ECAI 2014 : proceedings, Frontiers in artificial intelligence and applications (Print)*, volume 263, pages 1213–1214, Amsterdam [etc.], 2014. IOS Press
- [6] V. Groznik, A. Sadikov, M. Možina, J. Žabkar, D. Georgiev, and I. Bratko. Attribute visualisation for computer-aided diagnosis : a case study. In *ICHI 2014 : proceedings, 2014 IEEE International Conference on Healthcare Informatics*, pages 294-299, Los Alamitos (California); Washington; Tokyo, 2014. IEEE Computer Society, Conference Publishing Services.

- [7] V. Groznik, M. Možina, J. Žabkar, D. Georgiev, I. Bratko, and A. Sadikov. Development, debugging, and assessment of ParkinsonCheck attributes through visualisation. In A. Briassouli, J. Benois-Pineau, A. Hauptmann. *Health monitoring and personalized feedback using multimedia data*, pages 47–71, Cham [etc.], 2015. Springer. doi: [10.1007/978-3-319-17963-6_4](https://doi.org/10.1007/978-3-319-17963-6_4).
- [8] A. Sadikov, J. Žabkar, M. Možina, V. Groznik, D. Nyholm, and M. Memedi. Feasibility of spirometry features for objective assessment of motor symptoms in Parkinson's disease. In: J. H. Holmes, et al, editors, *Artificial intelligence in medicine*, volume 9105 of *Lecture Notes in Computer Science*, pages 267–276, Cham, 2015. Springer.
- [9] A. Sadikov, V. Groznik, M. Možina, J. Žabkar, D. Nyholm, M. Memedi, I. Bratko and D. Georgiev. Feasibility of spirometry features for objective assessment of motor function in Parkinson's disease. *Artificial intelligence in medicine*, volume 66, 2017. doi: [10.1016/j.artmed.2017.03.011](https://doi.org/10.1016/j.artmed.2017.03.011).

Potrdujem, da sem pridobila pisna dovoljenja vseh lastnikov avtorskih pravic, ki mi dovoljujejo vključitev zgoraj navedenega materiala v pričujočo disertacijo. Potrdujem, da zgoraj navedeni material opisuje rezultate raziskav, izvedenih v času mojega podiplomskega študija na Univerzi v Ljubljani.

POVZETEK

Tremorji so ena najpogostejših motenj gibanja, ki jih povezujemo predvsem z različnimi boleznimi živčevja. Ker obstaja več kot 20 različnih tipov tremorjev, je z vidika pravilnega zdravljenja pomembno, da jih znamo med seboj razločevati. V disertaciji smo se osredotočili na razločevanje med parkinsonskim, esencialnim in mešanim tremorjem, saj se ti najpogosteje pojavljajo.

V disertaciji smo se najprej lotili gradnje diagnostičnega modela za razločevanje med parkinsonskim, esencialnim in mešanim tremorjem na podlagi podatkov kliničnega pregleda, družinske anamneze in digitalne spirografije. Sam proces gradnje diagnostičnega modela je potekal z uporabo argumentiranega strojnega učenja. Ta nam je omogočil, da smo skozi proces elicitacije ekspertovega znanja (v našem primeru nevrologa) zgradili model, ki zajema trinajst pravil, ki so z medicinskega vidika smiselna. Sam postopek elicitacije znanja pa je pripomogel tudi k višji klasifikacijski točnosti končnega diagnostičnega modela v primerjavi z začetnim.

V končnem diagnostičnem modelu so spirografski atributi nastopili v več kot polovici pravil. To nas je motiviralo, da smo izdelali model, ki temelji zgolj na podatkih digitalne spirografije. Za potrebe gradnje razumljivega modela smo najprej zgradili attribute. Namen atributov je bil zajem medicinskega znanja, ki se nanaša tako na spirografijo kot tudi na samo domeno. Na podlagi več kot 500 različnih atributov smo z uporabo logistične regresije zgradili končni model, ki z 90% klasifikacijsko točnostjo razločuje med preiskovanci s tremorjem in tistimi, ki tremorja niso imeli.

Med samo gradnjo atributov smo želeli ugotoviti, kaj naši atributi zaznajo. Tako smo razvili metodo za vizualizacijo atributov na vrstah. Ta nam je pomagala pri gradnji atributov, poleg tega pa je uporabna tudi za vizualno razlago odločitev diagnostičnega modela. Samo vizualizacijo in posledično diagnostični model smo evalvirali s pomočjo treh neodvisnih ekspertov. Izkazalo se je, da sta tako sam diagnostični model kot tudi

vizualizacija smiselna in zajameta medicinsko znanje. Končni diagnostični model je vgrajen v prosto dostopni mobilni aplikaciji PARKINSONCHECK.

Ključne besede: umetna inteligenca, vizualizacija atributov, tremor, digitalna spirografija, parkinsonova bolezen, esencialni tremor

ABSTRACT

Tremors are one of the most common movement disorders primarily associated with various neurological diseases. Since there are more than 20 different types of tremors, differentiation between them is important from the treatment point of view. In the thesis, we focus on differentiation between three of the most common tremors: Parkinsonian, essential and mixed type of tremor.

Our first goal was to build a diagnostic model for distinguishing between Parkinsonian, essential and mixed type of tremors, based on clinical examination data, family history and digital spirometry. The process of building a model was carried out using argument-based machine learning which enabled us to build a decision model through the process of knowledge elicitation from the domain expert (in our case from a neurologist). The obtained model consists of thirteen rules that are medically sensible. The process of knowledge elicitation itself contributed to the higher classification accuracy of the final model in comparison with the initial one.

In the final diagnostic model, attributes derived from the spirometry were included in more than half of the rules. This motivated us to build a model based solely on the digital spirometry data. For the needs of constructing an understandable model, we first built several attributes which represented domain medical knowledge. We have built more than 500 different attributes which were used in a logistic regression to construct the final diagnostic model. The model is able to distinguish subjects with tremors from those without tremors with 90% classification accuracy.

During the process of attribute construction, we wanted to know what our attributes were detecting. Thus, we have developed a method for attribute visualisation on series. The method not only helped us with attribute construction, but it is also useful for visual interpretation of the diagnostic model's decisions. The visualisation method and consequently the decision model were evaluated with the help of three independent

neurology experts. The results show that both the diagnostic model and the visualisation are meaningful and cover medical knowledge of the domain. The final diagnostic model is built into the freely available PARKINSONCHECK mobile application.

Key words: artificial intelligence, attribute visualisation, tremor, digital spirometry, Parkinson's disease, essential tremor

ZAHVALA

Najprej bi se rada zahvalila mentorju, akad. prof. dr. Ivanu Bratku, za usmerjanje in konstruktivne pogovore, ki so me pripeljali do rezultatov, ki jih predstavljam v tem delu.

Zahvaljujem se prof. Bratku in prof. dr. Zvezdanu Pirtošku, ki sta omogočila in vzpodbujala sodelovanje Laboratorija za umetno inteligenco s Kliničnim oddelkom za bolezni živčevja Nevrološke klinike v Ljubljani. To sodelovanje je bilo ključno za razvoj tega dela.

Posebno zahvalo namenjam doc. dr. Aleksandru Sadikovu za ves čas, ki ga je žrtvoval za razpravljanje o idejah in za reševanje zapletov, na katere sem naletela med raziskovalnim delom. Hvala tudi za nenehno priganjanje in oporo v težkih trenutkih, kar mi je preprečilo, da bi "vrgla puško v koruzo".

Nadalje se zahvaljujem Martinu, Dejanu in Mateju za večurno "ciklanje" algoritma ABML in argumentiranje podatkov o tremorjih. Hvala celotni ekipi projekta Parkinson-Check - Aleksandru, Martinu, Juretu, Dejanu in Gurujevcem, kajti le s skupnimi močmi smo lahko dosegli zavidljive rezultate pri razpoznavanju tremorjev, ki so ugledali luč sveta tudi v dejanskem produktu. Že od vsega začetka aplikacija PARKINSONCHECK sproža zanimanje tako v znanstveni kot tudi v strokovni sferi.

Hvala dr. Dietrichu Haubenbergerju, doc. dr. Dagu Nyholmu in dr. Dejanu Georgievu za pomoč pri evalvaciji rezultatov vizualizacije napak. Hvala Alešu in Marku za predelavo zdravniške aplikacije, s katero je bila evalvacija hitrejša in učinkovitejša.

Iskreno se zahvaljujem Jerki za vse dragocene nasvete in predvsem veliko psihično oporo v trenutkih, ko sem to najbolj potrebovala. Zahvaljujem se ji tudi, ker si je vzela kar nekaj časa za iskanje in lovljenje slovničnih ter tipkarskih napak, ki so se vtihotapile v besedilo doktorata in se spretno skrivale pred mojimi očmi.

Tijani, Alenki in Anji sem hvaležna za njihovo razumevanje, ker sem jih občasno postavila nižje na lestvici prioritete. Nad vse cenim vse ure, ki so mi jih namenile, ko sem to potrebovala.

Zahvaljujem se še Timoteju, Aljažu, Cirilu, Tatjani in vsem ostalim, ki so na svoj način pripomogli k temu, da sem uspešno zaključila doktorski študij.

Nenazadnje se zahvaljujem svojim staršem za vso vzpodbudo in podporo, ki so mi jo dajali celo življenje in mi omogočili, da sem sledila svojim željam. Hvala.

— Vida Groznik, Ljubljana, junij 2018.

KAZALO

<i>Povzetek</i>	<i>i</i>
<i>Abstract</i>	<i>iii</i>
<i>Zahvala</i>	<i>v</i>
1 <i>Uvod</i>	1
1.1 Motivacija	2
1.2 Prispevki k znanosti	3
2 <i>Izhodišča in predznanje</i>	5
2.1 Tremor	6
2.2 Digitalna spirografija	7
3 <i>Izgradnja modela za razločevanje tremorjev na podlagi kliničnih podatkov in digitalne spirografije</i>	13
3.1 Metodologija	14
3.1.1 Argumentirano strojno učenje	14
3.1.2 Obravnavanje komorbidnosti	14
3.1.3 Interakcija med ekspertom in metodo ABCN2	15
3.1.4 Elicitacija znanja z uporabo metode ABML	17
3.2 Predstavitev domene	21
3.3 Rezultati in evalvacija modela	22
3.3.1 Podatki	22
3.3.2 Kvantitativna primerjava rezultatov	23
3.3.3 Kvalitativna primerjava začetnega in končnega modela	25

3.3.4	Analiza napak	28
3.3.5	Diskusija	29
4	<i>Izgradnja modela za razločevanje tremorjev na podlagi digitalne spirografije</i>	31
4.1	Ozadje	32
4.2	Predstavitev domene in predznanje	32
4.2.1	Predznanje	34
4.3	Predstavitev atributov	35
4.3.1	Atributi za odkrivanje asimetrij	35
4.3.2	Atributi za odkrivanje ekstremov	36
4.3.3	Atributi za odkrivanje napak/odstopanj	37
4.3.4	Atributi frekvenčne analize	38
4.3.5	Atributi za analizo sprememb v radiju	38
4.3.6	Atributi izpeljani iz večih testov	38
4.4	Rezultati in evalvacija modela	39
4.4.1	Podatki	39
4.4.2	Opis eksperimenta	39
4.4.3	Evalvacija rezultatov	41
4.4.4	Predstavitev izbranega modela	43
4.5	Diskusija	44
5	<i>Metoda za vizualno razlago diagnoz</i>	47
5.1	Metodologija	48
5.1.1	Algoritem	60
5.2	Podatki in evalvacija metode	62
5.2.1	Nastavitve parametrov	63
5.2.2	Zasnova eksperimenta	64
5.3	Rezultati	68
5.3.1	Analiza optimalnosti algoritma	74
5.4	Diskusija poskusa z eksperti	76
6	<i>Zaključki</i>	91
6.1	Zaključki	92
6.2	Prispevki k znanosti	94

<i>A</i>	<i>Začetni atributi, uporabljeni za izgradnjo modela za razločevanje tremorjev na podlagi kliničnih podatkov in digitalne spirografije</i>	<i>95</i>
<i>B</i>	<i>Primer sledi izvajanja algoritma</i>	<i>101</i>
<i>C</i>	<i>Navodila za izvedbo eksperimenta</i>	<i>131</i>
	<i>Literatura</i>	<i>137</i>

Uvod

1.1 Motivacija

Tremor je nehoteno gibanje enega ali večih delov telesa ter predstavlja eno najpogostejših motenj gibanja. Je simptom, ki ga povezujemo predvsem z različnimi boleznimi živčevja, med drugim s parkinsonovo boleznijo. Ker obstaja več vrst tremorjev, je z vidika pravičnega zdravljenja pomembno, da razumemo, kako se odražajo in jih znamo med seboj ločevati.

Že v začetku devetdesetih let prejšnjega stoletja so za pomoč pri razločevanju tremorjev začeli uporabljati metodo spirografije. Naloga preiskovanca je bila, da na list papirja s pisalom nariše levo-sučno spiralo z obema rokama. Pri tem je zdravnik opazoval risanje (hitrost, obotavljanja) in nato še končni rezultat. A to metodo so poznali le redki posamezniki, ki so se s tem ukvarjali večinoma raziskovalno. Z razvojem računalništva se je pojavila digitalna spirografija, ki je omogočala shranjevanje poteka risanja spirale in analizo teh podatkov. Kljub temu so, predvsem zaradi nezaupanja v novo tehnologijo, zdravniki še vedno uporabljali standardne in uveljavljene postopke za klinični pregled in postavitve diagnoze, ki pa niso vsebovali spirografskega testiranja. Ker je spirografija ostala v ozkih, raziskovalnih krogih, ni presenetljivo, da se v dvajsetih letih ni razvil sistem digitalne spirografije, ki bi omogočal, da bi na podlagi narisanih spiral in po potrebi dodatnih podatkov iz kliničnega pregleda avtomatsko razločeval med različnimi tipi tremorjev. Da bi izboljšali zaupanje v takšen sistem, bi bilo priporočljivo, da bi nudil zdravnikom tudi razlago za svojo odločitev. Ena od možnosti je vizualizacija odločitve na narisanih spiralah.

Vizualizacija je zelo močno orodje, saj nam omogoča grafični prikaz atributov in najdenih vzorcev, ki so jih zgradili snovalci sistema za podporo pri odločanju. Da bi zdravniki sprejeli takšen sistem in vanj zaupali, morajo biti vizualizacije za njih smiselne, predvsem pa se morajo skladati z njihovim medicinskim znanjem o domeni. Zato je nujno, da snovalci pri gradnji atributov, na podlagi katerih bodo sprejemale odločitve znotraj sistema, sodelujejo z zdravniki oz. domenskimi eksperti. To pomeni, da morajo zajeti čim več njihovega znanja o domeni, predvsem tistega, ki se ga niti sami ne zavedajo (način razmišljanja). Rekli bi lahko, da bi vizualizacija atributov, na podlagi katerih sistem (pa tudi zdravnik) poda svojo odločitev, pravzaprav vizualizirala znanje in razmišljanje zdravnikov.

Pri modeliranju je pogosto prisotna želja, da bi opisali domeno z uporabo atributov, saj nam je rezultat takšnega pristopa domač in razumljiv. Takšen opis domene pa nam

omogoča, da vizualizacijo uporabimo vsaj za:


- odkrivanje napak v atributih,
- odkrivanje novega znanja,
- razlago klasifikacije.

Izpostaviti je potrebno, da je bila v preteklosti vizualizacija morda nekoliko zanemarjena oz. se ji ni pripisovalo večjega pomena. V zadnjih letih pa se je izkazala za pomembno orodje pri sprejemanju različnih odločitev. Kot primer lahko navedemo njen velik pomen pri sprejemanju poslovnih odločitev, kjer z vizualizacijo poslovnih rezultatov, trendov in analiz poslovanja odločevalci dobijo popolnoma drugačen vpogled v poslovanje družbe. Podobno vzporednico lahko potegnemo tudi s sprejemanjem sistemov za podporo odločanju, ki jim uporabniki (v našem primeru zdravniki) niso zaupali. Takšne sisteme so videli kot črne škatle, ki jim podaš vhodne podatke, le-ta pa vrne rezultat. Če pa uporabnik ne razume, kako je sistem prišel do takšnega rezultata, dvomi v njegovo odločitev. Ravno zaradi tega tovrstni sistemi v začetku niso bili širše sprejeti, načrtovalci sistemov pa so bili primorani razmišljati, kako povečati zaupanje uporabnikov v tovrstne sisteme. Razlaga oz. vizualizacija odločitev je zato ena najpomembnejših stvari, ki jih mora sistem vsebovati. Z uvedbo le-teh se je zaupanje v sisteme občutno povečalo, v nekaterih primerih celo preveč. Dogaja se, da uporabniki ne zaupajo več svojemu čutu (logiki), temveč že slepo verjamejo rezultatu, ki ga vrne takšen sistem. Zato je potrebna pazljivost pri zasnovi tovrstnih sistemov, poraja pa se tudi vprašanje, ali je zaradi tega potrebna gradnja bolj konzervativnih odločitvenih modelov. Zato je še toliko bolj pomembno, da je razlaga odločitve, ki jo poda sistem, smiselna in razumljiva za uporabnika. Konec koncev razlaga predstavlja povratno informacijo računalnika ekspertu, zakaj se je odločil tako, kot se je.

1.2 Prispevki k znanosti

Prispevki k znanosti, ki so zajeti v disertaciji:

- Računalniški diagnostični model za ločevanje med esencialnim, parkinsonskim in mešanim tremorjem;
- Metoda za vizualno razlago diagnoz, izpeljanih iz vrst.



Metoda za gradnjo računalniškega diagnostičnega modela za ločevanje med različnimi tipi tremorjev je predstavljena v poglavju 3, kjer je predstavljen tudi končni model. Model vsebuje podatke kliničnega pregleda in digitalne spirografije. Na podlagi znanja, ki smo ga pridobili pri gradnji tega modela, smo zgradili računalniški diagnostični model, ki uporablja izključno podatke, ki jih dobimo z uporabo digitalne spirografije. Model in atributi, ki smo jih zgradili za uporabo v diagnostičnem modelu, so predstavljeni v poglavju 4.

Diagnoze, ki jih postavi računalniški diagnostični model nato razložimo z uporabo metode za vizualno razlago diagnoz. Metoda deluje za attribute, ki so definirani na vrstah. Metoda, evalvacija metode in njene omejitve so predstavljene v poglavju 5.

Deli doktorske disertacije so bili predstavljeni na različnih mednarodnih znanstvenih konferencah [1–6], v revijah [7, 8], tehničnem poročilu [9] ter poglavju v knjigi [10].

Izhodišča in predznanje

2.1 Tremor

Tremor je motorična motnja, ki se navzven izraža v obliki nehotene, večinoma ritmične tresavice določenega dela telesa. Največkrat prizadane roke, pogosti pa so tudi tremorji glave, nog, glasilk in telesa. Vzroki in mehanizmi nastanka različnih vrst tremorjev so kljub velikemu številu raziskav še vedno nejasni.

Poznamo več vrst tremorjev, ki jih ločujemo glede na okoliščine pojavljanja [11, 12, 14]:

Tremor v mirovanju (*angl.* rest tremor) se pojavlja, ko je prizadeti del telesa sproščen in podprt s podlago (npr. roke počivajo v naročju). Zanj je značilna frekvenca tresenja med 3 in 6 Hz [13] in se tipično zmanjša ali popolnoma izgine med izvajanjem hotenih gibov. Tremor se začne asimetrično, njegovo gibanje pa spominja na "štetje denarja" (*angl.* pill-rolling tremor). Tremor v mirovanju je značilen za parkinsonovo bolezen.

Akcijski tremor (*angl.* action tremor) se pojavi ob izvajanju zavestnih mišičnih krčenj in vključuje posturalni in kinetični tremor.

Posturalni tremor (*angl.* postural tremor) se pojavlja ob vzdrževanju določenega položaja prizadetenga uda (npr. držanje iztegnjenih rok pred seboj) [11]. Zanj je značilna frekvenca med 4 in 12 Hz, pojavlja pa se pri sindromu esencialnega tremorja. Kljub temu, da se lahko pojavi tudi pri bolnikih s parkinsonovo boleznijo, povezanost posturalnega tremorja s tremorjem v mirovanju ni jasna. [14]

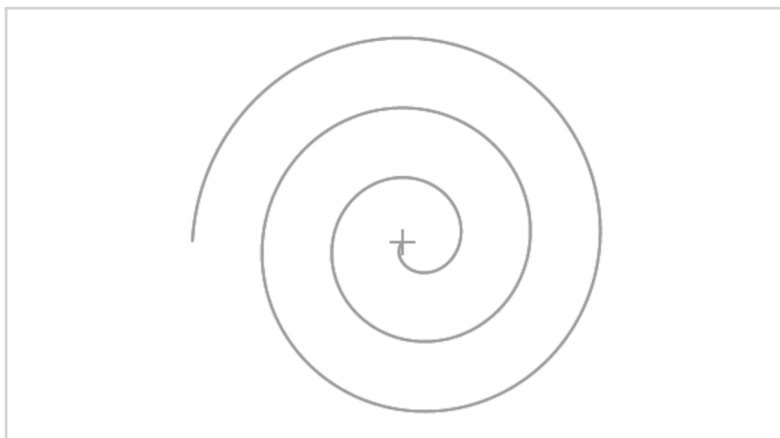
Kinetični tremor (*angl.* kinetic tremor) se pojavlja med prostovoljnim gibanjem udov. Kinetični tremor ima lahko prisotno tudi intencijsko komponento, ki se izraža v povečani amplitudi tremorja tik preden se prizadeti ud približa določeni tarči (npr. pri testu prst-nos-prst).

Opravlno-specifični kinetični tremor (*angl.* task-specific tremor) se pojavlja, ko oseba izvaja ali poskuša izvesti točno določeno nalogo. Najbolj pogosto se pojavlja tremor pri pisanju. Imenujemo ga primarni tremor pri pisanju (*angl.* primary writing tremor). [15].

Izometrični tremor (*angl.* isometric tremor) se pojavlja kot rezultat mišičnega krčenja pri pritisku ob mirujoč objekt. [12]

2.2 Digitalna spirografija

Z metodo risanja spirale merimo motorične sposobnosti preiskovanca ob risanju Arhimedove spirale (Slika 2.1), katere prednost je možnost testiranja motorike v vseh smereh. Kot metodo za ocenjevanje tremorjev jo priporoča tudi Združenje za motnje gibanja (*angl.* Movement Disorder Society) [11].



Slika 2.1: Arhimedova spirala.

Dokazano je bilo, da rezultati spirografije korelirajo z oceno UPDRS (*angl.* Unified Parkinson's Disease Rating Scale), kar je lahko koristno za ocenjevanje resnosti znakov parkinsonove bolezni [16].

Uporaba metode spirografije je zelo razširjena. Uporablja se jo za oceno akinezije pri parkinsonovi bolezni [17–19], diskinezije povzročene z zdravili [20], tremorjev pri multipli sklerozi [21–23] in esencialnega tremorja [24, 25]. Prav tako se lahko uporablja pri oceni učinkovitosti globoke možganske stimulacije [26].

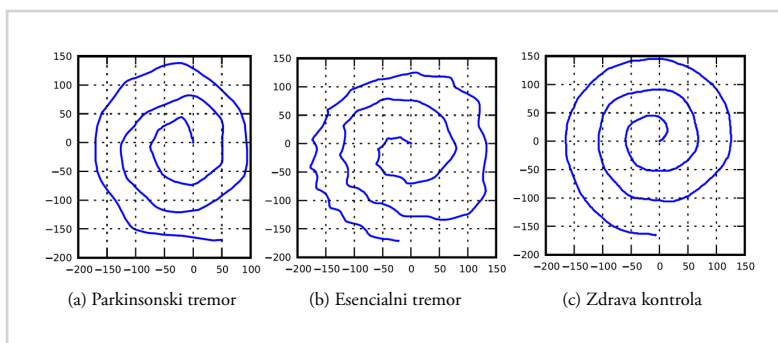
Tradicionalno se spirografija izvaja z risanjem spirale na listu papirja, narisane spirale pa se nato vizualno oceni. Zaradi pomanjkanja jasnih smernic za ocenjevanje in interpretacijo narisanih spiral, je klasična spirografija zelo subjektivna. Najbolj priznane smernice sta v 90. letih postavila Bain in Findley [27]. Tradicionalna metoda spirografije nam, razen slike spirale, ne omogoča zbiranja dodatnih podatkov o risanju kot npr. časovnega poteka risanja točk, pritiska pisala ipd.

V preteklosti so že bili poskusi objektivnega merjenja amplitude tremorja z analizo

skeneranih spiral, narisanih na papir [28, 29]. Kljub temu glavne pomanjkljivosti še vedno ostajajo. Ker ne merimo časovne komponente točk, ne moremo analizirati pomembnih komponent risanja, kot npr. hitrost risanja in njeno spreminjanje, frekvenca risanja ipd. Tovrstne pomanjkljivosti uspešno odpravi novejša metoda - digitalna spirografija.

Digitalna spirografija uporablja digitalne sisteme za snemanje poteka risanja spirale [30, 31]. Sistem sestavljajo osebni računalnik, na katerem teče namenska programska oprema, grafična tablica in elektronsko pisalo za grafično tablico. Omogoča nam snemanje in shranjevanje spiral v obliki časovnih vrst, kjer imamo za vsak par (x_i, y_i) podano še časovno komponento t_i . To nam omogoča podrobnejšo analizo spiral - npr. frekvenčno analizo. Na sliki 2.2 so predstavljeni tipični primeri spiral oseb s parkinsonskim tremorjem, esencialnim tremorjem in zdravih kontrol.

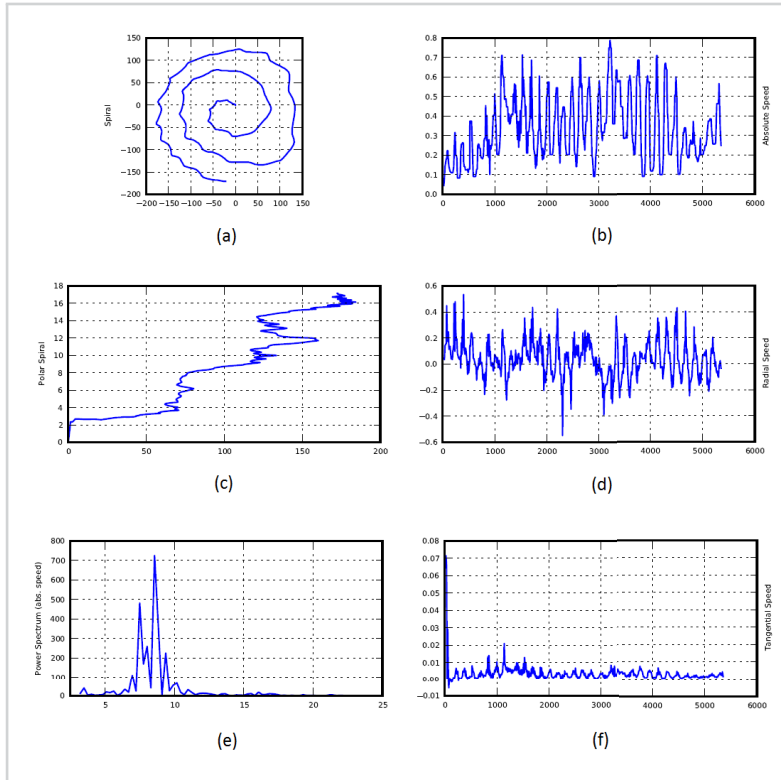
Slika 2.2: Tipični primeri spiral preiskovancev s (a) parkinsonskim tremorjem, (b) esencialnim tremorjem in (c) zdravih kontrol.



Tako dobljene časovne vrste nato ročno analizira ekspert in postavi diagnozo, saj sistemi, ki bi vsebovali odločitveni model za diagnostiko, ne obstajajo. Sistem, ki ga uporabljajo na Univerzitetnem kliničnem centru Ljubljana, na podlagi vhodne spirale izriše šest osnovnih grafov (slika 2.3), na podlagi katerih skuša nevrolog oceniti tip tremorja in njegovo resnost.

Graf (a) na sliki 2.3 prikazuje narisan spiralo v kartezičnem koordinatnem sistemu. Z njim je možno določiti nekatere tipične nepravilnosti v spirali, kot je stisnjenost na eni strani, kar je značilno za parkinsonski tremor ali nazobčanost spirale, kar je tipično za esencialni tremor. Žal pa so le redke spirale takšne, da lahko ekspert samo na podlagi tega grafa napove tip tremorja.

Graf (b) na sliki 2.3 prikazuje absolutno hitrost risanja spirale, ki jo izračunamo tako,



Slika 2.3: Šest osnovnih grafov, izpeljanih iz spirale, pridobljene z uporabo digitalne spirografije: (a) spirala v kartezijskem koordinatnem sistemu, (b) absolutna hitrost med risanjem spirale, (c) spirala v polarnem koordinatnem sistemu, (d) radialna hitrost, (e) frekvenčna analiza, in (f) tangencialna hitrost.

da vzamemo dve sosednji točki (X_i, y_i, t_i) in $(x_{i+1}, y_{i+1}, t_{i+1})$ ter izračunamo razdaljo med njima, ki jo delimo z razliko v času:

$$v = \frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{t_{i+1} - t_i}.$$

Graf (c) na sliki 2.3 prikazuje narisan Arhimedovo spiralo v polarnem koordinatnem sistemu. Arhimedova spirala je v polarnih koordinatah predstavljena v obliki monotono naraščajoče premice, kar olajša primerjavo idealne spirale z narisano.

Graf (d) na sliki 2.3 prikazuje radialno hitrost, ki meri spremembo radija v času in jo lahko povezujemo z gladkostjo spirale. V idealnih razmerah bi moral radij skozi čas monotono naraščati, medtem ko bi morala biti radialna hitrost konstantna. Pri spiralah, ki so tipične za osebe z esencialnim tremorjem, je v grafu radialne hitrosti razviden žagast vzorec. Radialno hitrost preračunamo z enačbo:

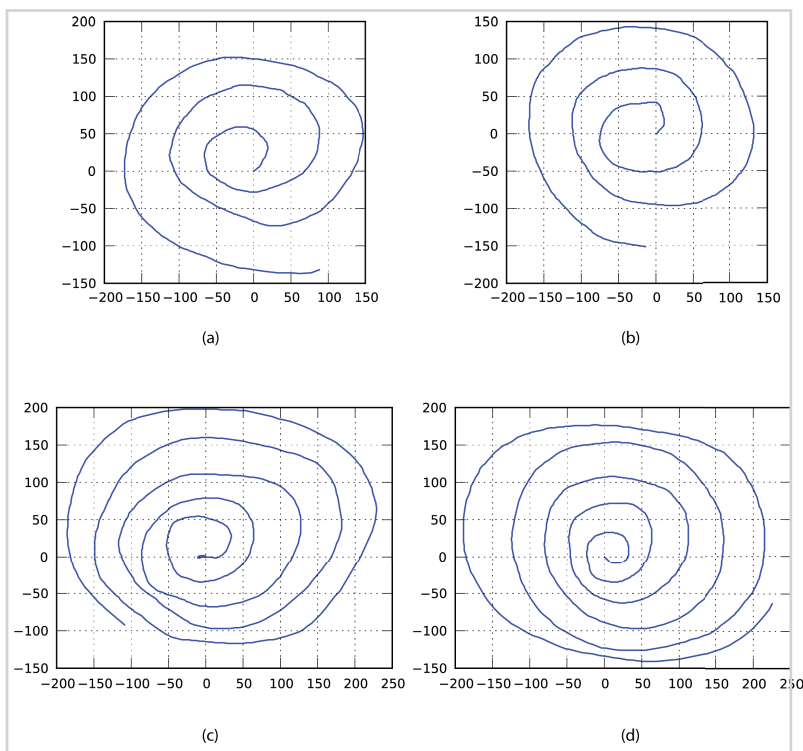
$$v_r = \frac{r_{i+1} - r_i}{t_{i+1} - t_i},$$

Graf (e) na sliki 2.3 prikazuje spekter frekvenčne analize absolutne hitrosti. S tem grafom se lahko identificira ritmično gibanje v različnih frekvenčnih območjih, kar je značilno za nekatere tremorje, npr. esencialni tremor.

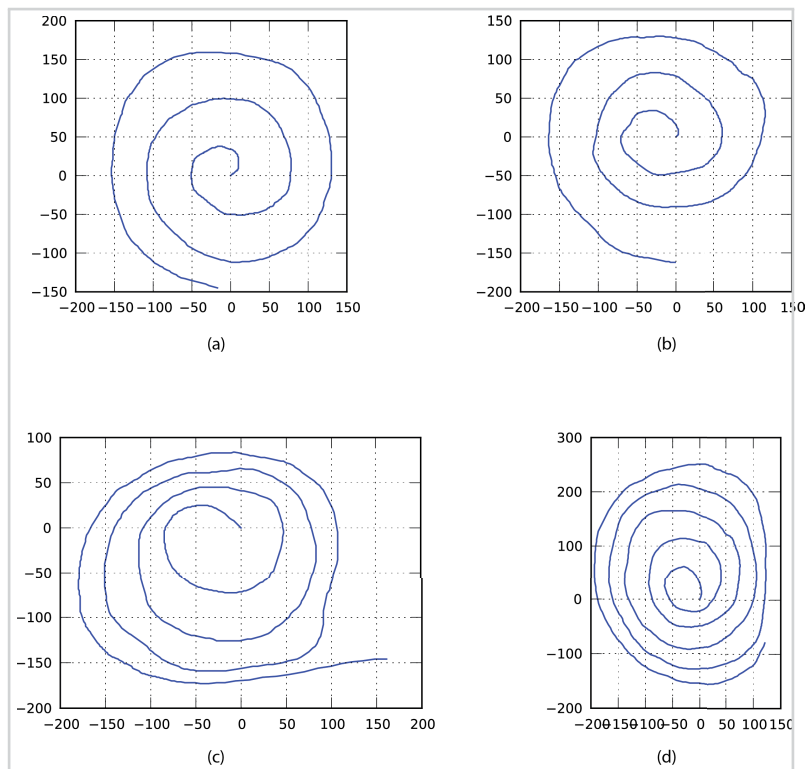
Graf (f) na sliki 2.3 prikazuje kotno hitrost, torej spremembo kota skozi čas, ki se praviloma med risanjem zmanjšuje:

$$\omega = \frac{\varphi_{i+1} - \varphi_i}{t_{i+1} - t_i}.$$

Na prvi pogled se bralcu morda lahko zazdi, da je problem popolnoma enostaven, saj ob pregledu tipičnih spiral na sliki 2.2 že laiku ni težko razločiti med posameznimi spiralami in njihovimi vzorci. Žal pa so takšne tipične spirale zelo redke. Velikokrat se namreč lahko zgodi, da oseba, ki je zdrava, nariše spiralo, ki na prvi pogled spominja na tipično spiralo osebe s parkinsonskim tremorjem in obratno. Oseba, ki ima parkinsonski tremor lahko nariše zelo lepo spiralo (razvidno iz slike 2.4). Glavni razlog za to je lahko v tem, da ima ta oseba le tremor v mirovanju, ko pa začne premikati roko (torej risati spiralo), tremor začasno izzveni in je spirala na pogled brez napak. Podobne primere najdemo pri osebah z esencialnim tremorjem (slika 2.5).



Slika 2.4: Primer atipične spirale, ki jo je narisal preiskovanec s parkinsonskim tremorjem.



Slika 2.5: Primer atipične spirale, ki jo je narisal preiskovanec z esenci- alnim tremorjem.

*Izgradnja modela za
razločevanje tremorjev na
podlagi kliničnih podatkov in
digitalne spirografije*

3.1 Metodologija

V tem poglavju so opisane bistvene sestavine našega pristopa: argumentirano strojno učenje (angl. argument-based machine learning - ABML), obravnavanje komorbidnosti ter interakcija med strokovnjakom in učnim programom (izpopolnitveni cikel ABML).

3.1.1 Argumentirano strojno učenje

ABML [32] je strojno učenje, nadgrajeno s koncepti argumentacije. Pri ABML se argumenti uporabljajo kot sredstva za izvabljanje znanja ekspertov prek razlag učnih primerov. Prednost pristopa je v tem, da se morajo eksperti osredotočiti samo na en specifičen primer naenkrat in posredovati znanje, ki je relevantno za ta primer. Uporabili smo metodo ABCN2 [32] - z argumenti razširjena metoda CN2 [33]. Metoda se nauči množice neurejenih verjetnostnih pravil iz primerov s podanimi argumenti (argumentirani primeri).¹

3.1.2 Obravnavanje komorbidnosti

Obravnavana problemska domena ima tri razrede: esencialni tremor (ET), parkinsonski tremor (PT) in mešani tremor (MT). Ker za razred MT velja, da sta prisotna tako esencialni kot tudi parkinsonski tremor (komorbidnost), je pri učenju pravil težko razlikovati med ET in MT ter med PT in MT. Rešitev tega problema je prevedba trirazrednega problema v dva dvorazredna problema.

V prvem primeru ET in MT združimo v razred EMT (*esencialni mešani*). Za vse preiskovance v tem razredu velja, da imajo nekaj znakov esencialnega tremorja. Pravila za EMT bodo tako v svojih pogojih vsebovala značilnosti, ki označujejo esencialni tremor, ne pa parkinsonskega tremorja. Čeprav EMT vsebuje paciente s parkinsonskim tremorjem (znotraj razreda MT), pa značilnosti, ki označujejo parkinsonski tremor, ne bodo vključene v pravila, saj niso relevantna za razlikovanje med razredoma EMT in PT.

V učnem problemu z EMT in PT se naučimo množice pravil zgolj za razred EMT. Pri tem se izognemo učenju pravil za PT, saj razred EMT vsebuje tudi paciente s parkinsonskim tremorjem; zaradi tega je malo verjetno, da bi učenje pravil za PT oblikovala

¹Več informacij o metodah ABML in ABCN2 je podanih v [32] in na spletni strani www.ailab.si/martin/abml.

dobra in razumljiva pravila za parkinsonski tremor.

Drugi dvorazredni problem je analogen prvemu, kjer PT in MT združimo v razred PMT (*parkinsonski mešani*). V tem primeru je rezultat učenja množica pravil za razred PMT.

Za diagnosticiranje novih primerov z induciranimi pravili potrebujemo mehanizem, ki omogoča sklepanje o novih primerih. V ta namen smo razvili tehniko, s katero lahko nove primere klasificiramo v enega izmed treh možnih razredov glede na inducirana pravila za razreda EMT in PMT.

Naj bo e primer, ki ga želimo klasificirati. Naj bo R_{EMT} množica pravil, ki pokrivajo primer e in napovedujejo razred EMT. Podobno pravila v R_{PMT} napovedujejo razred PMT in pokrivajo e . Naj bo verjetnost $P(e = EMT)$ (kjer je $e = EMT$ okrajšava za $class(e) = EMT$) verjetnost napovedanega razreda "najboljšega" pravila (z največjo verjetnostjo napovedanega razreda) iz R_{EMT} in $P(e = PMT)$ verjetnost napovedanega razreda "najboljšega" pravila iz R_{PMT} . Če je R_{EMT} (ali R_{PMT}) prazen, potem je $P(e = EMT) = 0$ (ali $P(e = PMT) = 0$). Naša metoda uporablja naslednjo formulo za izračun verjetnosti za tri razrede:

$$\begin{aligned} P(e = ET) &= 1 - P(e = PMT) \\ P(e = PT) &= 1 - P(e = EMT) \\ P(e = MT) &= P(e = EMT) + P(e = PMT) - 1, \end{aligned}$$

kjer $P(e = ET)$, $P(e = PT)$ in $P(e = MT)$ ustrezajo napovedanim verjetnostim razredov ET, PT in MT. V posebnih primerih, kjer je $P(e = EMT) + P(e = PMT) < 1$, je verjetnost $P(e = MT)$ negativna. To se zgodi v primeru, ko nobeno izmed pravil ne pokriva primera e (R_{EMT} ali R_{PMT} je prazen). V takšnih primerih nastavimo verjetnost $P(e = MT)$ na nič in nato normaliziramo verjetnosti $P(e = ET)$, $P(e = PT)$ in $P(e = MT)$ tako, da je njihova vsota ena. Primer e je klasificiran v razred z največjo napovedano verjetnostjo.

Opisani mehanizem je bil uporabljen v zanki za elicitacijo znanja, ki je predstavljena v naslednjem poglavju.

3.1.3 Interakcija med ekspertom in metodo ABCN2

Ker ne moremo pričakovati, da bi eksperti podali argumente za celotno učno množico, smo uporabili naslednjo zanko, s pomočjo katere smo izbrali *kritične primere*, ki jih je

ekspert razložil.

1. korak: Nauči se hipoteze z uporabo danih podatkov in metode ABCN2.
2. korak: Poišči "najbolj kritičen" primer in ga prikaži ekspertu. Če kritičnega primera ni mogoče najti, se postopek ustavi.
3. korak: Ekspert razloži dani primer; razlaga se za potrebe metode predstavi v obliki argumentov in doda k učnemu primeru.
4. korak: Vrni se na 1. korak.

Če želimo zaključiti postopek, moramo odgovoriti na naslednji vprašanji:

- (a) Kako izberemo kritične primere?
- (b) Kako lahko zagotovimo, da bomo dobili vse potrebne informacije za izbrani primer?

Prepoznavanje kritičnih primerov

Kritičen primer je primer, ki ga s trenutno hipotezo ni mogoče dobro razložiti. Ker naša metoda poda napovedi razredov v obliki verjetnosti, bomo najprej identificirali "najbolj problematičen" primer - to je primer z največjo verjetnostjo napake. Za ocenjevanje verjetnostne napake smo uporabili k -kratno prečno preverjanje, ki smo ga ponovili n -krat (npr. $n = 4, k = 5$), tako da se vsak primer testira n -krat. Kritični primer je izbran v skladu z naslednjima praviloma:

1. Če je problematičen primer iz razreda MT, le-ta postane kritičen primer.
2. Če je problematičen primer iz razreda ET (ali PT), bomo s to metodo ugotovili, katero pravilo, ki napoveduje PMT (ali EMT), je odgovorno za napačno klasifikacijo primera. Ker to pravilo pokriva naš problematičen primer, je zelo verjetno, da je le-to pravilo slabo. Zato bo kritičen primer postal primer iz razreda PT ali MT (oz. ET ali MT), ki ga pokriva to problematično pravilo. Nato bomo prosili eksperta, da razloži, zakaj je bil ta preiskovanec diagnosticiran s parkinsonskim tremorjem (ali esencialnim tremorjem). Razlaga tega primera načeloma povzroči, da bo algoritem ABCN2 problematično pravilo nadomestil z boljšim pravilom za razred PMT (ali EMT), za katerega upamo, da ne bo pokrivalo problematičnega primera.

Ali so ekspertovi argumenti dobri ali jih je treba izboljšati?

Sledi podrobnejši opis tretjega koraka zgornjega algoritma:

Korak 3a: Razlaga kritičnega primera. Če je primer iz razreda MT, prosimo eksperta, da razloži njegove znake, ki kažejo na parkinsonski in esencialni tremor (kar se zgodi, ko je problematični primer iz razreda MT) ali da razloži samo eno od bolezni. V drugih dveh primerih (ko je primer iz razreda ET ali PT) ekspert razloži samo znake, ki so relevantni za tisti razred, kateremu pripada primer. Ekspert nato artikulira množico razlogov, ki nakazujejo razred primera. Za podani argument velja, da naj vsebuje minimalno število razlogov. S tem se izognemo preveč specifičnim argumentom.

Korak 3b: Dodajanje argumentov k primeru. Argument se poda v naravnem jeziku, ki ga je potrebno prevesti v domenski opisni jezik (atribute). Če argument omenja koncepte, ki v domeni še niso prisotni, je potrebno te koncepte vključiti v domeno (kot nove attribute), preden se argument doda k primeru.

Korak 3c: Odkrivanje protiprimerov. Protiprimeri se uporabljajo za ugotavljanje, ali argument zadostuje za uspešno razlago kritičnega primera ali ne. Če ne zadostuje, ABCN2 izbere protiprimer, čigar razred je nasproten razredu kritičnega primera, vendar ga pravilo, inducirano iz podanih argumentov, pokriva.

Korak 3d: Izboljševanje argumentov s protiprimeri. Strokovnjak mora revidirati svoj prvotni argument ob upoštevanju protiprimera.

Korak 3e: Vrni se na korak 3c, če je najden protiprimer.

3.1.4 Elicitacija znanja z uporabo metode ABML

V našem primeru je postopek elicitacije znanja potreboval 19 iteracij. Med postopkom je bilo v domeno vključenih 17 novih atributov. Vsi ti atributi so bili izpeljani iz prvotnih atributov in temeljijo na razlagah, ki jih je podal ekspert. Novi atributi so predstavljeni v tabeli 3.1.

Argumentacija primerov iz razreda ET/PT

Na začetku postopka so bili uporabljeni samo prvotni atributi brez podanih argumentov. Primer, ki smo ga označili z E.2 (klasificiran kot ET v naši bazi podatkov), je

Tabela 3.1: Novi atributi, ki so bili izpeljani iz prvotnih atributov med postopkom elicitaricije znanja s pripadajočimi stopnjami manjkajočih vrednosti. Ti atributi temeljijo na razlagah, ki jih je podal ekspert.

Atribut	Opis	Število manjkajočih vrednosti n (%)
Sim.tremor.start	Obojestranski začetek tremorja	0 (0,00)
Diff.age.tremor.duration	Razlika med starostjo preiskovanca in trajanjem tremorja	14 (12,28)
Diff.disease.tremor.duration	Razlika med trajanjem bolezni in trajanjem tremorja	59 (51,75)
Sim.resting.tremor.up	Obojestransko enako izražen tremor v mirovanju - zgornje okončine	6 (5,26)
Sim.postural.tremor.up	Obojestransko enako izražen posturalni tremor zgornjih okončin	10 (8,77)
Sim.rigidity.up	Obojestransko enako izražena rigidnost zgornjih okončin	7 (6,14)
Sim.bare.speed.time	Obojestransko enaka transformacija hitrost-čas - prostoročno risanje	21 (18,42)
Sim.bare.radius.angle	Obojestransko enaka transformacija radij-kot - prostoročno risanje	23 (20,18)
Sim.template.speed.time	Obojestransko enaka transformacija hitrost-čas - risanje po predlogi	45 (39,47)
Sim.template.radius.angle	Obojestransko enaka transformacija radij-kot - risanje po predlogi	49 (42,98)
Bradykinesia	Prisotnost bradikinezije vsaj na eni strani	5 (4,40)
Resting.tremor.up	Prisotnost tremorja v mirovanju vsaj na eni strani	6 (5,26)
Postural.tremor.up	Prisotnost posturalnega tremorja vsaj na eni strani	10 (8,77)
Rigidity.up	Prisotnost rigidnosti vsaj na eni strani	7 (6,14)
Harmonics	Pojavitev harmonikov vsaj v enem primeru	10 (8,77)
Spiro.Parkinsonian.only	Vsi spirografske podatki so označeni kot parkinsonski	0 (0,00)
Spiro.Essential.only	Vsi spirografske podatki so označeni kot esencialni	0 (0,00)

bil prvi kritičen primer, ki ga je izbral naš algoritem. Strokovnjaka smo prosili, da pove, katere značilnosti E.2 nakazujejo, da primer pripada razredu ET. Izbral je naslednje značilnosti: tremor v mirovanju, rigidnost in bradikinezijo; bradikinezijo (ki je zastopana v dveh atributih - enim za levo stran in enim za desno stran) je izbral kot najpomembnejšo značilnost primera izmed treh. Strokovnjak je na podlagi svojega domenskega znanja podal naslednji odgovor: "E.2 pripada razredu ET, ker bradikinezija ni prisotna niti na levi niti na desni strani." Z uporabo splošnega domenskega znanja je razložil tudi, da stran (leva ali desna) nima posebne vloge pri razlikovanju med ET in PT.

Na podlagi strokovnjakove razlage je inženir znanja (*angl.* knowledge engineer) uvedel nov atribut BRADYKINESIA z možnima vrednostima *true* (bradikinezija je prisotna na levi strani ali na desni strani) in *false* (bradikinezija ni bila zaznana na nobeni strani). Istočasno sta bila prvotna atributa (za bradikinezijo) izključena iz domene - po mnenju strokovnjaka je zadostna informacija že to, da se bradikinezija pojavi na eni izmed strani.

Na podlagi strokovnjakove razlage je bil argument "BRADYKINESIA is *false*" dodan kot argument za ET h kritičnemu primeru E.2. Metoda ni našla protiprimerov in s tem se je zaključila prva iteracija. Pred naslednjo iteracijo so bila inducirana nova pravila. Ena od opaznih sprememb je bila ta, da se je pojavilo naslednje pravilo:

IF BRADYKINESIA = *false* THEN class = EMT;

Pravilo pokriva 20 učnih primerov; vsi ti primeri pripadajo razredu ET.

Argumentacija primerov iz razreda MT

V prej opisani iteraciji je nevrolog kritični primer E.2 klasificiral kot izključno ET. V eni izmed naslednjih iteracij pa je bil kritični primer E.61 klasificiran kot PT in kot ET. V takem primeru strokovnjaka prosimo, da navede, katere značilnosti so v prid razredu ET in katere značilnosti so v prid razredu PT. Ekspert je razložil, da prisotnost posturalnega tremorja govori v prid razredu ET, medtem ko prisotnost rigidnosti govori v prid razredu PT. Tudi tokrat se je zanesel na svoje domensko znanje in razložil, da ločeno obravnavanje atributov za levo in desno stran v ničemer ne vpliva na odločitev za razred ET ali razred PT, zato je predlagal, da se atributa za levo in desno stran združita v enega.

Tako smo v domeno uvedli atributa POSTURAL.TREMOR.UP in RIGIDITY.UP, ki sta nadomestila originalne attribute za posturalni tremor in rigidnost. Prvi je bil uporabljen kot argument za ET, slednji pa kot argument za PT – oba argumenta sta bila dodana h kritičnemu primeru E.61. Medtem ko metoda ni našla protiprimerov za ekspertov argument v prid razredu ET, je bil za njegov argument v prid razredu PT najden protiprimer - E.45, ki pripada razredu ET.

Eksperta smo tokrat prosili, da primerja kritični primer E.61 s protiprimerom E.45 in razloži, katera je najpomembnejša značilnost v prid PT, ki velja za E.61 in ne velja za E.45. Po ekspertovem mnenju je pojavitev harmonikov pri primeru E.45 (in neobstoj harmonikov pri primeru E.61) tisto, kar je tipično za ET. Atribut HARMONICS, ki smo ga že dodali v domeno, z možnima vrednostma *true* in *false*, je bil dodan k prejšnjemu

argumentu. Metoda je nato našla še en protiprimer - E.30 (ki ima razred ET). Ekspert je razložil, da se tremor pri E.30 za razliko od kritičnega primera ni pričel simetrično na obeh straneh telesa. Argument je bil dodatno razširjen z atributom SIM.TREMOR.START in dodan h kritičnemu primeru E.61. Novih protiprimerov nismo našli. S tem se je iteracija zaključila.

Izboljševanje argumentov

Ekspertove argumente lahko izboljšamo na tri različne načine:

1. s strani eksperta z uporabo protiprimerov, ki jih izbere naša metoda,
2. zgolj z našo metodo in
3. s strani eksperta na podlagi njegovih ugotovitev glede induciranih pravil.

Prvi način je bil opisan v razdelku 3.1.4, v nadaljevanju sta opisana druga dva načina.

V prvi iteraciji (opisani v razdelku 3.1.4) se je izkazalo, da so ekspertovi argumenti zadoščali, da so bila z našo metodo inducirana pravila z jasno porazdelitvijo. Včasih pa metoda samodejno najde dodatne omejitve, s pomočjo katerih izboljša ekspertov argument. To se je zgodilo v eni izmed naslednjih iteracij, kjer se je pojavil argument: "RESTING.TREMOR.UP IS *true* AND HARMONICS IS *false* AND SIM.TREMOR.START IS *false*." Z uporabo tega argumenta se je induciralo naslednje pravilo, ki je tudi del končnega modela:

```
IF RESTING.TREMOR.UP = true AND HARMONICS = false
AND SIM.TREMOR.START = false AND spiro.Essential.only = false
THEN class = PMT;
```

Metoda je ekspertov argument samodejno izboljšala z dodajanjem dodatne omejitve v zgornje pravilo. Atribut SPIRO.ESSENTIAL.ONLY je bil uveden s strani eksperta v eni prejšnjih iteracij. Njegov pomen je naslednji: če je kvalitativna ocena spirografskih podatkov pri katerikoli od osmih opazovanj (atributov) v originalnih podatkih esencialna in hkrati nobena od njih ni parkinsonska (ali katera druga), potem je vrednost SPIRO.ESSENTIAL.ONLY *true*, sicer pa *false*. Zgornje pravilo pokriva 14 primerov (vsi so iz razreda ET) in ga je ekspert izredno hvalil – eden od razlogov za to je, da učinkovito kombinira klinične podatke in rezultate spirografije.

Tretji način je bil v našem primeru le enkrat konkretiziran v postopku elicitacije znanja z metodo ABML. Pri končnem pregledu pravil je ekspert potrdil vsa inducirana pravila z izjemo enega:

```
IF POSTURAL.TREMOR.UP = true AND SIM.BRADYKINESIA = true  
THEN class = EMT;
```

Čeprav pravilo pokriva 23 primerov (od 47) in ima jasno porazdelitev, je ekspert menil, da je atribut SIM.BRADYKINESIA nesmiselen. To je bil samodejno inducirani del pravila iz ekspertovega argumenta za primer E.61, kot je opisano v 3.1.4. Na podlagi ekspertove razlage je bil argument razširjen v "POSTURAL.TREMOR.UP is true and BRADYKINESIA = false." Takšnih sprememb se ne sme delati po pregledu rezultatov na testnih podatkih, še pomembneje pa je, da se ekspert pri tem zaneša na domensko znanje. Naslednje pravilo je bilo inducirano iz ekspertovega argumenta:

```
IF POSTURAL.TREMOR.UP = true  
AND BRADYKINESIA = false  
THEN class = EMT;
```

Čeprav ima pravilo izrazito slabše pokritje, je ekspert menil, da je pravilo konsistentno z njegovim domenskim znanjem. Na tej točki je ekspert odobril vsa pravila v končnem modelu. S tem je bil iterativni postopek zaključen.

3.2 Predstavitev domene

Nabor podatkov je vseboval 122 preiskovancev, ki jim je bila s strani nevrologa postavljena diagnoza parkinsonski tremor, esencialni tremor ali mešani tip tremorja (parkinsonski in esencialni tremor hkrati). Te tri diagnoze so predstavljale tudi razrede našega klasifikacijskega problema.

Za postavitev diagnoze je specialist nevrolog opravil klinični nevrološki pregled vključno z družinsko anamnezo. Poleg tega je vsak preiskovanec opravil spirografsko testiranje, ki je potekalo v prostorih Laboratorija za motnje gibanja in hojo Kliničnega oddelka za bolezni živčevja na Nevrološki kliniki v Ljubljani. Zajem podatkov je potekal z uporabo grafične tablice Wacom z namenskim pisalom in posebne programske opreme Spiralography podjetja Wise Technologies, ki omogoča zajem podatkov in izris osnovnih grafov časovnih vrst (narisana spirala, transformirana spirala v polarni koordinatni sistem, spreminjanje hitrosti, frekvenčni spekter, spreminjanje pritiska pisala na tablico med risanjem). Vsak preiskovanec je narisal štiri spirale. Najprej je narisal spiralo, kjer je s pisalom sledil že narisani spirali na predlogi enkrat z desno in enkrat z

levo roko. Nato je ponovno narisal z vsako roko po eno spiralo, a tokrat prostoročno (brez predloge). Ekspert za spirografijo je nato pregledal narisane spirale in podal kvalitativno oceno spiral - ali so v njih vidni znaki esencialnega tremorja, parkinsonskega tremorja ali v njih ni vidnih posebnih anomalij.

Vsak preiskovanec je bil opisan s 69 atributi, pri predprocesiranju podatkov pa se je število atributov zmanjšalo na 47. Izločeni atributi so povečini vsebovali komentarje, ali pa so bili nepopolni in za gradnjo modela nepomembni oz. neuprabni. Polovica preostalih atributov je vsebovala podatke o zgodovini preiskovancev in kliničnem nevrološkem pregledu, druga polovica pa o opravljenem spirografskem testiranju. Začetni atributi, ki so bili uporabljeni v postopku elicitacije znanja, njihov kratek opis in delež manjkajočih podatkov so predstavljeni v tabeli A.1, ki se nahaja v prilogi A.

Med samim postopkom elicitacije znanja so bili nekateri atributi nadomeščeni z novimi. To je ponavadi pomenilo združitev dveh ali večih atributov na podlagi eksperimentalnega poznavanja obravnavane domene.

3.3 Rezultati in evalvacija modela

3.3.1 Podatki

Podatki, ki so bili uporabljeni za gradnjo in evalvacijo modela, so bili zbrani v dveh delih.

V prvem delu smo posneli 67 preiskovancev, ki smo jih naključno porazdelili v učno množico, ki je vsebovala 47 preiskovancev in v testno množico, ki je vsebovala 20 preiskovancev. Porazdelitev preiskovancev v razrede je razvidna iz tabele 3.2.

Tabela 3.2: Porazdelitev preiskovancev, zbranih v prvem delu v razrede znotraj učne in testne množice.

	parkinsonski tremor	esencialni tremor	mešan tremor
učna množica (delež)	12 (25,53 %)	22 (46,81 %)	13 (27,66 %)
testna množica (delež)	5 (25,00 %)	10 (50,00 %)	5 (25,00 %)

V drugem delu zbiranja podatkov smo posneli 55 preiskovancev. Vsi ti primeri so bili, v želji po izboljšanju evalvacije, dodani v testno množico. Pri predprocesiranju podatkov smo bili primorani zavreči podatke osmih preiskovancev, saj so pri njih manjkali vsi podatki kliničnega pregleda. Na koncu je bilo v testni množici skupno

67 primerov. Porazdelitev preiskovancev znotraj razredov je predstavljena v tabeli 3.3, iz katere je razvidno povečanje deleža preiskovancev s parkinsonskim tremorjem v primerjavi z ostalima dvema tremorjema. Razlog za to je, da smo v prvem delu pridobili podatke oseb, ki jih je bilo težko pravilno diagnosticirati in so že bile na voljo ekspertu. V drugem delu smo pridobili podatke zaporednih priskovancev, njihova porazdelitev pa ustreza incidenci pojavljanja tremorjev.

Tabela 3.3: Porazdelitev preiskovancev, zbranih v prvem in drugem delu v razrede znotraj učne in testne množice.

	parkinsonski tremor	esencialni tremor	mešan tremor
učna množica (delež)	12 (25,53 %)	22 (46,81 %)	13 (27,66 %)
testna množica (delež)	29 (43,28 %)	28 (41,79 %)	10 (14,93 %)

3.3.2 Kvantitativna primerjava rezultatov

Uspešnost modela smo merili z različnimi metodami strojnega učenja: ABCN₂, NB, DT, SVM, RF in LR. Testiranje je potekalo z uporabo prilagojene metode desetkratnega prečnega preverjanja, kjer smo za učenje 47 učnim primerom dodali 9/10 primerov iz testne množice in testirali na preostali 1/10 primerov.

Za samo klasifikacijo primera smo diagnostični model nadgradili z algoritmom PILAR [34], ki vsakemu pravilu dodeli utež, pri čemer upošteva korelacije med pravili. Verjetnost razreda za določeno pravilo dobimo s seštevkem uteži vseh pravil, ki pokrijejo primer in izračunom verjetnosti z uporabo logistične transformacije. Z verjetnostmi $P(e = EMT)$ in $P(e = PMT)$, ki jih vrne algoritem PILAR, izračunamo verjetnosti $P(e = ET)$, $P(e = MT)$ in $P(e = PT)$ z uporabo formul, ki so predstavljene v poglavju 3.1.2 in tako dobimo končno klasifikacijo primera v enega izmed treh razredov.

V tabeli 3.4 so predstavljeni rezultati (klasifikacijska točnost (CA), območje pod ROC krivuljo (AUC), specifičnost in senzitivnost) uporabljenih metod pred in po postopku elicitanje znanja. Pri CA in AUC so predstavljeni tudi intervali zaupanja.

Tabela 3.4: Rezultati apliciranih metod strojnega učenja pred in po postopku elicitacije znanja.

			ABCN ₂	NB	DT	SVM	RF	LR
CA	Pred		,82 ±,09	,82 ±,13	,60 ±,15	,71 ±,09	,76 ±,07	,76 ±,09
	Po		,91 ±,09	,88 ±,09	,64 ±,16	,78 ±,09	,78 ±,11	,78 ±,09
AUC	Pred		,95 ±,06	,98 ±,02	,77 ±,15	,92 ±,06	,93 ±,04	,89 ±,09
	Po		,96 ±,08	,98 ±,03	,79 ±,16	,94 ±,06	,93 ±,06	,91 ±,07
ET	Spec.	Pred	0,90	0,85	0,46	0,77	0,59	0,90
		Po	0,97	0,87	0,46	0,77	0,62	0,82
	Sens.	Pred	0,86	0,93	0,89	0,75	1,00	0,82
		Po	0,93	1,00	0,89	0,86	1,00	0,75
PT	Spec.	Pred	0,87	0,95	0,97	0,82	1,00	0,82
		Po	0,89	0,97	0,97	0,95	1,00	0,92
	Sens.	Pred	0,86	0,83	0,41	0,79	0,72	0,83
		Po	0,97	0,86	0,48	0,83	0,72	0,86
MT	Spec.	Pred	0,95	0,93	0,91	0,93	1,00	0,91
		Po	0,98	0,96	0,95	0,93	1,00	0,91
	Sens.	Pred	0,60	0,50	0,30	0,30	0,20	0,40
		Po	0,70	0,60	0,30	0,40	0,30	0,60

Začetna klasifikacijska točnost modela ABCN₂ je bila 0,82, kar je primerljivo s točnostjo modela, dobljenega z uporabo metode NB. Prav tako imata obe metodi visok AUC ter specifičnost in senzitivnost za posamezen razred (PT, ET, MT) preko 0,80. Izjema je le senzitivnost za MT, ki je pri vseh metodah nizka (0,60 ali manj). Končna klasifikacijska točnost (po izvedeni elicitaciji znanja) modela ABCN₂ je 0,91, prav tako pa so se nekoliko povečale tudi vrednosti za AUC, specifičnost in senzitivnost.

Za primerjavo uspešnosti metode ABCN₂ z ostalimi metodami smo izmerili še statistično signifikanco razlike med rezultati za klasifikacijsko točnost in območjem pod ROC krivuljo. Rezultati kažejo, da z izjemo odločitvenih dreves ni signifikantne razlike v rezultatih za AUC. Se pa razlika pojavi pri primerjavi klasifikacijske točnosti med ABCN₂ in vsemi ostalimi metodami z izjemo NB. Na podlagi tega lahko trdimo, da ABCN₂ daje vsaj primerljivo dobre rezultate za klasifikacijsko točnost v primerjavi z ostalimi metodami, ki se tipično uporabljajo na tovrstnih domenah.

3.3.3 Kvalitativna primerjava začetnega in končnega modela

Po koncu postopka elicitacije znanja smo imeli 13 pravil. Končni model je predstavljen v tabeli 3.5, za vsako pravilo pa je podano še število primerov iz testne oz. učne množice, ki jih pravilo pokrije in točnost posameznega pravila.

Klasifikacijske točnosti pravil so bile ocenjene z uporabo m -ocene,

$$CA(\text{pravila}) = \frac{s + m \times p_a}{n + m}$$

kjer je bil parameter m nastavljen na vrednost 2 (kar je privzeta vrednost parametra v ABCN2), s je število pozitivnih primerov, ki jih pokrije pravilo, p_a je apriorna verjetnost napovedanega razreda, n pa je število vseh primerov, ki jih pravilo pokrije.

Vsako pravilo sta neodvisno ovrednotila dva eksperta-nevrologa (noben od njiju ni sodeloval v našem postopku elicitacije znanja). Ugotovila sta, da vsa pravila pravilno nakazujejo napovedani razred. V nadaljevanju bomo predstavili ekspertove razlage za tri od teh pravil.

Devet pravil vsebuje attribute, ki se nanašajo na spirografske teste, tri od teh pa temeljijo izključno na spirografiji. Eno od teh je naslednje pravilo:

```
IF HARMONICS = true THEN class = EMT;
```

Pravilo pravi: če se v frekvenčnem spektru tremorja pojavijo harmoniki, je tremor esencialen. Znano je, da je pojav harmoničnih frekvenc značilen za ET.

Naslednje pravilo v končnem modelu temelji izključno na atributih kliničnega pregleda:

```
IF BRADYKINESIA = true
  AND RIGIDITY.UP = true
  AND RESTING.TREMOR.UP = true
THEN class = PMT;
```

Pravilo pravi: če ima pacient tremor v mirovanju, bradikinezijo in rigidnost, potem gre za PT. Kombinacija tremorja v mirovanju, bradikinezije in rigidnosti pravzaprav klinično določa PT.

Nenazadnje spodnje pravilo uspešno kombinira rezultate spirografije in kliničnega pregleda:

```
IF RESTING.TREMOR.UP = true
  AND HARMONICS = false
```

Tabela 3.5: Pravila v končnem modelu.

Številka pravila	Pravilo	Število pokritih primerov (učna množica)	Točnost pravila (učna množica)	Število pokritih primerov (testna množica)	Točnost pravila (testna množica)
1	IF Bradykinesia = false THEN class = EMT	20	0,9768	21	0,8319
2	IF Qualitative.spiral = essential THEN class = EMT	23	0,9796	29	0,9721
3	IF Harmonics = true THEN class = EMT	10	0,9574	16	0,9519
4	IF Spiro.Parkinsonian.only = false AND Postural.tremor.up.left > 0 THEN class = EMT	23	0,9796	30	0,7229
5	IF History = positive AND Bare.right.freq.range > 5 THEN class = EMT	2	0,8723	1	0,7114
6	IF Qualitative.spiral = Parkinsonian THEN class = PMT	8	0,9064	27	0,9367
7	IF Bradykinesia = true AND Rigidity.up = true THEN class = PMT	17	0,9507	34	0,9490
8	IF Bare.right.speed.time = Parkinsonian AND Tremor.neck ≤ 0. THEN class = PMT	15	0,9449	14	0,9478
9	IF Rigidity.up = true AND Harmonics = false AND Tremor.start = right side THEN class = PMT	10	0,922	10	0,9303
10	IF Bradykinesia = true AND Rigidity.up = true AND Resting.tremor.up = true THEN class = PMT	15	0,9449	21	0,9202
11	IF Resting.tremor.up = true AND Harmonics = false AND Spiro.Essential.only = false AND Sim.tremor.start = false THEN class = PMT	14	0,9415	15	0,9508
12	IF Bradykinesia = true AND Diff.age.tremor.duration ≤ 60 years AND Harmonics = false AND Sim.tremor.start = false THEN class = PMT	15	0,8273	13	0,8776
13	IF Postural.tremor.up = true AND Bradykinesia = false THEN class = EMT	16	0,9716	11	0,8565

AND SPIRO.ESSENTIAL.ONLY = *false*

AND SIM.TREMOR.START = *false*

THEN class = PMT;

Pravilo klasificira v razred PT na podlagi obstoja tremorja v mirovanju in neobstoja

harmonikov v frekvenčnem spektru tremorja. Pri tem niso vsi rezultati spirografije označeni za esencialne, začetek tremorja pa ni bil obojestranski. Vse te ugotovitve so bile v skladu z ekspertovim znanjem.

V začetnem modelu, pred začetkom postopka elicitacije znanja, je bilo 9 pravil (tabela 3.6). Pet od teh pravil je kot atribut vsebovalo trajanje bolezni ali starost ob nastopu bolezni. Po mnenju eksperta pa noben od teh dveh atributov ni izrazito informativen. Obe motnji se namreč lahko začeta pojavljati v vseh starostih. Medtem ko se ET tipično začne pojavljati prej, pacienti z ET obiščejo nevrologa šele več let po pojavu bolezni, zato njihova starost ob nastopu in trajanje bolezni včasih nista ustrezno zabeležena.

Tabela 3.6: Pravila v začetnem modelu.

Številka pravila	Pravilo
1	IF Qualitative.spiral = essential THEN class = EMT
2	IF Bare.right.speed.time = essential THEN class = EMT
3	IF Bradykinesia.right ≤ 0 THEN class = EMT
4	IF Rigidity.up.right ≤ 0 AND Bradykinesia.left ≤ 1 THEN class = EMT
5	IF Rigidity.up.right ≤ 1 AND Resting.tremor.up.left ≤ 2 AND Postural.tremor.up.left > 0 AND Disease.duration ≤ 9 years THEN class = EMT
6	IF Bare.right.speed.time = Parkinsonian AND Disease.duration ≤ 12 years THEN class = PMT
7	IF Rigidity.up.right > 0 AND Age ≤ 83 years THEN class = PMT
8	IF Rigidity.up.right > 0 AND Disease.duration ≤ 12 years THEN class = PMT
9	IF Bradykinesia.right > 0 AND Age ≤ 74 years THEN class = PMT

Za razliko od končnega modela je več pravil v začetnem modelu z medicinskega vidika nesmiselnih. To prikazujeta naslednja dva primera.

```
IF RIGIDITY.UP.RIGHT > 0
  AND AGE  $\leq 83$  years
THEN class = PMT;
```

Prvi del pravilno določa, da če se v zgornji desni okončini pojavi rigidnost, je to znak za PT. Vendar po mnenju eksperta stran (leva ali desna) nima posebne vloge pri razlikovanju med ET in PT. Drugi pogoj ni pravilen. Zgornja meja za starost ob nastopu bolezni ne obstaja niti za PT niti za ET.

Tudi naslednje pravilo se pojavi v prvotnem modelu:

```
IF RIGIDITY.UP.RIGHT ≤ 1
  AND RESTING.TREMOR.UP.LEFT ≤ 2
  AND POSTURAL.TREMOR.UP.LEFT > 0
  AND DISEASE.DURATION ≤ 9 years
THEN class = EMT;
```

Po mnenju eksperta pozitivna vrednost za RIGIDITY.UP.RIGHT v splošnem govori v prid PT. Podobno v splošnem velja, da pozitivna vrednost za RESTING.TREMOR.UP.LEFT prav tako govori v prid PT. Vendar zgornje pravilo očitno ne razlikuje med primeri, pri katerih sta vrednosti teh dveh atributov pozitivni oz. negativni. Poleg tega se, kot smo pojasnili zgoraj, ET tipično začne pojavljati pri nižji starosti kot PT, zato bi bil tu bolj logičen simbol \geq namesto \leq .

3.3.4 Analiza napak

Z uporabo pravil v končnem modelu smo pravilno klasificirali 61 od 67 primerov. Iz tabele 3.7 je razvidno, koliko primerov v posameznem razredu smo klasificirali v napačen razred. Domenskega eksperta smo prosili, da ponovno pogleda primere, ki so bili napačno klasificirani in pravila, ki so bila uporabljena pri posamezni klasifikaciji.

Tabela 3.7: Matrika pravih in napačnih razvrstitev primerov z uporabo končnega modela.

		Napovedani razred		
		ET	MT	PT
Pravi razred	ET	26	1	1
	MT	0	7	3
	PT	1	0	28

Po natančni evalvaciji se je ekspert strinjal z dvema računalniškima klasifikacijama,

saj je nevrolog v času diagnoze spregledal nekatere podrobnosti. Razred teh dveh primerov bi tako dejansko morali spremeniti.

Poleg tega je ekspert spremenil razred še v enem primeru. Tokrat se ni strinjal z računalniško evalvacijo, toda pravila so mu vseeno pomagala najti napako pri postavitvi prvotne diagnoze.

Domenskega eksperta smo prosili, da pregleda še 12 napačno klasificiranih primerov, ki so bili klasificirani z uporabo prvotnega modela (tabela 3.8). Čeprav se je ekspert strinjal z dvema računalniškima klasifikacijama in je bil po skrbnem pregledu razred spremenjen pri še dveh drugih primerih, so ostali še trije napačno klasificirani primeri najhujšega tipa: PT je bil dvakrat napačno klasificiran kot ET in ET je bil enkrat napačno klasificiran kot PT.

Tabela 3.8: Matrika zamenjav pri uporabi začetnega modela.

		Napovedani razred		
		ET	MT	PT
Pravi razred	ET	24	2	2
	MT	1	6	3
	PT	3	1	25

3.3.5 Diskusija

Evalvacija kaže, da je postopek elicitacije znanja z metodo ABML vodil do večje klasifikacijske točnosti. Klasifikacija je bila po postopku elicitacije znanja boljša tako pri metodi ABML kot tudi pri drugih metodah strojnega učenja, ki smo jih uporabili za primerjavo. Uspešnost metode ABML z vidika CA in AUC je vsaj primerljiva z uspešnostjo drugih sodobnih metod, ki so bile vključene v primerjavo. Z našega stališča sta glavna rezultata postopka elicitacije znanja z ABML razumljivost modela in medicinska smiselnost elicitiranega znanja. Primerjava začetnega in končnega modela s stališča razumljivosti jasno pokaže pozitiven učinek elicitacije znanja.

Prepričani smo, da je razumljivost ključen vidik, zaradi katerega je končna množica pravil primerna za uporabo v sistemu za podporo odločanju (*angl.* decision support system - DSS). To je pomembno za razlago, ki podpira odločitve (diagnoze), še zlasti v medicinskih domenah, kjer zdravljenje temelji na diagnozi. Dobre razlage lahko poleg

tega, da okrepijo zdravnikovo zaupanje v predlagane diagnoze, pomagajo zdravnikom odkriti potencialne napake oziroma jih pozovejo k ponovnemu razmisleku glede diagnoze. DSS kot tak se lahko uporablja kot drugo mnenje, še zlasti pri težjih primerih.

Dodatna prednost razumljivosti je še ta, da lahko DSS enostavno preoblikujemo v dragoceno učno orodje. Zato je pomembno, da so razlage razumljive, smiselne in pravilne. Evalvacija naših modelov s strani nevrologov jasno potrjuje koristi metode ABML. Medtem ko je bilo prvotno znanje težko razumljivo in včasih celo nelogično z medicinskega vidika, je bila končna množica pravil veliko bolj konsistentna z medicinskim znanjem.

Analiza matrik zamenjav za začetni in končni model ne kaže izrazitega znižanja stopnje resnosti narejenih napak pri klasifikaciji. Ker vse napačne diagnoze niso enako škodljive, bi bilo v prihodnje smiselno razmisliti o cenovno-občutljivem učenju, kar bi lahko pomenilo očitno izboljšavo za naš DSS.

Podrobnejši pregled končne množice pravil pokaže še en zanimiv rezultat. Devet od trinajstih pravil vsebuje attribute, ki se nanašajo na rezultate spirografskega testiranja, tri od teh devetih pa temeljijo izključno na spirografiji. Omenimo naj, da med samim postopkom elicitacije znanja nismo posebno vzpodbujali uporabe rezultatov spirografskega testiranja. Ta rezultat nakazuje, da je spirografija zelo koristna za razločevanje med posameznimi tipi tremorjev. Ta tri pravila so dajala zelo dobre rezultate tudi za osem pacientov, za katere nismo imeli kliničnih podatkov.

Navedeno upravičuje nadaljnje delo na spirografskih podatkih, ki se lahko uporabi kot presejalni test. Ena izmed možnosti je uporaba spirografskega testa na napravah, kot so tablični računalniki ali pametni telefoni.

*Izgradnja modela za
razločevanje tremorjev na
podlagi digitalne spirografije*

4.1 Ozadje

Motorične motnje so ene najpogostejših znakov pri osebah s parkinsonovo boleznijo. To nam omogoča, da z zbiranjem in analizo podatkov objektivno ocenimo resnost bolezni.

Raziskovalci se ocenjevanja resnosti in napredovanja bolezni lotevajo na različne načine. Izdelujejo sisteme za nadzorovanje in avtomatsko razpoznavanje hoje [35, 36], sisteme za nadzorovanje in oceno motenj v govoru in glasu [37–40] ter sisteme za nadzor tremorjev in bradikinezije [41–44]. V želji po skrajšanju čakalnih vrst za obisk in pregled pri nevrologu je veliko časa in denarja namenjenih izdelavi enostavnih in dostopnih sistemov za domačo uporabo, ki bi omogočali objektivno oceno resnosti tremorjev. Raziskovalci v ta namen izdelujejo mobilne [19] in nosljive sisteme [45–48], namenjene uporabi v domačem okolju.

Večina teh sistemov zgolj meri in zbira podatke, ki jih nato pošljejo na centralni strežnik v obdelavo računalniškemu sistemu ali samim raziskovalcem. Ravno zaradi pošiljanja podatkov na centralni strežnik se pojavljajo vprašanja o uporabnikovi varnosti in varovanju osebnih podatkov. Res je, da sistemi pošiljajo podatke preko zaščitene kanalov, a so zaradi nepoznavanja tehnologije uporabniki lahko do njih nezaupljivi in jih ne želijo uporabljati. Prav zato je potreba po sistemih, ki bi lokalno obdelovali podatke in dali rezultate še toliko večja.

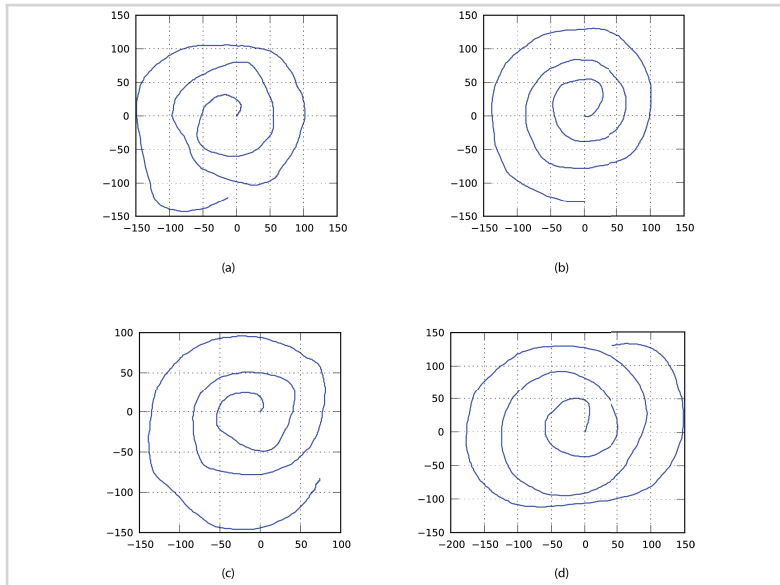
Že sedaj je na tržišču kar nekaj sistemov za domačo uporabo, a je njihova slabost v tem, da podajajo le informacijo o napredovanju bolezni (če jo uporabnik ima), ne omogočajo pa podajanja dejanske diagnoze oz. suma na diagnozo. Vsekakor je spremljanje napredovanja bolezni zelo pomembno, nič manj pomembno pa ni podajanje informacije o možni bolezni, ki bi jo uporabnik lahko imel.

4.2 Predstavitev domene in predznanje

Nabor podatkov vsebuje 159 preiskovancev, ki jim je bila s strani nevrologa postavljena diagnoza parkinsonski tremor (PT), esencialni tremor (ET), drug tip tremorja ali brez tremorja. Zaradi premajhnega števila preiskovancev, ki jim je bila postavljena diagnoza “drug tip tremorja”, so bili le-ti izločeni iz raziskave, preostale tri diagnoze pa so predstavljale tri razrede našega klasifikacijskega problema.

Vsak preiskovanec je moral s prstom na zaslon mobilnega telefona narisati štiri Arhimedove spirale. V prvem primeru je moral s kazalcem desne roke slediti črti na zaslonu

že izrisane spirale, v drugem primeru je moral spiralo narisati prostoročno (brez predloge), v tretjem in četrtem koraku pa je moral obe nalogi ponoviti še s kazalcem leve roke. Primer tako narisanih spiral je predstavljen na sliki 4.1.



Slika 4.1: Primer na mobilni telefon narisanih Arhimedovih spiral (a) z levo roko po predlogi, (b) z desno roko po predlogi, (c) z levo roko prostoročno (brez predloge) in (d) z desno roko prostoročno (brez predloge).

Vsaka spirala je predstavljena v obliki časovne vrste, kjer je vsaki točki (x, y) dodana še časovna komponenta t .

Cilj raziskave je bila izdelava odločitvenega modela za avtomatsko detekcijo potencialnih bolnikov s parkinsonovo boleznijo ali esencialnim tremorjem. Po tem, ko oseba nariše vse štiri spirale, mora biti sistem sposoben napovedati, ali so v narisanih spiralah prisotni znaki parkinsonskega oz. esencialnega tremorja ali teh znakov ni in je preiskovanec zdrav. Če se izkaže, da so prisotni znaki tremorja, je pomembno, da zna sistem ločiti med parkinsonskim in esencialnim tremorjem. Učni problem za izdelavo takšnega odločitvenega modela smo formulirali takole:

Na podlagi klasificiranih primerov (v razrede PT, ET, zdrav), kjer so za vsak primer podane štiri različne spirale, se nauči:

1. klasifikacijskega modela, ki zna napovedati ali ima oseba tremor ali ne, in

2. klasifikacijskega modela, ki zna ločiti med PT in ET.

Za potrebe gradnje atributov, ki bodo uporabljeni v odločitvenem modelu, je potrebno določeno nevrološko predznanje, da bi atributi zaradi potrebe po razumevanju odločitev dejansko vsebovali nevrološko znanje.

4.2.1 *Predznanje*

Odločitveni model, predstavljen v poglavju 3, vsebuje 13 pravil, od tega jih devet vsebuje attribute, dobljene s pomočjo spirografije. Naj izpostavimo tudi, da so tri pravila vsebovala izključno spirografske attribute. Med atributi, ki so se pojavili v končnem odločitvenem modelu, so bili harmoniki, frekvenčni razpon in kvalitativne ocene spirali (parkinsonska ali esencialna). Ti, bolj ali manj kvalitativni atributi, so bili osnova za izdelavo algoritmov za njihovo avtomatsko detekcijo, ki pa so bili nadgrajeni z dodatnim nevrološkim znanjem. Nevrologi lahko do neke mere vizualno ocenijo spirale in grafe, predstavljene v poglavju 2.2, njihovo predznanje pa temelji na njihovih večletnih kliničnih izkušnjah z digitalno spirografijo.

Za esencialni tremor je značilen posturalni tremor (tremor, ki se pojavi, ko imamo roke iztegnjene pred seboj) in kinetični tremor (ki se pojavi med prostovoljnim gibanjem udov) s frekvenco med 4 in 12 Hz. Za esencialni tremor načeloma velja, da je simetričen (leva in desna stran telesa sta enako prizadeti), se pa ponavadi začne v eni izmed rok in se nato preko vratu in glasilk razširi še na drugo stran telesa. Seveda obstaja tudi vrsta drugačnih oblik esencialnega tremorja - zgolj tremor glasu, zgolj tremor glave, zgolj tremor okončin ipd.

Za parkinsonski tremor je značilen tremor v mirovanju s frekvenco med 4 in 9 Hz in je eden najpogostejših motoričnih znakov za parkinsonovo bolezen. Ostali motorični znaki vključujejo bradikinezijo (upočasnjeno gibanje), rigidnost (povečan mišični tonus) in posturalno nestabilnost. Zanj velja, da je asimetričen, kar pomeni, da je ena stran telesa bolj prizadeta od druge.

Na podlagi teh dejstev se te značilnosti pojavljajo v spiralah v obliki žagastih vzorcev, ostrih kvadratnih robov, prisotnosti harmonikov in vrhov pri frekvenčni analizi absolutne hitrosti risanja, stisnjenosti spirale na eni strani in posledični asimetriji, asimetriji med spiralama, narisanimi z obema rokama ipd. Vse te značilnosti smo skušali zajeti v atributih, ki so predstavljeni v nadaljevanju.

4.3 Predstavitev atributov

V nadaljevanju so predstavljeni atributi, ki smo jih razvrstili v skupine glede na njihove lastnosti oz. glede na to, katere značilnosti spiral želimo z njimi detektirati.

4.3.1 Atributi za odkrivanje asimetrij

Atributi v tej skupini se uporabljajo za odkrivanje dveh različnih tipov asimetrij v spiralah:

1. asimetrij znotraj posamezne spirale;
2. asimetrij med spiralama, narisanimi z levo in desno roko.

Do asimetrije v posamezni spirali pride zaradi motenj gibanja pri risanju spirale (npr. tremor, rigidnost, bradikinezija) kar se odraža kot asimetrija na eni izmed strani spirale. Ponavadi se pojavlja pri osebah s parkinsonovo boleznijo. Asimetrije med spiralama, narisanimi z levo in desno roko, se pojavljajo pri skoraj vseh ljudeh, saj so le redki enako spretni z obema rokama. Je pa ta asimetrija bolj izražena pri osebah s PT in drugimi motnjami gibanja (npr. distonija).

Vsi atributi te skupine so predstavljeni v tabeli 4.1. Omeniti še velja, da je začetek/izhodišče spirale vedno nastavljen na točko (0, 0).

Tabela 4.1: Atributi za odkrivanje asimetrij.

Atribut	Kratek opis
<code>x_{symm}</code>	Atribut meri asimetrijo spirale na osi X. Izračuna se ga kot razliko med najbolj levo in najbolj desno točko spirale od osi Y, deljeno z vsoto teh dveh razdalj.
<code>y_{symm}</code>	Atribut meri asimetrijo spirale na osi Y. Izračuna se ga kot razliko med največjo in najmanjšo točko spirale od osi X, deljeno z vsoto teh dveh razdalj.
<code>angVsRad.template.distance</code>	Atribut meri asimetrijo spiral, narisanih z levo in desno roko. Spirale so narisane po predlogi, računanje pa se vrši na spiralah, predstavljenih v polarnih koordinatah. Rezultat predstavlja koren kvadrata napake (RMSE) med obema spiralama.

Izračun vrednosti atributa `andVsRad.template.distance` se vrši na spiralah, narisanih z levo in desno roko po predlogi. S tem zagotovimo, da sta spirali približno enake dolžine. Izračun se vrši na spiralah, predstavljenih v polarnih koordinatah, kjer prvo polovico obrata zanemarimo/izločimo, saj so ti podatki ponavadi nezanesljivi. Preostale točke spiral so nato enakomerno vzorčene do dolžine krajše spirale. Skupno vzorčimo 300 točk (kot), za katere nato interpoliramo njihove pare (radij) za obe spirali. Na točkah (radijih), ki smo jih dobili z interpolacijo, izračunamo koren kvadrata napake (angl. root mean squared error – RMSE), kar predstavlja vrednost atributa.

4.3.2 Atributi za odkrivanje ekstremov

Atributi v tej skupini so bolj ali manj elementarni in se jih lahko aplicira na praktično poljubno časovno vrsto. Osnovna ideja je, da s temi atributi predstavimo število, velikost in porazdelitev vrhov v časovnih vrstah, ki predstavljajo različne hitrosti – absolutno hitrost, radialno hitrost in kotno hitrost. S temi atributi želimo zaznati žagaste vzorce, saj je za njih značilna pogosta menjava smeri. Fiziološko gledano se ti vzorci pojavljajo kot posledica tremorja, saj se hitrost in smer risanja hitro spreminjata. To je še posebej značilno za osebe z ET, saj je pri njih frekvenca tremorja višja, kar se odraža v večjem številu manjših ekstremov v krajšem časovnem obdobju.

V tabeli 4.2 so na kratko predstavljeni vsi atributi za odkrivanje ekstremov.

Tabela 4.2: Atributi za odkrivanje ekstremov.

Atribut	Kratek opis
<code>nPeaks</code>	Atribut predstavlja število sprememb smeri risanja, normalizirano z dolžino časovne vrste.
<code>avgP</code>	Predstavlja povprečno velikost sosednjih ekstremov.
<code>stdevP</code>	Predstavlja standardno deviacijo razlik vrednosti sosednjih ekstremov.
<code>t.rho</code>	Predstavlja Spearmanovo korelacijo med podanim zaporedjem meritev in časovnimi žigi zaporedja.
<code>incDev</code>	Predstavlja če in kako se standardna deviacija sosednjih ekstremov spreminja skozi čas.

Izračun atributa `stdevP` poteka tako, da najprej v vrsti poiščemo vrednosti vseh

lokalnih minimumov in lokalnih maksimumov. Ko te vrednosti poznamo, sledi izračun razlik v vrednostih sosednjih minimumov in maksimumov. Dobljene razlike nato uporabimo v formuli za izračun standardne deviacije.

4.3.3 Atributi za odkrivanje napak/odstopanj

Za odkrivanje napak v spirali smo zgradili nekaj atributov, ki primerjajo narisano spiralo z optimalno spiralo in so predstavljeni v tabeli 4.3. Atributi *err*, *err0* in *errBF* merijo napako v polarnem koordinatnem sistemu. V ta namen smo morali najprej pretvoriti točke iz kartezičnega koordinatnega sistema $(x(t), y(t))$ v polarni koordinatni sistem:

$$r = \sqrt{x(t)^2 + y(t)^2}$$

$$\varphi(t) = \text{atan2}(y(t), x(t))$$

Za izgradnjo atributov *err*, *err0* in *errBF* smo z uporabo metode najmanjših kvadratov točkam (φ, r) izračunali prilegajoče točke in nato izračunali koren povprečja kvadratov napak (angl. root mean squared error – RMSE).

Tabela 4.3: Atributi za odkrivanje napak/odstopanj v narisanih spiralah.

Atribut	Kratek opis
<i>err</i>	RMSE med narisano in optimalno spiralo.
<i>err0</i>	RMSE med narisano in optimalno spiralo, ki poteka skozi točko (0, 0).
<i>errBF</i>	RMSE med narisano in optimalno spiralo s translacijo narisane spirale v točko (0, 0).
<i>rmse100</i>	RMSE med narisano spiralo in optimalno ukrivljenostjo pri velikosti okna 100 točk.
<i>rmse150</i>	RMSE med narisano spiralo in optimalno ukrivljenostjo pri velikosti okna 150 točk.
<i>rmseSum</i>	Vsota RMSE preko vseh velikosti oken.
<i>rmseMin</i>	Minimalen RMSE preko vseh velikosti oken.

Na podlagi vizualne analize spiral je možno sklepati, da je ukrivljenost spirale lahko pomemben atribut. Pri arhimedovi spirali pa se le-ta z naraščujočim kotom monotono zmanjšuje, kar pa za spirale preiskovancev (zdravih ali s tremorjem) ne velja. Namesto

ukrivljenosti smo zato izračunali približek radijev ukrivljenosti s prilagajanjem krožnic zaporedju točk z različnimi velikostmi okna: 10, 20, 30, 50, 75, 100, 125, 150, 175. Večja velikost okna posledično pomeni več glajenja. Nato smo na te točke prilagodili ravne črte in izračunali RMSE. Za pomembna sta se izkazala atributa `rmse100` in `150`, kjer številki 100 in 150 predstavljata velikost okna. Izračunali smo še vsoto RMSE (atribut `rmseSum`) in minimalen RMSE (atribut `rmseMin`) preko vseh oken.

4.3.4 Atributi frekvenčne analize

V poglavju 4.2.1 smo omenili nekatere značilnosti parkinsonskega in esencialnega tremorja. Med drugim smo izpostavili, da za ET velja, da se pojavlja v frekvenčnem območju med 4 in 12 Hz, medtem ko je za PT značilno območje med 4 in 9 Hz. To pomeni, da se frekvenčna območja med seboj prekrivajo. Ker pa nam vseeno lahko podajo določeno informacijo, atributi v tej skupini opisujejo značilnosti frekvenčne analize na grafu absolutne hitrosti risanja spirale in so predstavljeni v tabeli 4.4.

Tabela 4.4: Atributi, ki izhajajo iz frekvenčne analize.

Atribut	Kratek opis
<code>freq.dev</code>	Standardna deviacija gostote moči spektra absolutne hitrosti med 3 in 22,5 Hz.
<code>peakAmpXOut</code>	Zaznava izstopajoče vrhove v frekvenčnem spektru.

4.3.5 Atributi za analizo sprememb v radiju

Na podlagi vizualne ocene spirale je bilo predvsem pri spiralah, ki jih je narisala oseba z ET očitno, da se lahko žagaste vzorce detektira tudi na podlagi sprememb v radiju. Če spiralo rišemo enakomerno, se radij enakomerno povečuje, saj ves čas rišemo navzven, torej stran od izhodišča spirale. Spremembe radija se najlažje detektirajo na grafu radialne hitrosti. V tej skupini se nahaja le en atribut, `percNeg`, ki meri, kolikšen del dolžine spirale je bil risan proti izhodišču spirale. Številka torej predstavlja odstotek dolžine spirale, ko je radialna hitrost negativna.

4.3.6 Atributi izpeljani iz večih testov

Večina do sedaj predstavljenih atributov se nanaša na karakteristike ene spirale (pri asimetriji dveh spirale). Velikokrat se lahko zgodi, da je prizadeta le ena stran telesa

(posledično le ena roka) in še to se lahko odraža le na eni izmed spiral (npr. samo na spiralah, narisanih prostoročno). Zato smo se odločili, da za vse attribute, ki so bili do sedaj predstavljeni izračunamo minimum, maksimum, povprečje in vsoto atributa na vseh (štirih) spiralah, ki jih je narisal preiskovanec.

4.4 Rezultati in evalvacija modela

4.4.1 Podatki

Podatki, ki so bili uporabljeni za izgradnjo in evalvacijo odločitvenega modela so bili zbrani na Univerzitetnem kliničnem centru v Ljubljani. Posnetih je bilo 159 preiskovancev. Ker smo za gradnjo odločitvenega modela potrebovali čim bolj točne podatke, je vse preiskovance pregledal tudi nevrolog, ki jih je razvrstil v štiri različne skupine: parkinsonski tremor, esencialni tremor, drug tip tremorja in brez tremorja (zdrava kontrola). Preiskovance, za katere je bilo določeno, da imajo drug tip tremorja, smo izločili. Ostalo je 143 preiskovancev.

Preostale tri skupine preiskovancev so predstavljale razrede v našem odločitvenem problemu. Pri izgradnji končnega odločitvenega modela smo, po posvetu z eksperti, izločili preiskovance, ki so zaužili zdravilo z učinki levodope manj kot tri ure pred testiranjem. Končni nabor je tako vseboval podatke za 124 preiskovancev. Začetna in končna porazdelitev preiskovancev v posamezne razrede je razvidna iz tabele 4.5.

Tabela 4.5: Porazdelitev preiskovancev v razrede v začetnih in končnih podatkih za tem, ko so bili izločeni preiskovanci iz skupine “drug tip tremorja”.

	Parkinsonski tremor	Esencialni tremor	Brez tremorja
Začetni podatki (delež)	38 (26,57%)	19 (13,29%)	86 (60,14%)
Končni podatki (delež)	21 (16,94%)	17 (13,71%)	86 (69,35%)

4.4.2 Opis eksperimenta

Atribute, opisane v poglavju 4.3, smo uporabili pri gradnji klasifikacijskih modelov. Pri tem smo uporabili štiri različne metode strojnega učenja: logistično regresijo (angl. logistic regression - LR), naivni bayesov klasifikator (angl. naive bayes classifier - NB), metodo podpornih vektorjev (angl. support vector machines - SVM) in metodo naključnih gozdov (angl. random forests classifier - RF). Prvi dve metodi smo si

izbrali, ker so tako zgrajeni modeli dokaj enostavno razločljivi in jih je lahko preveriti s predznanjem. Drugi dve smo vključili za primer, če bi se pojavili kakšni pomembni nelinearni vzorci, ki jih s prvima dvema ne bi zajeli.

Kot je bilo že predstavljeno v poglavju 4.2, je klasifikacijski problem sestavljen iz dveh korakov:

prvi korak: v spiralah poišči znake, ki nakazujejo na ET ali PT;

drugi korak: če so bili v prvem koraku najdeni znaki, določi, ali gre za znake ET ali PT.

Problem je možno rešiti na dva načina:

1. nauči se en model, ki se ga uporabi v prvem in drugem koraku, ali
2. za prvi korak se nauči ločeni model, kjer so podatki za PT in ET združeni v isti razred.

Oba pristopa sta bila testirana na dvorazrednem problemu (prvi korak):

učenje na dveh razredih: klasifikacijski model je zgrajen iz podatkov, ki so razdeljeni v dva razreda (zdrav, združena PT in ET),

učenje na treh razredih: klasifikacijski model je zgrajen iz podatkov, ki so razdeljeni v tri razrede (zdrav, PT, ET) in testiran na dvorazrednih podatkih.

Parametri metod

Za gradnjo modelov smo uporabili implementacijo metod v knjižnici Orange [49]. Metoda LR je vzeta iz knjižnice *liblinear* [50], kjer smo uporabili privzete nastavitve parametrov. Pri NB smo parameter m v m -oceni verjetnosti [51] nastavili na 20, da smo preprečili prekomerno prilagajanje modela testnim podatkom. Parametri metode SVM so se optimizirali avtomatsko z uporabo interne navzkrižne validacije (angl. internal cross validation). Število dreves pri metodi RF je bilo nastavljeno na 150.

Uporaba diskretizacije

Iz predstavitve atributov v poglavju 4.3 je razvidno, da je večina le-teh zveznih. Tako smo bili postavljeni pred vprašanje, ali dobljene vrednosti diskretiziramo ali ne. Dva razloga sta govorila v prid temu:

1. v medicinskih domenah se večinoma določajo meje, kdaj je neka vrednost v območju normalnega stanja in kdaj ne;
2. zaradi majhnega števila učnih primerov se z diskretizacijo lahko izognemo prekomernemu prilagajanju podatkom.

Zavedati se je potrebno, da nekatere metode, kot je naprimer LR, delujejo bolje z zveznimi vrednostmi.

Vse metode smo zato testirali dvakrat - z uporabo diskretizacije in brez nje. Uspešnost metod smo primerjali z uporabo klasifikacijske točnosti (angl. classification accuracy - CA), območja pod ROC krivuljo (angl. area under ROC curve - AUC) in brierjeve ocene (angl. Brier score - Brier). Da bi lažje ocenili uspešnost metod, smo za primerjavo vzeli še klasifikator, ki primere klasificira v večinski razred (angl. majority classifier - Majority).

4.4.3 Evalvacija rezultatov

Rezultati uspešnosti metod z uporabo diskretizacije in brez nje so predstavljeni v tabeli 4.6. Iz nje je razvidno, da vse metode delujejo bolje od klasifikatorja Majority. Presenetljivo pa je dejstvo, da vse metode delujejo bolje z uporabo diskretizacije. Že v poglavju 4.4.2 smo omenili, da bi za metodo LR pričakovali, da bo bolje delovala na zveznih in ne na diskretnih vrednostih.

Tabela 4.6: Primerjava uspešnosti izbranih metod (z uporabo diskretizacije in brez nje) na trirazrednem problemu.

Metoda	Brez diskretizacije			Z diskretizacijo		
	CA	AUC	Brier	CA	AUC	Brier
Majority	0,694	0,500	0,472	0,694	0,500	0,472
LR	0,799	0,837	0,390	0,835	0,930	0,235
NB	0,784	0,872	0,342	0,822	0,900	0,304
RF	0,790	0,872	0,309	0,813	0,925	0,263
SVM	0,786	0,889	0,303	0,812	0,914	0,261

Razlog, da vse metode (tudi metoda LR) delujejo bolje z diskretizacijo, bi lahko bil v tem, da smo uporabili diskretizacijo na osnovi entropije [52], ki minimizira entropijo

razredov v učnih podatkih. Torej bi lahko domnevali, da je imela pri tem večjo vlogo metoda izbire podmnožice značilk (angl. feature subset selection - FSS) in ne sama diskretizacija. Preverili smo, ali razlog res v tem.

Iz tabele 4.7 lahko razberemo, da je na boljše rezultate vendarle vplivala diskretizacija in ne zgolj FSS. Prva vrstica prikazuje rezultate LR brez diskretizacije, druga pa rezultate LR z diskretizacijo. V tretji vrstici so prikazani rezultati LR, kjer so diskretizirani vsi atributi, četudi razlika v entropiji ni bila dovolj visoka. Na ta način smo se izognili FSS. V zadnji vrstici so prikazani rezultati učenja na zveznih atributih, uporabljeni pa so le tisti atributi, ki so bili izbrani pri klasični diskretizaciji. S tem smo dosegli to, da je bil uporabljen le FSS del diskretizacije. Iz rezultatov lahko vidimo, da LR z uporabo diskretizacije brez FSS deluje nekoliko bolje od LR na zveznih podatkih, medtem ko LR na zveznih podatkih z uporabo FSS deluje slabše. Iz tega lahko sklepamo, da je pri uspešnosti LR z diskretizacijo in FSS imela večji doprinos sama diskretizacija podatkov.

Tabela 4.7: Primerjava rezultatov dobljenih z LR brez diskretizacije, LR z diskretizacijo (LR + disc), LR z vsiljeno diskretizacijo vseh atributov (LR + forced disc.) in LR, kjer je bila diskretizacija uporabljena za izbiro podmnožice značilk, sami atributi pa so ostali zvezni (LR + FSS).

Metoda	CA	AUC	Brier
LR	0,799	0,837	0,390
LR + disc	0,835	0,930	0,235
LR + forced disc.	0,806	0,926	0,282
LR + FSS	0,777	0,870	0,325

V tabeli 4.8 prikazujemo rezultate testiranja modelov na dvorazrednem problemu. Ti kažejo, da se vsi modeli obnesejo bolje, ko se učijo iz podatkov, razdeljenih v tri razrede v primerjavi z modeli, ki se učijo iz podatkov, razdeljenih v dva razreda. Razlog bi lahko bil v tem, da so spirale, ki jih nariše oseba s PT tako različne od spiral, ki jih nariše oseba z ET, da njihovo združevanje ne doprinese k izboljšavi modelov.

Na podlagi tega smo se odločili, da za naš klasifikacijski problem ne bomo uporabljalih dveh različnih modelov za posamezen korak klasifikacije, ampak le enega.

Tabela 4.8: Primerjava uspešnosti izbranih metod na dvo-razrednem problemu. Metode so se učile iz podatkov razdeljenih v dva ali tri razrede.

Metoda	Učenje dveh razredov			Učenje treh razredov		
	CA	AUC	Brier	CA	AUC	Brier
Majority	0,694	0,500	0,472	0,694	0,500	0,472
LR	0,889	0,932	0,172	0,902	0,951	0,153
NB	0,894	0,903	0,196	0,895	0,938	0,185
RF	0,888	0,917	0,197	0,909	0,942	0,172
SVM	0,894	0,920	0,174	0,896	0,926	0,162

4.4.4 Predstavitev izbranega modela

Na podlagi dognanj iz 4.4.3 smo izbrali model, dobljen z logistično regresijo (LR) na diskretiziranih atributih in ga uporabili v obeh korakih klasifikacijskega problema.

Logistična regresija je zasnovana za uporabo na dvorazrednih problemih. Ker imamo v našem primeru tri razrede (zdrav, PT, ET), smo se morali naučiti tri modele: *zdrav proti PT + ET*, *PT proti zdrav + ET* in *ET proti zdrav + PT*. V tabeli 4.9 je predstavljenih osem najpomembnejših atributov modela *zdrav proti tremor (PT ali ET)* s pripadajočimi betami (linearnimi koeficienti), dobljenimi z uporabo LR.

Tabela 4.9: Osem najpomembnejših atributov modela (*zdrav proti tremor (PT ali ET)*) dobljenim z LR in uporabo diskretizacije.

Atribut	Beta
Prior	3,359
LF.err > 18,894	-1,532
RF.rmse100 < 7,566	+1,061
RF.rmseSum < 206,47	+1,043
LF.ang.avgP > 1,290	+0,897
LT.ang.peaks > 0,017	-0,821
RF.percNeg0 > 0,405	-0,800
angVsRad.template.distance > 236,36	-0,764
LF.Ysymm > 0,234	-0,735

Najpomembnejši atribut je LF_{err} , ki predstavlja razdaljo oz. napako med narisano spiralo in optimalno spiralo, narisano prostoročno (brez predloge) z roko, ki ni vodilna (npr. z levo roko, če je bil preiskovanec desničar). Trije atributi, ki sledijo, so prav tako izpeljani iz spirala, ki so bile narisane prostoročno. Iz tega lahko sklepamo, da preiskovanci s tremorjem lažje narišejo spiralo, če imajo na voljo črto, ki ji lahko sledijo.

4.5 Diskusija

Klasifikacijski problem, ki je predstavljen v poglavju 4.2, je po svoji naravi problem klasifikacije časovnih vrst. V literaturi je zaslediti, da se je tovrstnih problemov najbolje lotiti z uporabo metode k-najbližjih sosedov (angl. k-nearest neighbour - kNN), ki uporablja metodo dinamičnega ukrivljanja časa (angl. dynamic time warping - DTW), torej kNN-DTW [53]. Naši začetni poskusi so pokazali, da je metoda kNN-DTW na naših podatkih delovala slabše od gradnje modelov z uporabo zgrajenih atributov. Iz tega lahko sklepamo, da predznanje, ki je zakodirano v attribute, doprinese več informacije učnim metodam kot same časovne vrste. Zaradi tega smo opustili nadaljnjo gradnjo modelov zgolj na časovnih vrstah.

Iz rezultatov, predstavljenih v poglavju 4.4.3 je razvidno, da je uspešnost posameznih metod med seboj enakovredna. Ker smo želeli razumljive modele, smo se na koncu odločili med NB in LR, saj so linearni modeli lažje razumljivi in razložljivi. Z njimi namreč lahko ugotovimo, kateri atributi so pomembnejši, kar nam je pomagalo tudi pri zaznavanju napak v podatkih in definiciji atributov.

Izbrani model je bil implementiran v mobilni aplikaciji za zaznavanje ET in PT, ParkinsonCheck, in deluje na večini mobilnih platform. Uporabniški vmesnik aplikacije je prikazan na sliki 4.2.



Slika 4.2: Uporabniški vmesnik aplikacije ParkinsonCheck. Slika prikazuje glavni meni aplikacije, zaslon za risanje spirale in vizualiziran rezultat testiranja.





*Metoda za vizualno razlago
diagnoz*

V poglavju 4 smo predstavili računalniški diagnostični model za ločevanje med zdravi mi preiskovanci in tistimi s parkinsonskim ali esencialnim tremorjem. Model temelji na atributih, razvitih za to domeno, in logistični regresiji. Ko pride do vprašanja zaupanja v sisteme, ki uporabljajo tovrstne diagnostične modele, se izkaže, da jih eksperti (v našem primeru nevrologi) le stežka sprejemajo, saj ne razumejo ozadja odločitev. Zato je pomembno, da sistemi ne delujejo po principu “črne škatle” (*angl.* black-box), ampak morajo ekspertom poleg svoje odločitve podati tudi tehtno razlago za takšno odločitev.

V našem primeru smo po pogovoru z eksperti (nevrologi) prišli do ugotovitve, da bi bilo najbolje razlago podati v obliki vizualizacije napak v spiralah. To pomeni, da bi označili zgolj tiste dele s strani preiskovancev narisanih spiral, ki pomembno prispevajo k določeni odločitvi računalniškega diagnostičnega modela. S tem bi pridobili zaupanje pri zdravnikih, saj bi jim pokazali, kaj je algoritem zaznal v narisanih spiralah kot problematično. Tako na nek način vzpostavimo komunikacijo z zdravnikom in mu argumentiramo odločitev algoritma.

Tovrstna vizualizacija je koristna predvsem za eksperte (zdravnike) in razvijalce odločitvenega modela. Skozi razvoj metode smo ugotovili, da lahko s pomočjo vizualizacije odkrijemo tudi novo znanje, služi pa nam lahko tudi kot orodje za vizualno odkrivanje napak v zasnovi atributov.

5.1 Metodologija

Cilj našega dela je vizualizirati najbolj zanimive oz. najpomembnejše dele vrst, ki so pomembno prispevali k vrednosti posameznega atributa v odločitvenem modelu.

Definirajmo najprej nekaj osnovnih pojmov, ki so pomembni za našo metodologijo.

Definicija 1: Vrsta T je zaporedje točk t_1, t_2, \dots, t_n .

Točke so običajno številske vrednosti, ki jih lahko uporabimo za izračun vrednosti nekega atributa. Vrednosti so večinoma pridobljene z meritvami, interpolacijami, izračuni ipd.

Definicija 2: Atribut A je funkcija vrste T : $A(T)$.

Definicija 3: Interval I je podzaporedje sosednjih točk v vrsti T : t_i, \dots, t_j .

Pri tem velja, da je najkrajše podzaporedje dolžine ena (torej ena točka iz vrste), najdaljše pa je določeno z dolžino celotne vrste T .

Koristnost množice intervalov

Koristnost množice intervalov predstavlja razmerje med *prispevkom* množice in njeno *ceno*: $\text{koristnost} = \frac{\text{prispevek}}{\text{cena}}$.

Prispevek množice intervalov bo definiran v razdelku "Prispevek množice intervalov in nevtralne vrednosti", *cena* množice intervalov pa v razdelku "Cena intervalov".

Definicija 4: Naj bo W množica vseh možnih intervalov I .

Naš problem je:

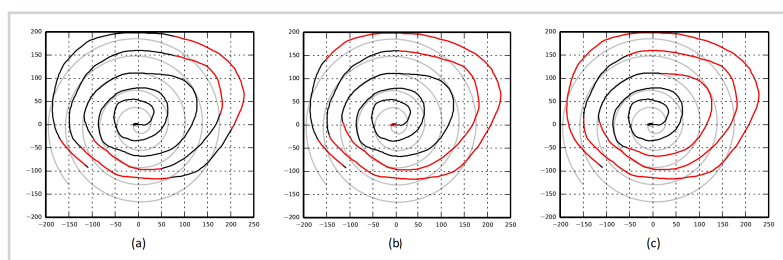
Poišči $W_a = \{I_1, I_2, \dots, I_k\}$, kjer je $W_a \subset W$, pri čemer W_a predstavlja množico z največjo koristnostjo.

Množica je lahko dodatno omejena s številom elementov, glede na zahteve, ki izhajajo iz potreb domene, pri kateri se uporablja (npr. označimo le delež osnovne vrste, da uporabnika opozorimo na najpomembnejše dele).

Dodatne zahteve glede omejitev se določijo za vsak atribut posebej glede na domeno in samo uporabo oz. namen vizualizacije. V našem primeru smo pri vseh atributih določili, da sme množica W_a vsebovati največ 25% vseh točk v vrsti T . Za takšno omejitev smo se odločili, ker se pri nekaterih atributih lahko zgodi, da se z vsako novo točko, ki jo dodamo v množico W_a , poveča koristnost. Tako bi se v najslabšem primeru lahko zgodilo, da bi nam algoritem označil celotno vrsto. Naša želja je, da označimo le tiste dele vrste, ki so pomembno prispevali k vrednosti posameznega atributa in ne vseh delov, ki so kakorkoli prispevali h končni vrednosti. Če je teh točk zelo veliko in bi to rezultiralo v označitvi večjega dela vrste, sama vizualizacija izgubi svoj pomen. V našem primeru smo preizkusili različne nastavitve te omejitve (25%, 50% in 75% vseh točk v vrsti), pri čemer je bilo mnenje eksperta, da je za to domeno smiselno označiti največ 25% vseh točk.

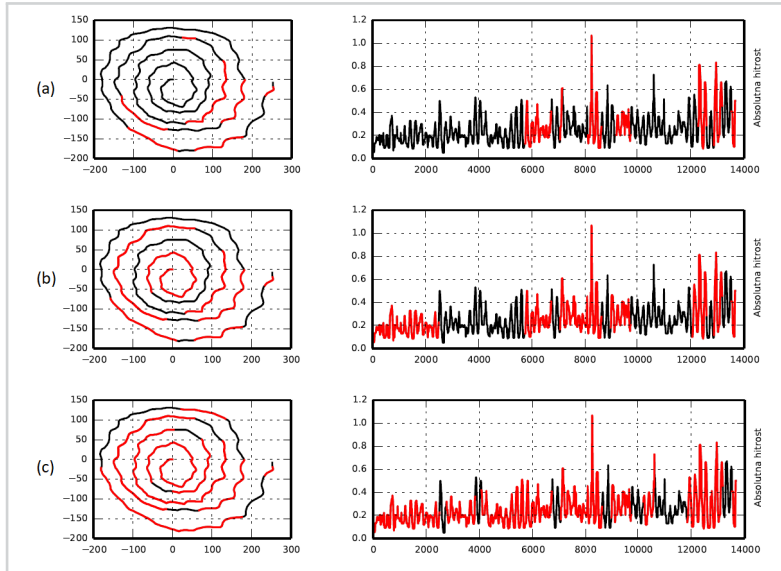
Na sliki 5.1 je prikazan primer označitve spirale, na kateri je bil za klasifikacijo najbolj pomemben atribut err . Vidimo lahko, da so na sliki (a) označeni intervali točk, ki najbolj odstopajo od idealne spirale (narisana s sivo). Na sliki (b) so označene iste točke, kot na sliki (a), dodanih pa je še enkrat toliko točk, ki pa že manj odstopajo od idealne spirale. Na sliki (c) so dodatno označene točke, ki še manj odstopajo od idealne spirale.

Slika 5.1: Primer označitve za atribut err z različnimi nastavitvami največjega dovoljenega deleža označenih točk osnovne spirale: (a) 25%, (b) 50% in (c) 75%.

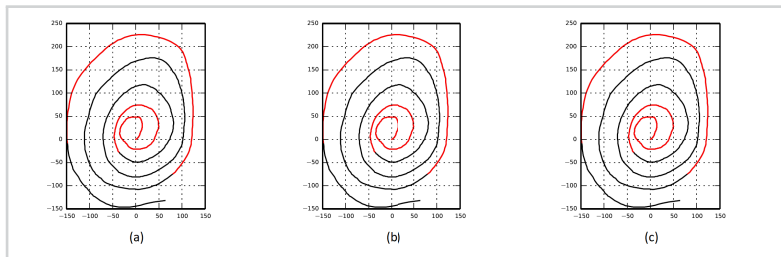


Na sliki 5.2 je prikazan primer označitve spirale, kjer je bil za klasifikacijo najbolj pomemben atribut $absSp.stdevP$, ki predstavlja standardno deviacijo razlik vrednosti sosednjih ekstremov na podatkih o absolutni hitrosti risanja spirale (sam postopek izračuna je opisan v poglavju 4.3.2). Slike v levem stolpcu prikazujejo narisano spiralo z označenimi predeli, kjer so bile standardne deviacije sosednjih ekstremov največje. Slike v desnem stolpcu prikazujejo grafe absolutne hitrosti risanja spirale, z označenimi predeli z največjo standardno deviacijo sosednjih ekstremov. Ponovno smo uporabili različne nastavitve za določanje, kolikšen delež točk smemo označiti. Iz grafov hitrosti je razvidno, da so na slikah (a) označeni deli, kjer so razlike med sosednjimi ekstremi največje, medtem ko so na slikah (b) in (c) označeni tudi deli, kjer so te razlike vidno manjše. Omeniti velja, da smo attribute, ki izhajajo iz podatkov o hitrostih, označevali na animiranih spiralah, saj le tako lahko izpostavimo predele, kjer je prihajalo do razlik.

Na sliki 5.3 lahko opazimo, da so na vseh treh slikah, kljub različnim nastavitvam največjega dovoljenega deleža označenih točk, označeni isti deli spirale. To pomeni, da je množica W_a že sama po sebi vsebovala manj kot 25% vseh točk v vrsti in sama omejitev največjega dovoljenega deleža označenih točk ni prišla do izraza.



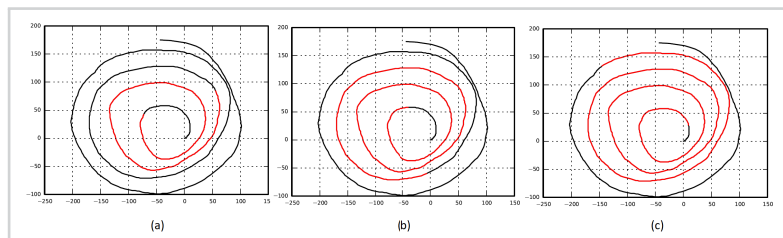
Slika 5.2: Primer označitev za atribut `abs-Sp.stdevP` z različnimi nastavitvami največjega dovoljenega deleža označenih točk osnovne spirale: (a) 25%, (b) 50% in (c) 75%.



Slika 5.3: Primer označitev za atribut `rmse150` z različnimi nastavitvami največjega dovoljenega deleža označenih točk osnovne spirale: (a) 25%, (b) 50% in (c) 75%.

Na sliki 5.4 prikazujemo isti atribut, kot na sliki 5.3, le da je tu omejitev največjega deleža označenih točk prišla do izraza. Ali bo ta nastavev za posamezen atribut dejansko prišla do izraza, je odvisno od same vrste, na kateri računamo atribut in označujemo točke.

Slika 5.4: Primer označitve za atribut `rmse150` z različnimi nastavitvami največjega dovoljenega deleža označenih točk osnovne spirale: (a) 25%, (b) 50% in (c) 75%.



Ker se mnenja uporabnikov, kolikšen delež označenih točk je smiselno označiti, lahko bistveno razlikujejo, bi bila v sami aplikaciji ta omejitev vgrajena kot nastavljen parameter. Tako bi se lahko uporabnik sam odločil, največ kolikšen delež spirale bi želel imeti označen. To bi bilo enostavno izvedljivo z drsnikom, ki bi omogočal hitre spremembe nastavitve parametra.

Prispevek množice intervalov in nevtralne vrednosti

Definicija 5: Če je T' spremenjena vrsta T , kjer smo vrednosti točk znotraj intervala I zamenjali z vrednostmi točk iz nevtralne vrste za atribut A , je *prispevek intervala* (angl. contribution) definiran kot $|A(T) - A(T')|$.

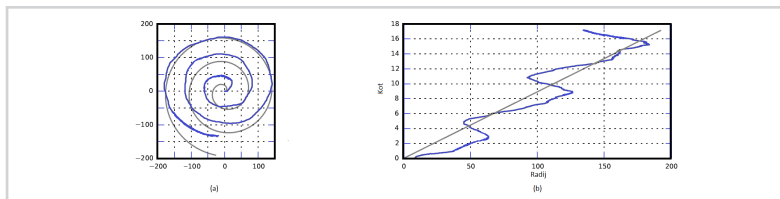
Če je T' spremenjena vrsta T , kjer smo vrednosti točk, ki so v množici intervalov W_a zamenjali z vrednostmi točk iz nevtralne vrste za atribut A , je *prispevek množice intervalov* W_a definiran kot $|A(T) - A(T')|$.

Prispevek intervala je sam po sebi zelo dober pokazatelj, kateri intervali so najpomembnejši. Tako bi se lahko odločili in izrisali n intervalov, ki imajo najvišji prispevek. Če si predstavljamo, da imamo intervale dolžine ena, se v najslabšem primeru lahko zgodi, da intervali, ki imajo največji *prispevek*, niso sosednji in bomo tako obarvali posamezne točke, ki so razpršene po časovni vrsti. Za večino atributov to ni najbolj

primerno, saj nam razpršene točke ne pomenijo kaj dosti z vidika razumevanja odločitve računalniškega modela.

Definicija 6: Nevtralne vrste so tiste vrste, pri katerih je vrednost atributa A enaka 0.

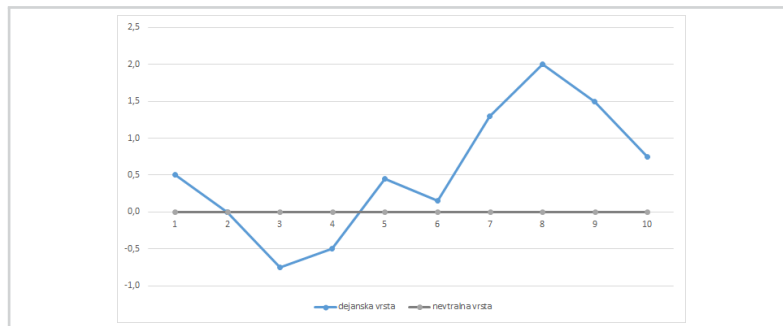
Takšna definicija nevtralne vrste je ustrezna za atribute, ki na nek način merijo odstopanje od idealne vrste. To so takšne vrste, ki so zaželeno, ampak v realnosti verjetno nedosegljive. V naših primerih, predstavljenih v poglavjih 3 in 4, kjer so morali preiskovanci narisati spirale, je idealna vrsta arhimedova spirala, za katero velja, da risanje poteka z enakomerno hitrostjo brez sunkovitih sprememb. Takšna spirala je v polarnem koordinatnem sistemu predstavljena kot naraščajoča premica. V realnosti je torej praktično ničelna verjetnost, da bi preiskovanec narisal idealno arhimedovo spiralo. Primer spirale, ki jo je narisal preiskovanec in pripadajoča idealna spirala sta prikazani na sliki 5.5.



Slika 5.5: Primer spirale, ki jo je narisal preiskovanec (modra barva) in idealna spirala (siva barva). (a) kartezični koordinatni sistem, (b) polarni koordinatni sistem.

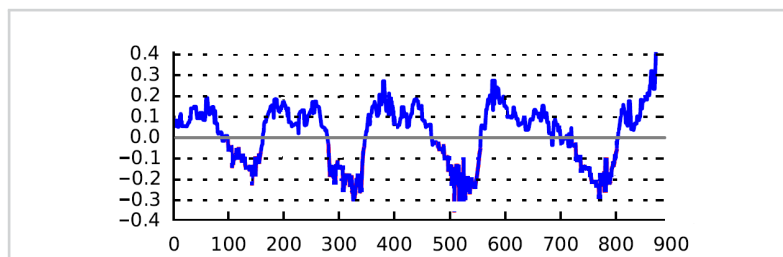
Nevtralne vrednosti se določijo glede na atribut, ki ga računamo. V nadaljevanju je predstavljen nekaj primerov atributov in pripadajočih nevtralnih vrst. Kot prvi primer vzemimo atribut err , ki računa koren povprečja kvadratov napak (angl. root mean squared error - RMSE). Za primer si oglejmo napako glede na vrsto $(0, 0, 0, \dots)$, torej odstopanje dane vrste od 0. Primerna nevtralna vrsta za ta primer je vrsta $(0, 0, 0, \dots)$. Na sliki 5.6 je prikazan hipotetičen primer vrste, na kateri računamo vrednost atributa err . Z modro je narisana vrsta T , s sivo barvo pa nevtralna vrsta.

Slika 5.6: Hipotetičen primer vrste T (modra barva), na kateri računamo vrednost atributa err . S sivo barvo je narisana nevtralna vrsta.

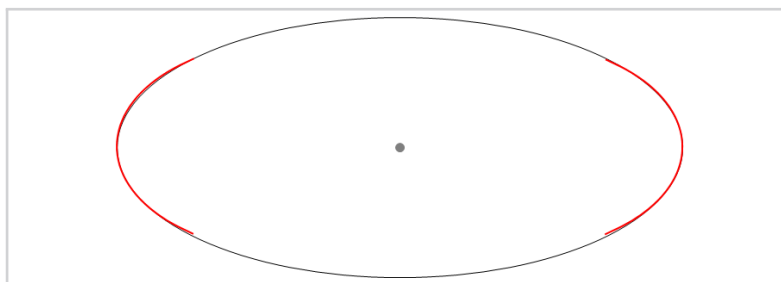


Za atribut err je določitev nevtralne vrste trivialen problem, saj lahko izhajamo iz same definicije izračuna atributa. Podobno velja za atribut $percNeg$, ki predstavlja kolikšen delež spirale je bil risan proti izhodišču spirale. Vrednost atributa torej predstavlja delež vrste, ko je radialna hitrost negativna. Tudi v tem primeru imajo točke v nevtralni vrsti vrednost 0 (torej $(0, 0, 0, \dots)$). Primer vrste, ki predstavlja radialno hitrost risanja spirale in pripadajoča nevtralna vrsta za atribut $percNeg$ sta prikazani na sliki 5.7.

Slika 5.7: Hipotetičen primer vrste T (modra barva), na kateri računamo vrednost atributa $percNeg$. S sivo barvo je narisana nevtralna vrsta.



Predlagani algoritem deluje tudi na atributih, kot je ploščina lika, npr. elipse. V tem primeru je nevtralna vrsta tista, v kateri imajo vse točke vrednost, ki je enaka vrednosti središča lika. V tem primeru bi algoritem označil točke na elipsi, kot je prikazano na sliki 5.8.



Slika 5.8: Primer označitve točk, ki imajo največji vpliv pri izračunu ploščine elipse. S sivo barvo je označeno središče elipse, ki predstavlja vrednosti nevtralne vrste.

Za predstavljene atribute velja, da vsaka točka v vrsti vrednost atributa kvečjemu poveča. To pomeni, da je prispevek posamezne točke, pri konstantni dolžini vrste, nenegativen.

Pri izboru točk nevtralne vrste (za posamezen atribut) moramo izbrati takšne točke, za katere je vrednost atributa A enaka 0 . Ko določamo nevtralno vrsto, izhajamo iz definicije in pomena atributa. Na primer:

- če je atribut odstopanje vrste T od idealne vrste, je nevtralna vrsta enaka idealni vrsti;
- če atribut meri delež točk, za katere velja nek pogoj (npr. vrednosti točk so negativne, pozitivne, večje od povprečja ipd.), potem vrednosti točk v nevtralni vrsti ustrezajo omejitvi tega pogoja (vrednost 0 , povprečna vrednost, ipd.);
- če atribut na nek način računa odstopanje od povprečja (npr. standardna deviacija), so vrednosti v nevtralni vrsti enake povprečni vrednosti točk v vrsti T ;
- če atribut pri izračunu uporablja vrednost odmika od neke točke (npr. pri ploščini kroga je to središče kroga), so nevtralne vrednosti enake vrednosti te točke (središče kroga).

Cena intervalov

Bolj kot večje število razpršenih intervalov si želimo, da se posamezni intervali povežejo v večje intervale. Takšna vizualizacija ima lahko s kognitivnega vidika večji učinek

in lahko pomeni več informacij za eksperta. To smo skušali doseči z vpeljavo *cene* intervalov.

Definicija 7: *Ceno* (angl. cost) intervalov definiramo takole:

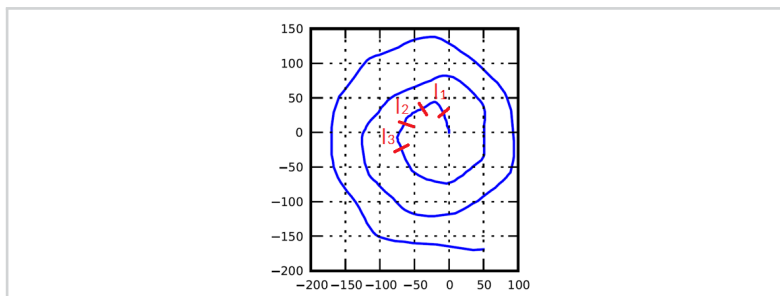
$$\text{cost} = \sum_i \omega * |I| + (1 - \omega) * \log_2(1 + |I|)$$

Cena je sestavljena iz dveh delov, v obeh pa kot argument nastopa dolžina intervala ($|I|$).

Z logaritmskim delom ($\log_2(1 + |I|)$) smo dosegli to, da naša metoda preferira manj daljših intervalov kot več krajših. Uporaba le logaritmskega dela formule se je izkazala za neprimerno, saj preveč kaznuje krajše intervale in smo posledično kot najpomembnejši del časovne vrste vedno dobili en zelo dolg interval. Zato smo formulo uravnotežili tako, da so prišli do izraza tudi nekoliko krajši intervale.

Parameter v enačbi za ceno pa je utež ω . Utež ω nastavimo ročno za vsak atribut, odvisno predvsem od pomena atributa. Z utežjo določimo, ali pri določenem atributu preferiramo krajše intervale na račun večjega števila le-teh ali daljše intervale na račun manjšega števila intervalov. Nastavitve uteži ni mogoče avtomatizirati, saj je odločitev o tem, kaj in kako želimo imeti pobarvano, v veliki meri kognitivna odločitev. Utež je realno število med 0 in 1.

Slika 5.9: Primer spirale z označenimi tremi intervali I_1 , I_2 in I_3 , pri čemer je dolžina posameznega intervala dve enoti. Intervala I_1 in I_2 ter I_2 in I_3 sta sosednja.



Za lažjo ponazoritev je na sliki 5.9 prikazan hipotetičen primer spirale z označenimi tremi intervali I_1 , I_2 in I_3 , pri čemer je dolžina posameznega intervala dve enoti.

Tabela 5.1: Primer cen in koristnosti posameznih intervalov, ki so enake dolžine (dolžine dve) ter kombinacij intervalov (pri čemer lahko intervala I_1 in I_2 ter I_2 in I_3 paroma združimo in tako dobimo dva intervala $I_{1,2}$ in $I_{2,3}$ dolžine štiri; prav tako lahko v en interval ($I_{1,2,3}$) dolžine šest združimo vse tri intervale). Utež ω , ki je potrebna pri izračunu cene, je nastavljena na 0,5. Intervali se med seboj ne prekrivajo.

Interval	Dolžina int.	Število int.	Prispevek	Cena	Koristnost
I_1	2	1	4,0	1,79	2,23
I_2	2	1	5,5	1,79	3,07
I_3	2	1	6,0	1,79	3,35
$I_{1,2}$	4	1	10,0	3,16	3,16
$I_{2,3}$	4	1	12,5	3,16	3,96
$I_{1,3}$	2	2	11,0	3,58	3,07
$I_{1,2,3}$	6	1	16,0	4,40	3,64

Intervala I_1 in I_2 ter I_2 in I_3 sta sosednja. Iz podatkov v tabeli 5.1 je razvidno, da je v primeru združitve sosednjih intervalov (I_1 in I_2 oz. I_2 in I_3), njihova *cena* nižja, kot če bi v množico dodali intervala I_1 in I_3 posamično, ki nista sosednja in ju ni mogoče združiti. Prav tako lahko opazimo, da je v primeru združitve vseh treh intervalov njihova skupna *koristnost* manjša od *koristnosti*, ki jo dobimo, če združimo le intervala I_2 in I_3 . To pomeni, da interval I_1 premalo prispeva h končni koristnosti združenega intervala in je zato bolj smiselno združiti le intervala I_2 in I_3 .

Omejitve predlagane metode vizualizacije

Predlagana metoda vizualizacije je primerna za attribute, ki na nek način merijo odstopanje od idealne vrste in za katere velja:

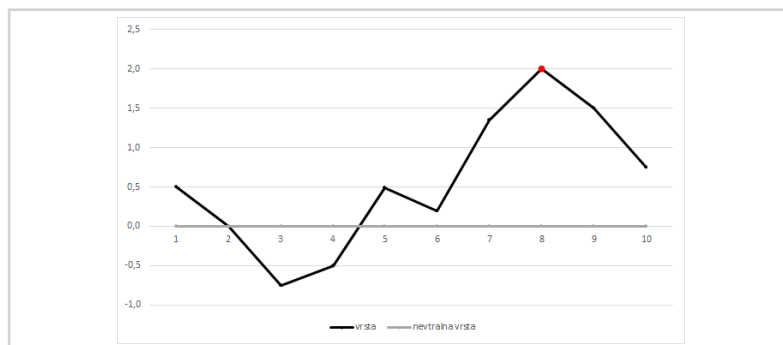
- da je izračunana vrednost atributa numerična vrednost;
- da je prispevek posamezne točke k vrednosti atributa neodvisen od prispevka katere koli druge točke v vrsti.
Primer atributa, za katerega to ne velja, je maksimalna frekvenca spektra podane vrste;
- da vsaka točka v vrsti vrednost atributa kvečjemu poveča. Torej je prispevek

posamezne točke nenegativen.

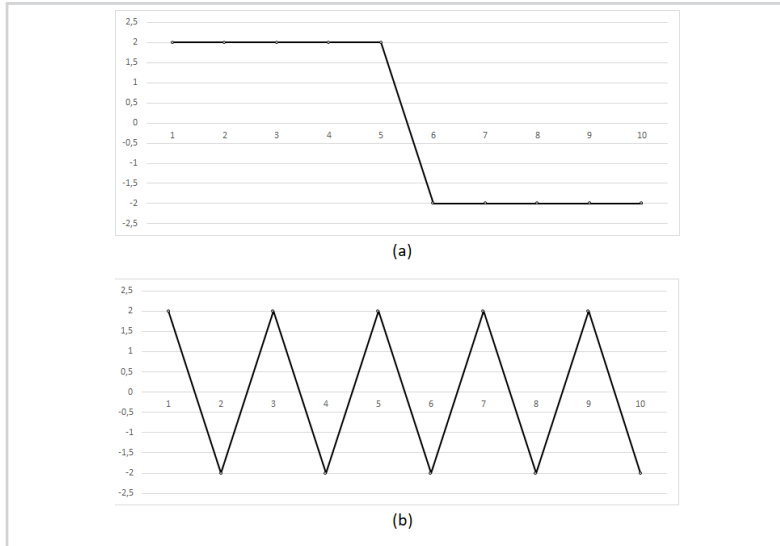
Poleg atributov, ki ustrezajo zgornjim omejitvam (nekaj smo jih predstavili v podpoglavju “Prispevek množice intervalov in nevtralne vrednosti”), se metoda lahko uporabi tudi za nekatere druge attribute, pri katerih je potrebno določene pogoje nekoliko omiliti oz. prilagoditi algoritem. Nekaj primerov navajamo v nadaljevanju.

Vzemimo atribut, ki izračuna maksimalno vrednost v vrsti. V tem primeru lahko, neodvisno od dejanske vrste, vse vrednosti v nevtralni vrsti nastavimo na $-\infty$. Na voljo imamo tudi druge vrednosti v nevtralni vrsti, pri čemer moramo upoštevati vrednosti točk v podani vrsti. Tako so lahko primerne vse točke na odprtem intervalu $(-\infty, \max(T))$. S takšno nastavitvijo vrednosti točk v nevtralni vrsti prekršimo definicijo 6, ki pravi, da morajo biti vrednosti v nevtralni vrsti takšne, da je vrednost atributa, izračunana na nevtralni vrsti, enaka 0. Temu se lahko izognemo tako, da zahtevamo, da je maksimalna vrednost v vrsti pozitivna. V tem primeru bi nevtralno vrsto nastavili na $(0, 0, 0, \dots)$. Primer je predstavljen na sliki 5.10.

Slika 5.10: Primer vrste T (črna barva), na kateri računamo atribut - največja vrednost v vrsti. S sivo barvo je narisana nevtralna vrsta, z rdečo pa točka, ki bi jo označil naš algoritem.



Na podoben način bi nastavili vrednosti točk nevtralne vrste za atribut, ki bi iskal točko z najmanjšo vrednostjo v vrsti ali atribut, ki bi izračunal razpon točk v vrsti ($\max - \min$; vrednosti na odprtem intervalu $(\min(T), \max(T))$). Kljub temu, da bi predlagani algoritem v principu deloval za te tri attribute (največja vrednost, najmanjša vrednost in razpon vrednosti), obstajajo določene omejitve, kjer algoritem rešitve ne bi našel. To bi se zgodilo v primeru, ko bi bila v vrsti več kot ena točka z najvišjo ali najnižjo vrednostjo. Na sliki 5.11 sta prikazana dva primera takšnih vrst.



Slika 5.11: Hipotetična primera vrste T , na kateri predlagani algoritem za atribute največja vrednost, najmanjša vrednost in razpon ne bi deloval in ne bi označil nobene točke.

S prilagoditvijo algoritma na način, da bi bili začetni intervali daljši od ene točke, bi v primeru (a), ki je prikazan na sliki 5.11 algoritem našel rešitev, če bi bili intervali dolgi vsaj toliko, kolikor je zaporednih točk z najvišjo vrednostjo. Kljub temu pa rešitve z uporabo predlaganega algoritma ne bi dobili v primeru (b) (razen, če bi bil interval dolg toliko, kolikor je dolga celotna vrsta, pri čemer bi vizualizacija izgubila pomen).

Druga možna prilagoditev predlaganega algoritma, ki bi delovala tako za primer (a), kot tudi za primer (b) s slike 5.11, bi bila naslednja. Kot vemo, prispevek intervala računamo tako, da vrednosti atributa na vrsti T ($A(T)$) odštejemo vrednost atributa na vrsti T' ($A(T')$). Če intervali na začetku vsebujejo le eno točko, bo prispevek vseh intervalov enak 0. Kar pomeni, da bo tudi koristnost vseh intervalov enaka 0 in posledično nobenega intervala ne bomo vključili v rešitev. Ker vemo, da je vrednost atributa v tem primeru neodvisna od dolžine vrste (saj iščemo vsaj eno točko, ki ima največjo ali najmanjšo vrednost), lahko spremenimo način računanja prispevka tako, da je prispevek intervala I kar enak vrednosti atributa na tem intervalu ($A(I)$). Tako bo algoritem pravilno določil koristnost posameznega intervala in našel vse tiste točke, ki imajo največjo oz. najmanjšo vrednost v vrsti.

Predlagana metoda je primerna za večino pomembnih atributov, ki smo jih zgradili za potrebe gradnje diagnostičnega modela za razločevanje tremorjev na podlagi digitalne spirografije. Atributi so predstavljeni v poglavju 4, tu pa navajamo tiste, za katere je predlagana metoda vizualizacije primerna: `Xsymm`, `Ysymm`, `nPeaks`, `avgP`, `stdevP`, `err`, `err0`, `errBF`, `rmse100`, `rmse150` ter `percNeg`.

Atributi, predstavljeni v poglavju 4, za katere predlagana metoda vizualizacije ni primerna, so: `angVsRad`, `template.distance`, `t.rho`, `incDev`, `rmseSum`, `rmseMin`, `freq.dev`, `peakAmpXOut` ter dodatni atributi, ki merijo minimum, maksimum, povprečje in vsoto (posameznih naštetih atributov) na vseh štirih spiralah, ki jih je narisal preiskovanec. Omeniti velja, da sta med najpomembnejšimi atributi v končnem diagnostičnem modelu le dva atributa (`rmseSum` in `angVsRad.template.distance`), za katera predlagana metoda vizualizacije ni primerna.

5.1.1 Algoritem

V nadaljevanju je predstavljena psevdokoda algoritma za iskanje najpomembnejših odsekov na časovni vrsti (algoritem 1).

Kot smo že nekajkrat omenili, nam model, predstavljen v poglavju 4, za vsak obravnavani atribut določa mejo, kdaj se le-ta uporablja v odločitvenem procesu in kdaj ne. Zato je prvi korak našega algoritma izračun vrednosti obravnavanega atributa na posamezni časovni vrsti (torej za posamezen test). Če je vrednost atributa dovolj visoka (torej preko meje, ki jo določa odločitveni model), nadaljujemo z iskanjem najpomembnejših odsekov in vizualizacijo, sicer označevanja odsekov z uporabo tega atributa ne izvedemo, saj le-ta ni vplival na samo klasifikacijo primera.

V prilogi B je prikazan primer sledi izvajanja algoritma za atribut RMSE.

Algoritem 1 Algoritem za iskanje najpomembnejših odsekov v vrsti.

Vhodi:

- vrsta T - zaporedje točk t_1, t_2, \dots, t_n ,
- ω - utež za zračun cene intervala,
- ustavitvena vrednost (thr) - največji delež točk, ki jih lahko označimo.
- nevtralna vrsta ($neutr$) - zaporedje točk v_1, v_2, \dots, v_n .

Izhoda:

- največja koristnost,
 - seznam, katere točke so izbrane za označitev.
-

function POISCIINTERVALE($T, \omega, thr, neutr$)

$N \leftarrow \text{length}(T)$ ▷ Izračunaj dolžino vrste T .

▷ Seznam, v katerem hranimo podatke o izbranih točkah.

$best_selector \leftarrow \underbrace{[0, \dots, 0]}_N$ ▷ 1 - točka je izbrana, 0 - točka ni izbrana.

$att_val \leftarrow \text{CALCATT}(T, neutr, best_selector)$ ▷ Izračunaj vrednost atributa.

$best_benefit \leftarrow 0$

▷ Število označenih točk mora biti manjše od deleža točk, ki jih

while ($\sum_{j=0}^{\text{length}(best_selector)-1} best_selector[j]) < thr \cdot N$ *do* ▷ lahko označimo.

$change \leftarrow \text{FALSE}$

$selector \leftarrow best_selector$

for $x \leftarrow 0, 1, \dots, N - 1$ *do*

if $selector[x] = 1$ *then*

continue

else

$selector[x] \leftarrow 1$ ▷ V množico za označitev dodamo novo točko.

$prispevek \leftarrow | att_val - \text{CALCATT}(T, neutr, selector) |$ ▷ Izračunaj prispevek.

$il \leftarrow \text{INTERVALENGTHS}(selector)$ ▷ Izračunaj dolžine intervalov.

$cost \leftarrow \sum_{i=0}^{\text{length}(il)-1} il[i] \cdot \omega + (1 - \omega) \cdot \log_2(il[i] + 1)$ ▷ Cena intervalov.

$benefit \leftarrow \begin{cases} \frac{prispevek}{cost} & \text{if } cost \neq 0 \\ 0 & \text{otherwise} \end{cases}$ ▷ Koristnost množice intervalov/točk.

if $benefit > best_benefit$ *then*

$best_benefit \leftarrow benefit$ ▷ Shrani trenutno največjo koristnost.

$change \leftarrow \text{TRUE}$

$best_selector \leftarrow selector$ ▷ Točke z največjo koristnostjo.

$selector[x] \leftarrow 0$

if $\neg change$ *then*

```

break
return (best_selector, best_benefit)

```

▷ Vrni seznam točk za označitev
▷ in njihovo koristnost.

Funkcija `CALCATT(T , $neutr$, $selector$)` izračuna vrednost atributa na vrsti.

Vhod:

- vrsta T - zaporedje točk t_1, t_2, \dots, t_n ,
- nevtralna vrsta ($neutr$) - zaporedje točk v_1, v_2, \dots, v_n ,
- $selector$ - seznam (dolžine N), ki določa, katere točke so izbrane za označitev; 1 - točka je izbrana, 0 - točka ni izbrana.

Izhod:

- vrednost atributa.

```

function CALCATT( $T$ ,  $neutr$ ,  $selector$ )
  if INTERVALLENGTHS( $selector$ ) = [] then
    return att_val( $T$ )
  else
     $T' \leftarrow T$ 
    for  $id \leftarrow 0, 1, \dots, N - 1$  do
      if  $selector[id] == 1$  then
         $T'[id] = neutr[id]$ 
    return att_val( $T'$ )

```

▷ Nobena točka ni bila izbrana.
▷ Izračunaj in vrni vrednost atributa na osnovni vrsti T .
▷ Točka je bila izbrana.
▷ Nadomesti vrednost točke v vrsti T' z vrednostjo
▷ istoležne točke v nevtralni vrsti.
▷ Izračunaj in vrni vrednost atributa na spremenjeni vrsti T' .

Funkcija `INTERVALLENGTHS(p)` izračuna seznam dolžin intervalov enic v seznamu p .

Primer: `INTERVALLENGTHS([0,0,1,1,1,0,1,0,0,0,1,1,0])=[3,1,2]`.

5.2 Podatki in evalvacija metode

Podatki, ki so bili uporabljeni za evalvacijo metode vizualizacije, so bili zbrani za potrebe gradnje odločitvenega modela, predstavljenega v poglavju 4. Vzeli smo podmnožico zbranih podatkov, ki je vsebovala 45 preiskovancev, od tega 15 preiskovancev s parkinsonovo boleznijo, 15 preiskovancev z esencialnim tremorjem in 15 zdravih kontrol.

Odločitveni model iz prejšnjega poglavja je pravilno napovedal razred za devet preiskovancev s parkinsonovo boleznijo, sedem preiskovancev z esencialnim tremorjem in

vse zdrave kontrole.

Evalvacijo metode za vizualno razlago diagnoz smo izvedli v sodelovanju z eksperti s področja spirografije. Postavili smo spletno aplikacijo¹, s katero smo ocenjevalcem prikazovali spirale in zbirali njihove odgovore. Aplikacija omogoča animacijo posamezne spirale, tako da so ocenjevalci lahko videli tudi spremembe v hitrostih risanja in morebitna obotavljanja.

K sodelovanju smo povabili pet ekspertov, ki delujejo večinoma v tujini. Sodelovanje so potrdili trije izmed njih.

Nevrolog D. N. z Uppsala University se ukvarja z raziskovanjem učinkovanja globoke možganske stimulacije in učinkovine levodopa pri bolnikih s parkinsonovo boleznijo ter kako se premajhni oz. preveliki odmerki le-te odražajo pri spirografiji.

Specializant nevrologije D. G. z University College London in Nevrološke klinike Univerzitetnega kliničnega centra v Ljubljani se raziskovalno ukvarja z globoko možgansko stimulacijo pri bolnikih s parkinsonovo boleznijo, več let pa se ukvarja tudi s spirografijo pri kliničnih pregledih.

Nevrolog D. H. z National Institutes of Health (ZDA) in Medical University of Vienna se raziskovalno ukvarja z motnjami gibanja, predvsem s parkinsonovo boleznijo, distonijami, esencialnim tremorjem in psihogenimi motnjami gibanja. Med drugim je objavil več člankov na temo uporabe spirografije za vrednotenje esencialnega tremorja.

Navodila, ki smo jih posredovali ocenjevalcem, so v prilogi C.

5.2.1 Nastavitve parametrov

Za potrebe vizualizacije smo morali nastaviti dva parametra. Prvi je parameter ω , ki nastopa v *cost* funkciji. Ta parameter je potrebno ročno nastaviti za vsak atribut. V tabeli 5.2 so predstavljene nastavitve parametra za posamezen atribut, ki smo ga vizualizirali.

Drugi parameter, ki smo ga morali nastaviti, je bila ustavitvena vrednost. Ta je bila za vse attribute nastavljena tako, da smo smeli označiti največ 25% vrste.

¹Spletna aplikacija je dostopna na naslovu <https://www.parkinsoncheck.net/vidadoctor/>. Za dostop potrebujete uporabniško ime in geslo, ki ju dobite pri avtorici tega dela.

Tabela 5.2: Nastavitve parametra ω za posamezen atribut, ki smo ga vizualizirali.

Atribut	ω
LT ₃ .err	0,4
RT ₃ .errtBF	0,4
RT ₃ .erro	0,5
LFerr	0,5
LFerro	0,4
RF.percNego	0,5
RF.rmse100	0,6
RF.rmse150	0,3
absSp.stdevP	0,3
radSp.stdevP	0,3

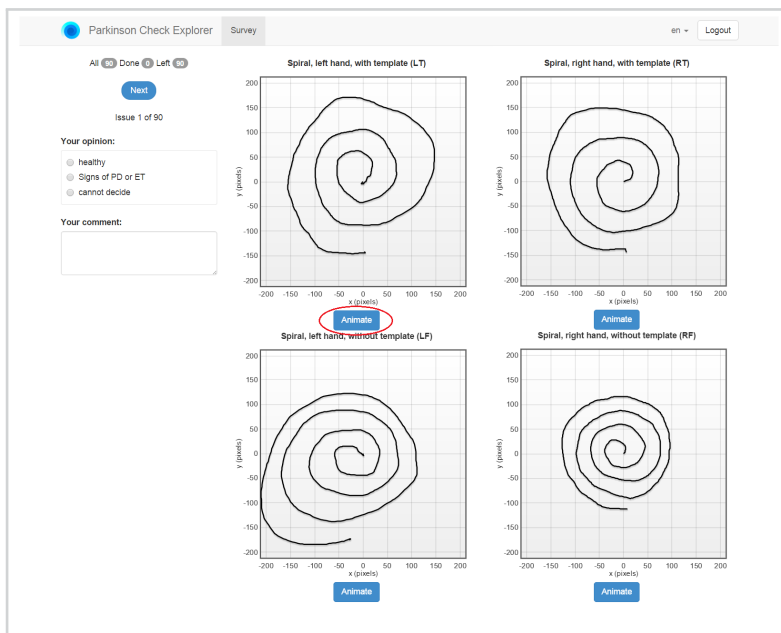
5.2.2 Zasnova eksperimenta

Eksperiment smo zasnovali v dveh korakih:

1. V prvem koraku smo ekspertom pokazali po štiri spirale, ki so jih narisali preiskovanci. Primer spiral enega izmed preiskovancev je viden na sliki 5.12.

Zgornja leva spirala je bila narisana z ne-vodilno roko (v nadaljevanju bomo govorili o spiralah, narisanih z levo roko) po predlogi (preiskovanec je moral čim bolj natančno slediti črti že narisane spirale). Spodnja leva spirala je bila prav tako narisana z levo roko, a tokrat "prostoročno" (brez podane slike spirale). Desna zgornja spirala je bila narisana z vodilno roko (v nadaljevanju bomo govorili o spiralah, narisanih z desno roko) po predlogi, spodnja desna spirala pa je bila narisana prostoročno z desno roko. Ocenjevalci so lahko posamezno spiralo animirali s pritiskom na gumb "Animate" in jo zaustavili s pritiskom na gumb "Stop". Na podlagi teh spiral je ocenjevalec podal svoje mnenje o preiskovancu, ki je spirale narisal. Možnosti, med katerimi so lahko izbirali, so bile:

- zdrav,
- vidni so znaki parkinsonove bolezni ali esencialnega tremorja,
 - vidni so znaki parkinsonove bolezni,



Slika 5.12: Primer štirih spiral, ki so jih morali ocenjevalci oceniti.

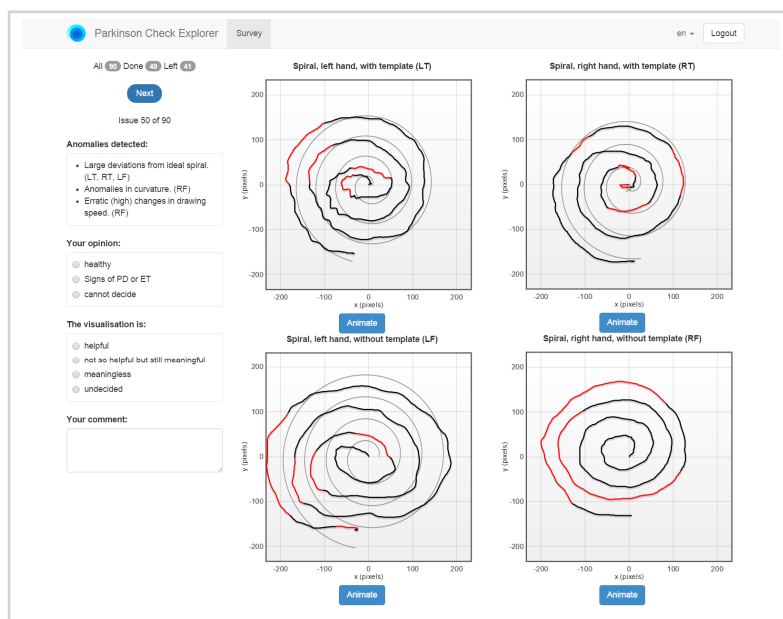
- vidni so znaki esencialnega tremorja,
- oboje (vidni so znaki parkinsonove bolezni in esencialnega tremorja),
- ne morem se odločiti, ali pripadajo znaki parkinsonovi bolezni ali esencialnemu tremorju,
- ne morem se odločiti, ali je preiskovanec zdrav ali ne.

Kot je razvidno iz možnih odgovorov, se je moral ocenjevalec najprej odločiti med tem, ali je preiskovanec zdrav ali ima prisotne znake parkinsonove bolezni oz. esencialnega tremorja. V primeru slednjega se je potem nadalje odločal med tem, kateri bolezni znaki dejansko pripadajo. Torej se je ocenjevalec moral odločati na podoben način, kot deluje naš odločitveni model iz poglavja 4.

Da bi bolje razumeli, zakaj so se ocenjevalci odločili za posamezen odgovor, smo jih prosili, da podajo komentar na svojo odločitev.

Ko so ocenili spirale vseh 45 preiskovancev, je sledil drugi korak.

2. V drugem koraku smo ekspertom ponovno pokazali po štiri spirale istih 45 preiskovancev v enakem vrstnem redu kot v prvem koraku. Razlika je bila v tem, da so bili tokrat delčki nekaterih (ali vseh) spiral obarvani z rdečo barvo (na statični sliki) oz. z modro barvo (v animaciji). Primer obarvanih spiral je predstavljen na sliki 5.13.



Slika 5.13: Primer vizualizacije pomembnih delov v spiralah.

Obarvani deli spiral predstavljajo najpomembnejše dele spiral, ki jih je odkrila naša metoda za določen atribut. Na stacionarni sliki smo prikazovali attribute, ki niso vezani na hitrosti (npr. odstopanje od idealne spirale), v animacijah pa smo prikazovali attribute, ki so izpeljani iz različnih hitrosti (npr. hitro spreminjanje hitrosti - pospeški in pojemki). Pri spiralah, kjer smo označevali attribute, ki iščejo odstopanja narisane spirale od idealne spirale, smo poleg barvanja delov spirale z rdečo barvo izrisali še idealno spiralo. S takšno vizualizacijo je lažje videti napake v narisanih spiralah. Poleg vizualnih oznak smo za vsako označeno anomalijo podali še kratko razlago, kaj je računalnik v posamezni spirali zaznal. Razlage, ki smo jih podali ocenjevalcem, so bile:

- “Visoka odstopanja od idealne spirale” - narisana spirala je odstopala od “idealne” spirale. Pri čemer “idealna” spirala ni spirala s predloge, ampak je to spirala, ki je bila določena kot najbolj priljegajoča se spirala narisani spirali (angl. best-fit spiral).
- “Znaten delež spirale je bil risan proti izhodišču spirale” – radij arhimedove spirale mora monotonno naraščati. Izmerili smo delež spirale, kjer se je radij zmanjševal.
- “Anomalije v ukrivljenosti” - računalnik je zaznal anomalije v ukrivljenosti, kot so npr. ostri koti v spirali.
- “Neenakomerne (visoke) spremembe v hitrosti risanja” - hitrost risanja ni bila konstantna, ampak se je občutno spreminjala (naraščala in padala) na nekaterih delih spirale.
- “Neenakomerne (visoke) spremembe v radialni hitrosti” - pomeni, da radialna hitrost (radialna hitrost meri hitrost proti ali stran od začetne točke spirale) ni bila konstantna, ampak se je občutno spreminjala (naraščala in padala) na nekaterih delih spirale.

Poleg posamezne razlage smo v oklepaju še zapisali, na katerih spiralah je bila določena anomalija zaznana in tudi prikazana.

Na podlagi obarvanih spiral in razlag zaznanih anomalij so ocenjevalci ponovno podali svoja mnenja o osebi, ki je spirale narisala. Možnosti, med katerimi so izbirali, so bile enake kot v prvem koraku. Poleg tega smo jih povprašali še, kakšna se jim je zdela vizualizacija. Izbirali so lahko med naslednjimi odgovori:

- bila mi je v pomoč pri podajanju mnenja,
- ni mi bila v pomoč pri podajanju mnenja, ampak je obarvala smiselne predele spiral,
- obarvani predeli spiral niso smiselni,
- ne morem se odločiti.

Tudi tokrat smo jih prosili, da podajo komentar na svojo odločitev pri ocenjevanju spiral in pri ocenjevanju smiselnosti vizualizacije.

5.3 Rezultati

V nadaljevanju so predstavljeni rezultati evalvacije naše metode, interpretacija rezultatov in komentarji pa so podani v diskusiji.

V tabeli 5.3 so predstavljeni rezultati klasifikacije primerov v razreda *zdrav* in *znaki bolezni* ocenjevalca D. G. v prvem koraku, torej brez vizualizacije atributov. V tabeli niso prikazani primeri, za katere se ocenjevalec ni mogel odločiti, v kateri razred bi jih uvrstil. Ocenjevalec je v razred *zdrav* pravilno uvrstil 13 primerov (od 15), v razred *znaki bolezni* pa 21 primerov (od 30). Klasifikacijska točnost v prvem koraku je bila 0,83.

Tabela 5.3: Matrika zamenjav ocenjevalca D. G. pri klasifikaciji v razreda *zdrav* in *znaki bolezni* - prvi korak, brez vizualizacije atributov.

		Napovedani razred	
		Zdrav	Znaki bolezni
Razred	Zdrav	13	1
	Znaki bolezni	6	21

V tabeli 5.4 so predstavljeni rezultati klasifikacije primerov v prej omenjena razreda s strani istega ocenjevalca v drugem koraku, kjer so bili vizualizirani atributi. Tokrat je ocenjevalec v razred *zdrav* pravilno uvrstil sedem primerov (od 15) in v razred *znaki bolezni* 27 primerov (od 30). Klasifikacijska točnost ocenjevalca D. G. v drugem koraku je bila 0,87.

Tabela 5.4: Matrika zamenjav ocenjevalca D. G. pri klasifikaciji v razreda *zdrav* in *znaki bolezni* - drugi korak, z vizualizacijo atributov.

		Napovedani razred	
		Zdrav	Znaki bolezni
Razred	Zdrav	7	4
	Znaki bolezni	1	27

V tabeli 5.5 prikazujemo rezultate klasifikacije primerov v razreda *zdrav* in *znaki bolezni* za ocenjevalca D. N. v prvem koraku. Ocenjevalec je v razred *zdrav* pravilno

uvrstil 13 primerov (od 15), v razred *znaki bolezni* pa 17 primerov (od 30). Klasifikacijska točnost je bila 0,68.

Tabela 5.5: Matrika zamenjav ocenjevalca D. N. pri klasifikaciji v razreda *zdrav* in *znaki bolezni* - prvi korak, brez vizualizacije atributov.

		Napovedani razred	
		Zdrav	Znaki bolezni
Razred	Zdrav	13	2
	Znaki bolezni	12	17

Rezultati klasifikacije v razreda *zdrav* in *znaki bolezni* ocenjevalca D. N. v drugem koraku, kjer so bili na spiralah vizualizirani atributi, so podani v tabeli 5.6. Ocenjevalec je v razred *zdrav* pravilno razvrstil devet primerov (od 15) in v razred *znaki bolezni* 24 primerov (od 30). Klasifikacijska točnost je bila tokrat 0,79.

Tabela 5.6: Matrika zamenjav ocenjevalca D. N. pri klasifikaciji v razreda *zdrav* in *znaki bolezni* - drugi korak, z vizualizacijo atributov.

		Napovedani razred	
		Zdrav	Znaki bolezni
Razred	Zdrav	9	4
	Znaki bolezni	5	24

Ocenjevalec D. H. je v prvem koraku v razred *zdrav* pravilno uvrstil pet primerov (od 15), v razred *znaki bolezni* pa 28 primerov (od 30). Klasifikacijska točnost ocenjevalca v tem koraku je bila 0,79. Podrobnejši rezultati so razvidni iz tabele 5.7.

Rezultati klasifikacije primerov v razreda *zdrav* in *znaki bolezni* v drugem koraku, kjer so bili na spiralah vizualizirani atributi, so predstavljeni v tabeli 5.8. Ocenjevalec D. H. je v razred *zdrav* pravilno uvrstil 2 primera (od 15) in v razred *znaki bolezni* 28 primerov (od 30). Klasifikacijska točnost ocenjevalca je bila v drugem koraku 0,77.

V nadaljevanju prikazujemo podrobnejše rezultate klasifikacij s strani naših ekspertov.

Tabela 5.9 prikazuje matriko zamenjav ocenjevalca D. G. v koraku, v katerem nismo vizualizirali atributov. Iz tabele je razvidno, da se za štiri primere ni mogel odločiti, v

Tabela 5.7: Matrika zamenjav ocenjevalca D. H. pri klasifikaciji v razreda *zdrav* in *znaki bolezni* - prvi korak, brez vizualizacije atributov.

		Napovedani razred	
		Zdrav	Znaki bolezni
Razred	Zdrav	5	8
	Znaki bolezni	1	28

Tabela 5.8: Matrika zamenjav ocenjevalca D. H. pri klasifikaciji v razreda *zdrav* in *znaki bolezni* - drugi korak, z vizualizacijo atributov.

		Napovedani razred	
		Zdrav	Znaki bolezni
Razred	Zdrav	2	8
	Znaki bolezni	1	28

kateri razred bi jih uvrstil. Enega zdravega je napačno razvrstil v razred PB in dva primera iz razreda PB je napačno razvrstil v razred ET. Pravilno je razvrstil 9 primerov iz razreda PB in 8 primerov iz razreda ET. Klasifikacijska točnost ocenjevalca na tro-razrednem problemu (zdrav, PB, ET) je bila 0,77.

Tabela 5.9: Matrika zamenjav ocenjevalca D. G. - prvi korak, brez vizualizacije atributov.

		Napovedani razred					Ne morem se odločiti
		Zdrav	Znaki PB oz. ET			Ne morem se odločiti	
			PB	ET	Oboje		
Razred	Zdrav	13	1	0	0	0	1
	PB	3	9	2	1	0	0
	ET	3	0	8	1	0	3

V drugem koraku, kjer smo vizualizirali attribute, je ocenjevalec D. G. napačno razvrstil štiri primere iz razreda *zdrav* v razred *znaki bolezni*, ter dva primera iz razreda PB

v razred ET in tri primere iz razreda ET v razred PB. Pravilno je razvrstil 10 primerov iz razreda PB in devet primerov iz razreda ET. Za šest primerov se ni mogel odločiti, v kateri razred spadajo. Klasifikacijska točnost ocenjevalca v drugem koraku na tro-razrednem problemu je bila 0,76. Matrika zamenjav je predstavljena v tabeli 5.10.

Tabela 5.10: Matrika zamenjav ocenjevalca D. G. - drugi korak, z vizualizacijo atributov.

		Napovedani razred					Ne morem se odločiti
		Zdrav	Znaki PB oz. ET			Ne morem se odločiti	
			PB	ET	Oboje		
Razred	Zdrav	7	2	0	2	0	4
	PB	1	10	2	1	1	0
	ET	0	3	9	1	0	2

Matrika zamenjav za ocenjevalca D. N. v prvem koraku, ko ni bilo vizualiziranih atributov, je v tabeli 5.11. Ocenjevalec se za en primer ni mogel odločiti, ali spada med zdrave ali med tiste, ki izkazuje znake bolezni, še za enega pa se ni mogel odločiti, ali izkazuje znake PB ali ET. Napačno je v razred PB klasificiral dva primera iz razreda *zdrav*. Pravilno je klasificiral pet primerov iz razreda PB in šest primerov iz razreda ET, napačno pa je dva primera iz razreda PB uvrstil v razred ET ter tri primere iz razreda ET v razred PB. V tem primeru je bila klasifikacijska točnost ocenjevalca na trirazrednem problemu 0,56.

Tabela 5.12 prikazuje matriko zamenjav ocenjevalca D. N. v drugem koraku, kjer smo vizualizirali attribute. Tokrat se ocenjevalec za tri primere ni mogel odločiti, v kateri razred bi jih umestil. Štiri primere iz razreda *zdrav* je napačno klasificiral v razred PB. Pravilno je klasificiral devet primerov iz razreda PB in pet primerov iz razreda ET, napačno pa je klasificiral devet primerov iz razreda ET v razred PB in en primer iz razreda PB v razred ET. Klasifikacijska točnost je bila na trirazrednem problemu 0,55.

Rezultati klasifikacije primerov v razrede v prvem koraku s strani ocenjevalca D. H. so predstavljeni v matriki zamenjav 5.13. Ocenjevalec se za tri primere ni mogel odločiti, v kateri razred spadajo. Za osem primerov iz razreda *zdrav* je določil, da spadajo v razred *znaki bolezni*, od tega trije primeri v razred PB, dva v razred ET, za

Tabela 5.11: Matrika zamenjav ocenjevalca D. N. - prvi korak, brez vizualizacije atributov.

		Napovedani razred					
		Zdrav	Znaki PB oz. ET			Ne morem se odločiti	
			PB	ET	Oboje	Ne morem se odločiti	
Razred	Zdrav	13	2	0	0	0	0
	PB	8	5	2	0	0	0
	ET	4	3	6	0	1	1

Tabela 5.12: Matrika zamenjav ocenjevalca D. N. - drugi korak, z vizualizacijo atributov.

		Napovedani razred					
		Zdrav	Znaki PB oz. ET			Ne morem se odločiti	
			PB	ET	Oboje	Ne morem se odločiti	
Razred	Zdrav	9	4	0	0	0	2
	PB	4	9	1	0	0	1
	ET	1	9	5	0	0	0

preostale tri pa se ni mogel odločiti. Poleg tega je pet primerov iz razreda PB klasificiral v razred ET, enega v razred *zdrav*, za tri pa se ni mogel odločiti znaki katere bolezni so izraženi. Tri primere iz razreda ET je napačno klasificiral v razred PB, medtem ko se za enega ni mogel odločiti, katere bolezenske znake nakazujejo spirale. Iz tega sledi, da je bila klasifikacijska točnost na trirazrednem problemu 0,6.

Tabela 5.14 prikazuje matriko zamenjav ocenjevalca D. H. v drugem koraku, kjer smo vizualizirali attribute. Tokrat se ocenjevalec za šest primerov ni mogel odločiti, ali spirale izkazujejo znake bolezni ali ne, največ od tega, pet primerov, izhaja iz razreda *zdrav*. Poleg tega je štiri primere razvrstil v razred PB, tri v razred ET, za enega pa je napovedal, da izkazuje tako znake PB kot tudi znake ET. Za primere, ki spadajo v razred PB je v treh primerih napovedal, da spadajo v razred ET, v štirih pa se ni mogel odločiti, ali spadajo v razred ET ali PB, v vsakem primeru pa naj bi izkazovali znake

Tabela 5.13: Matrika zamenjav ocenjevalca D. H. - prvi korak, brez vizualizacije atributov.

		Napovedani razred					
		Zdrav	Znaki PB oz. ET			Ne morem se odločiti	
			PB	ET	Oboje	Ne morem se odločiti	
Razred	Zdrav	5	3	2	0	3	2
	PB	1	6	5	0	3	0
	ET	0	3	10	1	0	1

bolezni. Za enega je napačno napovedal, da spada v razred *zdrav*. Pri primerih iz razreda ET je za vse razene enega napovedal, da spirale izkazuje znake bolezni. Od tega je za en primer napovedal, da spada v razred PB, za enega, da izkazuje tako znake PB kot tudi znake ET, za pet primerov se ni mogel odločiti za znake katere bolezni gre. Klasifikacijska točnost na trirazrednem problemu je tako 0,57.

Tabela 5.14: Matrika zamenjav ocenjevalca D. H. - drugi korak, z vizualizacijo atributov.

		Napovedani razred					
		Zdrav	Znaki PB oz. ET			Ne morem se odločiti	
			PB	ET	Oboje	Ne morem se odločiti	
Razred	Zdrav	2	4	3	1	0	5
	PB	1	7	3	0	4	0
	ET	0	1	7	1	5	1

Poleg klasifikacije primerov v razrede smo ocenjevalce prosili še, da ocenijo vizualizacijo. Skupni rezultati po ocenjevalcih so predstavljeni v tabeli 5.15. Ocenjevalec D. G. se za osem primerov ni mogel odločiti, ali mu je bila vizualizacija v pomoč ali ne, ocenjevalec D. N. in D. H. pa za šest primerov. Z izjemo treh primerov, ko je ocenjevalec D. N. ocenil, da vizualizacija ni obarvala smiselnih predelov na spirali in petih primerov, ko je enako oceno podal ocenjevalec D. H., so vsi trije ocenjevalci pozitiv-

no ocenili smiselnost in uporabnost vizualizacije. Ocenjevalec D. G. je v 37 primerih določil, da mu je bila vizualizacija v pomoč, ocenjevalcu D. N. je bila v 18 primerih v pomoč, v 18 primerih pa je ocenil, da je vizualizacija označila smiselne odseke spiral. Ocenjevalcu D. H. je bila vizualizacija v pomoč pri 11 primerih, za 23 primerov pa je označil, da je označila smiselne dele spiral.

Tabela 5.15: Ocene uporabnosti in smiselnosti vizualizacije.

Vizualizacija	D. G.	D. N.	D. H.
v pomoč	37	18	11
ni bila v pomoč, je označila smiselne odseke spiral	0	18	23
ni obarvala smiselnih predelov	0	3	5
ne morem se odločiti	8	6	6

5.3.1 Analiza optimalnosti algoritma

Kot je razvidno iz samega opisa algoritma, naš algoritem za reševanje problema iskanja napomembnejših odsekov vrste temelji na požrešnem pristopu. Razlog za tak pristop je v prostorski in časovni kompleksnosti problema. Če bi želeli pregledati vse možne kombinacije točk, bi morali preiskati 2^n kombinacij, pri čemer je n število točk v vrsti, ki jo analiziramo. V našem primeru so imele posamezne vrste med 1000 in 2500 točk, kar pomeni, da bi morali pregledati med 2^{1000} in 2^{2500} kombinacij točk. To pa je že z vidika časovne kompleksnosti neizvedljivo.

Ker je naš algoritem v splošnem požrešen, smo želeli vedeti, kako so naši rezultati primerljivi z rezultati, ki bi jih dobili, če bi lahko uporabili popoln algoritem. V ta namen smo naredili analizo optimalnosti algoritma, ki je bila narejena na manjšem številu točk, poleg tega pa smo spreminjali nekaj osnovnih nastavitev, ki jih uporablja požrešen algoritem.

Največje število točk, na katerem smo še lahko pognali popoln algoritem, je bilo 20. Ker nas je zanimalo, koliko na rezultate našega algoritma vpliva število točk, smo zgenerirali različne vrste dolžine 20, 16, 12 ter 8 točk na intervalu $[-1,5, 1,5]$. Točke so bile izbrane naključno z uniformno porazdelitvijo. Pri vrstah dolžine 20 točk smo varirali tudi največji delež točk, ki ga lahko označimo (25% oz. 50%) ter utež ω (0,1, 0,5 in 0,9). Za različne kombinacije nastavitev smo zgenerirali po 100 različnih vrst,

dobljeni rezultati pa so predstavljeni v tabeli 5.16. Primerjava je narejena na podlagi dobljene koristnosti, ki jo dobimo s popolnim oz. požrešnim algoritmom, v tabeli pa navajamo, kolikšen delež rezultata, dobljenega s popolnim algoritmom, smo dosegli s požrešnim.

Tabela 5.16: Analiza delovanja požrešnega algoritma v primerjavi s popolnim algoritmom.

Test	število točk	interval	največji delež označenih točk	ω (v %)	povprečje (v %)	min (v %)	max
test 1	20	[-1,5, 1,5]	25%	0,5	98,93	89,75	100
test 2	20	[-1,5, 1,5]	50%	0,5	96,21	83,04	100
test 3	20	[-1,5, 1,5]	25%	0,1	93,89	76,04	100
test 4	20	[-1,5, 1,5]	25%	0,9	99,73	95,62	100
test 5	16	[-1,5, 1,5]	25%	0,5	99,49	92,72	100
test 6	12	[-1,5, 1,5]	25%	0,5	99,33	89,92	100
test 7	8	[-1,5, 1,5]	25%	0,5	99,80	91,24	100

Iz tabele 5.16 je razvidno, da so rezultati našega (požrešnega) algoritma primerljivi z rezultati popolnega algoritma. V povprečju je naš algoritem dosegal vrednosti, ki se gibljejo okoli 98,5% vrednosti rezultatov popolnega algoritma. Iz tabele je razvidno, da ima vpliv na rezultate največji dovoljeni delež točk, ki jih lahko označimo, pri čemer so rezultati slabši, če dovolimo, da algoritem označi največ 50% vseh točk v vrsti. Razlika v rezultatih je v povprečju 2,72%. Večja razlika se pokaže pri minimalni vrednosti, ki jo dosežemo z našim algoritmom v primerjavi s popolnim. Ta znaša okoli 6,71%. V našem primeru smo pri označevanju spiral največji delež označenih točk nastavili na 25% s čimer smo, kot kaže, dobili rezultate, ki so bolj primerljivi z optimalnimi rešitvami.

Pri večini testov, ki so podani v tabeli, smo parameter ω nastavili na 0,5. Razlog za to je, da smo v realnem primeru imeli pri večini atributov, ki smo jih izrisovali, le-tega nastavljenega med 0,3 in 0,6. Za občutek, kolikšen vpliv ima ta parameter na delovanje algoritma, smo naredili še preizkus, kjer je bil parameter nastavljen na 0,1 (*test 3*) in 0,9 (*test 4*). Rezultati se občutno poslabšajo, ko je ω nastavljen na 0,1. Rezultat je od *testa 1*, kjer je parameter ω nastavljen na 0,5, slabši za 5,04%, prav tako pa smo pri *testu 5* dobili tudi najslabši rezultat, ki je bil 76,04%. Pri nastavitvi parametra ω na 0,9 so bili rezultati nekoliko boljši od nastavitve parametra na 0,5.

Sama dolžina vrste pri istih nastavitvah ostalih parametrov ima le manjši vpliv na povprečne vrednosti rezultatov testov. Povprečni doseženi rezultati so bili 98,93% (20 točk), 99,49% (16 točk), 99,33% (12 točk) in 99,80% (8 točk). V najslabšem, primeru

se povprečni rezultat med najdaljšimi in najkrajšmi vrstami razlikuje za 0,9%.

5.4 Diskusija poskusa z eksperti

V nadaljevanju sledi podrobnejša analiza rezultatov, kjer se bomo najprej osredotočili na primere, ki so bili napačno klasificirani v razred *zdrav*, nato pa še na primere, ki so bili napačno klasificirani v razreda PB ali ET.

Prvi ocenjevalec, D. G., je v prvem koraku za tri primere, ki spadajo v razred PB in za tri primere iz razreda ET napovedal, da ne izkazujejo znakov bolezni. Pri podrobnejšem pregledu napak pri klasifikaciji je moč videti, da je za en primer iz razreda PB v drugem koraku (ko so bili vizualizirani pomembnejši odseki spiral) spremenil razred iz *zdrav* v PB. Za ta primer je v komentarju zapisal, da je v spiralah videti popačenja ter da je na eni strani bolj zgoščena, medtem ko je v prvem koraku navedel, da ni videti odstopanj. Pri drugem primeru iz razreda PB si je v drugem koraku ponovno premislil in podal mnenje, da spirale nakazujejo znake bolezni, a se ni mogel odločiti, ali so znaki za ET ali PB. V komentar je zapisal, da so v spiralah vidna popačenja in rahlo valovanje, kar bi lahko nakazovalo tudi na mešan tip tremorja. Za tretji primer iz razreda PB pa v drugem koraku ni spremenil mnenja iz prvega koraka in je primer ponovno klasificiral v razred *zdrav*. Ocenjevalec je za vse tri primere iz razreda ET, ki jih je v prvem koraku razvrstil v razred *zdrav*, v drugem koraku spremenil razred. Dva primera je razvrstil v razred PB, pri obeh pa je v komentar zapisal, da je v spiralah zaznal asimetrijo. Za tretji primer pa se ni mogel odločiti, ali spirale izkazujejo znake bolezni ali ne. V komentarju je navedel, da se ne more odločiti, ali bi primer razvrstil v razred *zdrav* ali PB.

Drugi ocenjevalec, D. N., je v prvem koraku osem primerov, ki spadajo v razred PB in štiri primere, ki spadajo v razred ET klasificiral v razred *zdrav*. Izpostaviti velja, da je pet od šestih primerov, ki jih je napačno klasificiral prvi ocenjevalec (D. G.), napačno v razred *zdrav* razvrstil tudi ocenjevalec D. N., za šestega pa se ni mogel odločiti, ali izkazuje znake bolezni ali ne. Štiri primere iz razreda PB je ocenjevalec D. N. v drugem koraku ponovno klasificiral v razred *zdrav*, v komentarju pa je zapisal, da je bila klasifikacija zahtevna. Za dva od teh primerov je ocenil, da vizualizacija ni pokazala smiselnih odsekov, pri dveh pa naj bi bili označeni smiselni odseki spiral, a mu pri klasifikaciji niso pomagali. Za en primer iz razreda PB se v drugem koraku ni mogel odločiti, ali spirale kažejo znake bolezni ali ne, pri vizualizaciji pa je označil, da niso bili obarvani smiselni predeli spiral. Preostale tri primere je v drugem koraku pravilno klasificiral v razred PB, vizualizacija pa je po njegovi oceni označila smiselne predele spiral. Pri

treh primerih iz razreda ET, ki jih je ocenjevalec v prvem koraku označil za zdrave, je v drugem koraku spremenil razred v PB z opombo, da spirale nakazujejo diskinezijo. Pri dveh primerih je že v prvem koraku navedel, da je morda prisotna diskinezija, pri obeh primerih pa je nato označil, da mu je vizualizacija pomagala pri odločitvi za razred PB. Pri četrtem primeru iz razreda ET, ki ga je v prvem koraku klasificiral v razred *zdrav*, v drugem koraku ni spremenil klasifikacije.

Pri tretjem ocenjevalcu, D. H., je moč opaziti, da je le enkrat klasificiral primer v razred *zdrav* namesto v razred PB, pri čemer v drugem koraku ni spremenil klasifikacije iz prvega koraka.

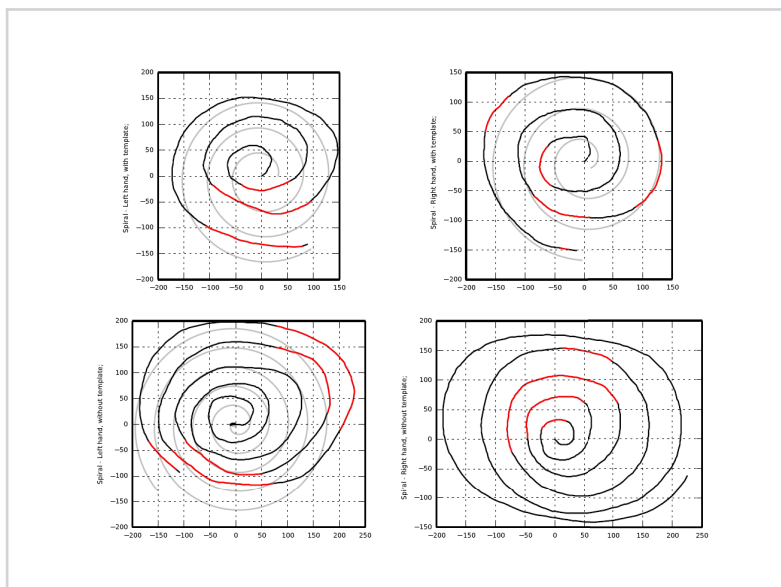
Pri primerjavi klasifikacij ekspertov in njihovih napak smo ugotovili, da so se vsi trije motili pri klasifikaciji istih šestih primerov, ki smo jih podrobneje pogledali. Razred in klasifikacije posameznega primera s strani vseh treh ocenjevalcev so predstavljeni v tabeli 5.17.

Tabela 5.17: Razredi posameznega primera in klasifikacije vseh treh ocenjevalcev v prvem in drugem koraku.

primer	razred	D. G.		D. N.		D. H.	
		1. korak	2. korak	1. korak	2. korak	1. korak	2. korak
33	PB	zdrav	PB	zdrav	zdrav	ET	znaki bolezni
47	PB	zdrav	zdrav	zdrav	ne morem se odločiti	znaki bolezni	PB
141	ET	zdrav	ne morem se odločiti	zdrav	PB	PB	znaki bolezni
156	ET	zdrav	PB	ne morem se odločiti	PB	PB	znaki bolezni
378	ET	zdrav	PB	zdrav	zdrav	ne morem se odločiti	PB
387	PB	zdrav	znaki bolezni	zdrav	zdrav	znaki bolezni	PB

Primer 33, ki je prikazan na sliki 5.14, ni tipičen predstavnik razreda PB. Spirali, ki sta bili narisani po predlogi (zgoraj levo in zgoraj desno), nimata očitnih ostrih kotov ali za PB značilne asimetrije. Na podlagi teh dveh spiral bi težko trdili, da izkazujeta znake kakršne koli bolezni. Spodnja leva spirala ima nekoliko več ostrih kotov, še bolj pomembna pa je stisnjenost spirale. Za oboje je lahko razlog tudi ta, da je bila spirala narisana z ne-vodilno roko. Spodnja desna spirala ponovno ne vsebuje posebnosti. V tem primeru nam je vizualizacija pomagala pri opažanju odstopanj predvsem na levih dveh slikah, kjer je jasno razvidno odstopanje med narisano in idealno spiralo ter zgostitev na eni in razredčenost na drugi strani (asimetrija). Ocenjevalec D. G. je v

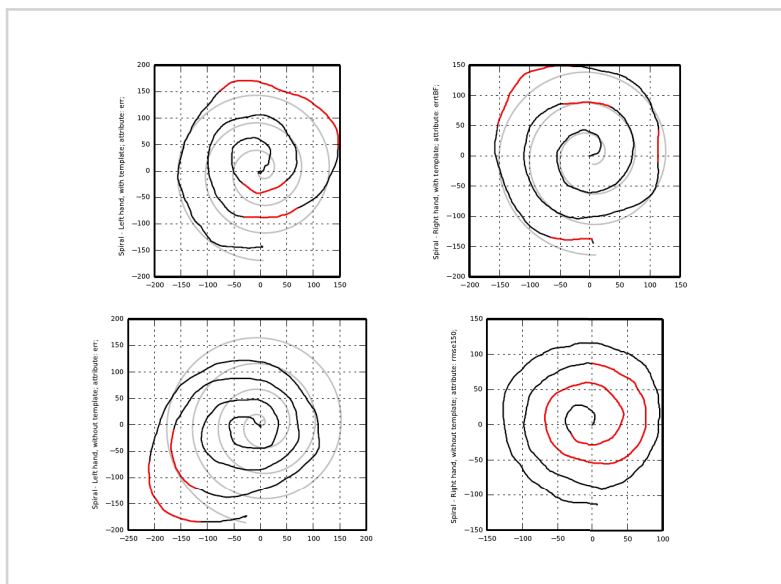
drugem koraku pravilno napovedal razred primera, ocenjevalec D. H. je označil, da so prisotni znaki bolezni (a se ni mogel odločiti med ET in PB), medtem ko je ocenjevalec D. N. vztrajal pri svoji prvotni napovedi, da spirale ne izkazujejo znakov bolezni.



Slika 5.14: Primer 33 z označenimi najpomembnejšimi odseki spiral

Na sliki 5.15 so prikazane spirale za primer 47. Na levih dveh spiralah je sicer vidnih nekaj ostrih kotov, na spodnji spirali tudi asimetričnost, medtem ko desni spirali ne izkazujejo očitnih posebnosti. Na spiralah ni razvidne očitne ritmičnosti (ki je značilna za ET), ostrejši koti na levih spiralah bi lahko nakazovali na diskinezijo (nehoteni gibi) in posledično vidnejše ostre kote. Kljub temu je imel s klasifikacijo tega primera težavo tudi naš odločitveni model, saj ga je namesto v razred PB klasificiral v razred ET. Ocenjevalec D. H. je tako pravilno napovedal, da spirale izkazujejo znake bolezni, v drugem koraku pa je pravilno napovedal tudi razred.

Primer 141 (slika 5.16) je netipičen predstavnik razreda ET. Spirale nimajo značilnih ritmičnih vzorcev, zgornji dve pa imata nekaj ostrejših kotov, ki so bolj značilni za PB oz. diskinezijo. Zgornja leva spirala je bila narisana z ne-vodilno roko, kar bi lahko bil razlog za obarvano popačenost. Težava pri spodnji levi spirali je v tem, da ima

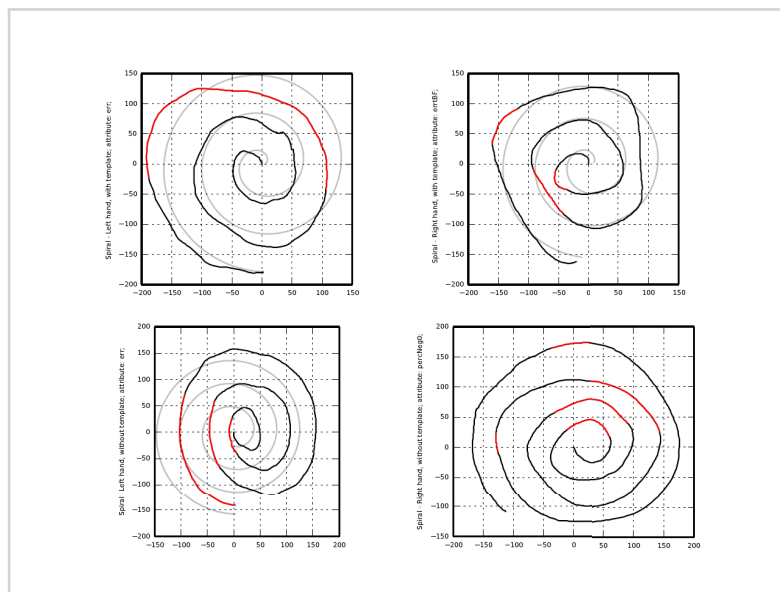


Slika 5.15: Primer 47 z označnimi najpomembnejšimi odseki spiral

zamaknjeno središče, kar privede do navidezne asimetričnosti, ki pa je pravzaprav ni. Do takšne napake pogosto pride pri prostoročnem risanju. Spodnja desna spirala ne vsebuje posebnosti, oglatost pri zadnjem zavoju pa bi lahko pripisali tudi nepazljivosti oz. neskoncentriranosti testiranca. Ekspert bi se tako odločal predvsem na podlagi zgornjih dveh spiral, ki pa nista tipični za razred ET. To potrjujeta tudi oceni dveh ekspertov, ki sta primer klasificirala v razred PB.

Tudi primer 156, prikazan na sliki 5.17, ni tipičen predstavnik razreda ET. Na spiralah ni opaziti ritmičnih vzorcev, vidnih je le nekaj ostrejših kotov na zgornji desni in spodnji levi spirali. Spirale nimajo opazne asimetrije, z vizualizacijo vidimo le manjša odstopanja od idealne spirale. Na podlagi teh spiral ni presenetljivo, da so vsi ocenjevalci primer napačno klasificirali, večinoma v razred PB. Ta primer je napačno klasificiral tudi naš odločitveni model, ki ga je razvrstil v razred PB.

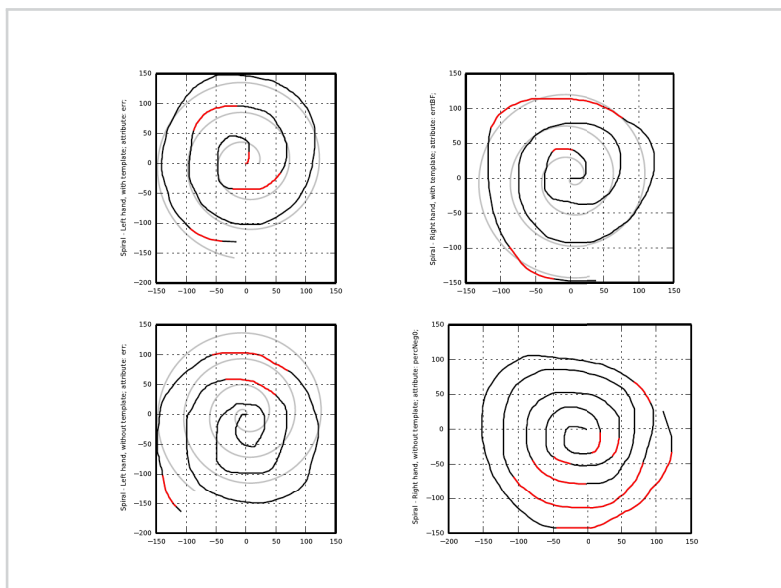
Spirale za primer 378 (ET), ki so prikazane na sliki 5.18, na predlogah ne kažejo posebnosti. Spodnja leva spirala je nekoliko sploščena, vizualizacija pa prikazuje večja odstopanja od idealne spirale. Vzrok za sploščenost je lahko rigidnost ne-vodilne roke,



Slika 5.16: Primer 141 z označenimi najpomembnejšimi odseki spirale

ki je znak za PB. Spodnja desna spirala ima neposrečen začetek spirale (ni v središču oz. težišču spirale), zato je vidna asimetrija med desno in levo stranjo, opaziti pa je tudi rahlo zgoščenost na levi strani. Oboje je znak za PB. Iz teh razlogov so razumljive klasifikacije ekspertov, ki so se odločali med tem, ali primer klasificirati v razred *zdrav* ali PB. Nihče od njih niti v komentarjih ni omenjal, da bi spirale lahko nakazovale tudi na ET. Odločitveni model je prav tako primer razvrstil v razred PB, saj ni našel dovolj značilnosti za razred ET.

Primer 387, čigar spirale so prikazane na sliki 5.19, pripada razredu PB. Zgornja leva spirala sicer je nekoliko oglata, a ji ob dejstvu, da je bila narisana z ne-vodilno roko, ne moremo dati velikega vpliva na končno odločitev. Spodnja leva slika to potrjuje. Razlog za stisnjenost na dnu je lahko v tem, da je bilo področje risanja omejeno z ekranom, kar nakazuje tudi zaključek spirale. Pri spiralah na desni strani ni opaziti večjih anomalij, spodnja slika je morda nekoliko stisnjena, kjer je ponovno lahko razlog v omejenosti risalne površine. Na neoznačenih spiralah sta dva ocenjevalca klasificirala primer v razred *zdrav*, tretji pa je ocenil, da spirale nakazujejo znake bolezni. Pri

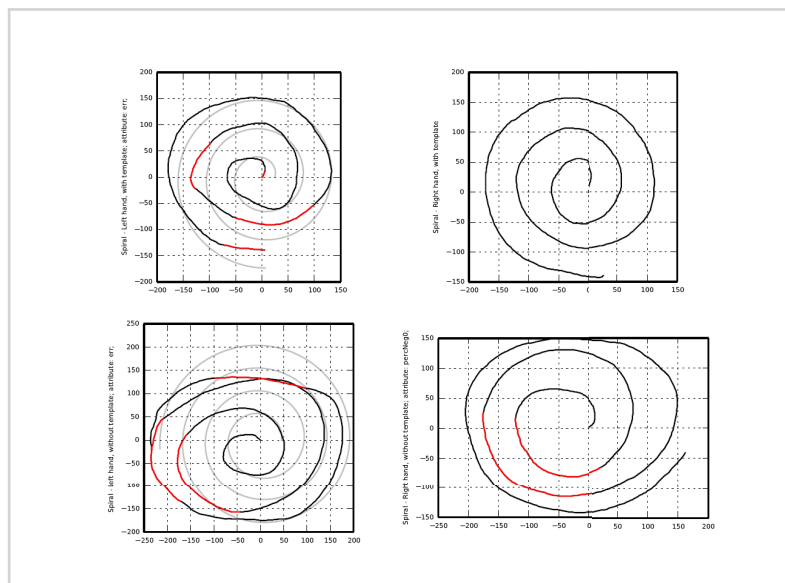


Slika 5.17: Primer 156 z označenimi najpomembnejšimi odseki spiral

označenih spiralah se je slednji odločil, da so znaki tipični za PB, ocenjevalec D. G. je menil, da so vidni znaki bolezni, a se ni mogel odločiti med ET in PB, ocenjevalec D. N. pa je še vedno menil, da spirale ne kažejo na znake bolezni. Da je primer težko pravilno klasificirati, nakazuje tudi klasifikacija našega odločitvenega modela, ki je primer razvrstil v razred ET.

Na podlagi slik spiral, ki smo jih predstavili je očitno, da je teh šest primerov zelo težko pravilno klasificirati le na podlagi spiral. Kljub temu se je izkazalo, da je klasifikacija pomagala pri odločitvi, da spirale vendarle nakazujejo znake bolezni, v nekaterih primerih se je, na podlagi vizualizacije, ekspert pravilno odločil glede samega razreda, kamor sodi primer.

Pri analizi klasifikacij primerov iz razreda *zdrav* se je izkazalo, da so bile ocene ekspertov v prvem koraku točnejše od ocen v drugem koraku. Razlog, da so v drugem koraku za nekaj primerov namesto ocene *zdrav* podali oceno, da izkazujejo znake bolezni, je zagotovo tudi v tem, da smo tudi na spiralah, ki so bile s strani našega modela sicer označene za zdrave, izrisali attribute, ki so na njih prožili. Tega v pravem sistemu

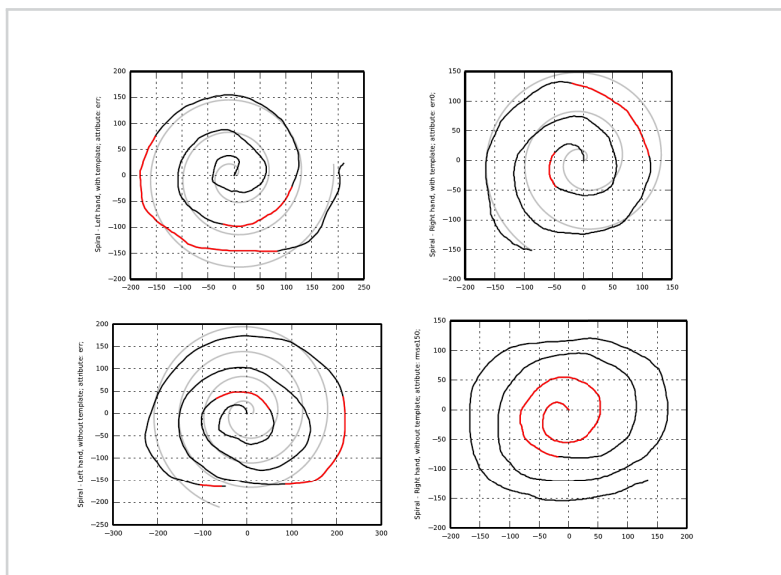


Slika 5.18: Primer 378 z označnimi najpomembnejšimi odseki spirale

za podporo odločanju zelo verjetno ne bi naredili (razen v primeru, ko bi ekspert to izrecno želel).

Ocenjevalec D. G. je v prvem koraku en primer iz razreda *zdrav* kasificiral v razred PB, enako klasifikacijo je podal v drugem koraku. V komentarju je zapisal, da spirale niso videti normalne in da so zgoščene na eni strani. V drugem koraku je na podlagi vizualizacije detektiral asimetrijo in popačenja, zaradi česar je ocenil, da primer spada v razred PB. Za en primer se v prvem koraku ni mogel odločiti, ali izkazuje znake bolezni ali ne, v komentarju pa je navedel, da je opaziti rahel tremor. V drugem koraku je označil, da spirale nakazujejo znake bolezni, pri čemer spirale, narisane po predlogi, nakazujejo ET, prostoročno narisane spirale pa PB.

V drugem koraku je isti ekspert kar šestim primerom iz razreda *zdrav*, ki so bili v prvem koraku pravilno klasificirani, spremenil razred. Za štiri izmed njih ni bil več prepričan, ali so v spiralah vidni znaki bolezni ali ne. Pri vseh štirih je v prvem koraku napisal, da ni opaznih znakov tremorja ali odstopanj od normale. V drugem koraku pa je pri vseh štirih navedel, da sicer meni, da so spirale normalne, ampak da je zaznal določene znake, ki bi lahko pomenili prisotnost ET (v dveh primerih) oz. PB (v drugih



Slika 5.19: Primer 387 z označenimi najpomembnejšimi odseki spirala

dveh primerih). Pri preostalih dveh primerih je, kljub trditvam v prvem koraku, da so spirale lepe in simetrične, v drugem koraku opazil znake, ki nakazujejo na PB, pri eni pa tudi ET. Pri slednji je sicer ocenil, da vizualizacija ni označila smiselnih delov spirala.

Ocenjevalec D. N. je dva primera iz razreda *zdrav* v obeh korakih napačno klasificiral v razred PB. Pri obeh je bila odločitev za to hitrost risanja spirala, ki je bila upočasnjena, kar je tipično za PB. V drugem koraku je štiri primere, ki jih je v prvem koraku pravilno klasificiral v razred *zdrav*, klasificiral v napačen razred. Dva je razvrstil v razred PB, pri dveh pa se ni mogel odločiti, ali so prisotni znaki bolezni ali ne. Zaznal je sicer nekaj anomalij, ki pa so lahko posledica površnega risanja.

Največ napak pri razvrščanju primerov iz razreda *zdrav* je naredil ocenjevalec D. H., ki je v prvem koraku le pet primerov (od 15) klasificiral v pravi razred. Tri primere je v prvem koraku razvrstil v razred PB, od katerih je v drugem za dva označil, da se ne more odločiti, ali vsebujejo znake bolezni, pri enem pa je ohranil razred iz prvega koraka. Za dva primera se v prvem koraku ni mogel odločiti, ali njune spirale vsebujejo znake bolezni, za oba pa je v drugem koraku določil, da spadata v razred PB. Še dva primera je v prvem koraku klasificiral v razred ET, pri čemer je za enega ohranil razred

tudi v drugem koraku, za drugega pa se v drugem koraku ni mogel odločiti, ali so v spiralah prisotni znaki bolezni ali ne. V treh primerih se v prvem koraku ni mogel odločiti, ali spirale nakazujejo na PB ali ET. Za dva primera je nato v drugem koraku določil, da znaki nakazujejo na ET, pri enem pa naj bi bili znaki tipični za PB. Za dva primera, ki jih je v prvem koraku pravilno klasificiral v razred *zdrav*, je v drugem koraku spremenil razred. V enem primeru naj bi spirale izkazovale znake za PB in ET hkrati, v drugem primeru pa se ni mogel odločiti, ali so v spiralah prisotni znaki bolezni ali ne. V komentarju je za slednjega napisal, da tremor ni jasno viden in da gre lahko le za površno risanje ali diskinezijo pri PB.

Pri zadnjem ocenjevalcu se je izkazalo, da že v osnovi podaja bolj pesimistične napovedi (le redko klasificira primer v razred *zdrav*), ob dodatni vizualizaciji problematičnih delov pa se to še dodatno izrazi. Očitno je, da se z vizualizacijo na ocenjevalca zlahka vpliva, saj je kar v 25 primerih (od 45) v drugem koraku spremenil svojo prvotno napoved. Slednje lahko pomeni tudi, da ni prepričan v svoje ocenjevalne sposobnosti in bolj zaupa zunanjim kazalnikom. Ocenjevalec je tudi sicer večkrat jasno izrazil svoje stališče, da meni, da je izključno na podlagi spiral nemogoče postaviti diagnozo ali gre za PB ali ET. Tako je uspešnost modela, ki smo ga predstavili v poglavju 4, presenetila tudi njega.

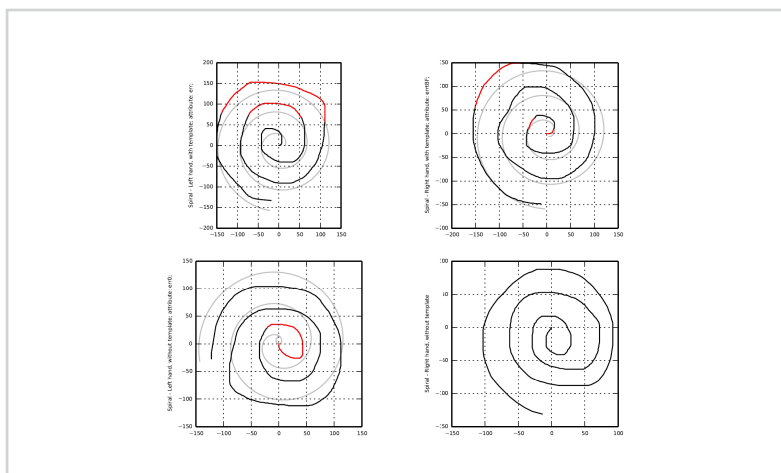
Za razumevanje napak, ki so jih naredili ocenjevalci pri primerih iz razreda *zdrav*, smo pregledali primere, pri katerih sta napako naredila vsaj dva ocenjevalca. Teh primerov je bilo sedem in so predstavljeni v tabeli 5.18.

Tabela 5.18: Klasifikacije vseh treh ocenjevalcev najbolj težavnih primerov iz razreda *zdrav* v prvem in drugem koraku.

primer	razred	D. G.		D. N.		D. H.	
		1. korak	2. korak	1. korak	2. korak	1. korak	2. korak
165	zdrav	zdrav	ne morem se odločiti	zdrav	ne morem se odločiti	PB	ne morem se odločiti
173	zdrav	zdrav	ne morem se odločiti	zdrav	ne morem se odločiti	ne morem se odločiti	PB
236	zdrav	ne morem se odločiti	PB in ET	PB	PB	PB	PB
246	zdrav	zdrav	zdrav	zdrav	PB	ne morem se odločiti	PB
276	zdrav	zdrav	ne morem se odločiti	PB	PB	znaki bolezni	PB
401	zdrav	zdrav	PB	zdrav	PB	zdrav	PB in ET
436	zdrav	PB	PB	zdrav	zdrav	znaki bolezni	ET

Pri pregledu slik smo ugotovili, da so spirale primerov 165 (slika 5.20), 173 (slika 5.21), 236 (slika 5.22), 401 (slika 5.25) in 436 (slika 5.26) povečini okorne, vsebujejo ostre kote, opazimo lahko zgostitve na eni strani, nekatere spirale pa imajo tudi nazobčan vzrec. Rečemo lahko, da ni nič nenavadnega, da so jih ocenjevalci klasificirali v napačen razred.

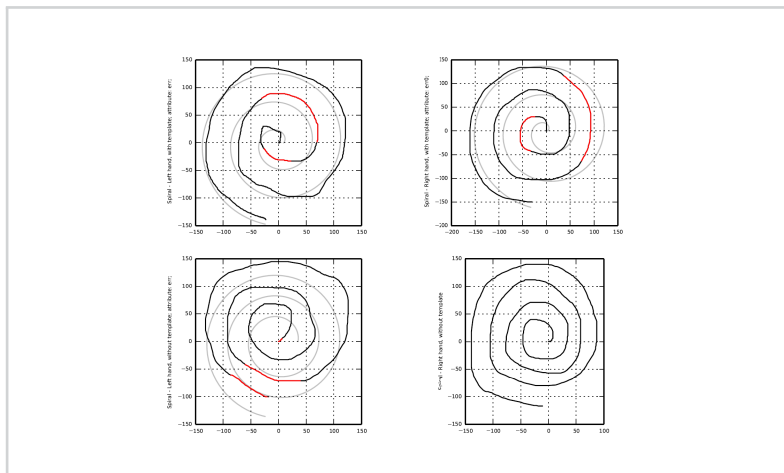
Nekoliko pa presenečajo napačne klasifikacije primerov 246 (slika 5.23) in 276 (slika 5.24), saj so spirale povečini gladke, simetrične in ne kažejo očitnih vzorcev za PB ali ET. Naj omenimo še, da je vse primere naš odločitveni model pravilno klasificiral v razred *zdrav*, verjetnost za prisotnost znakov bolezni pa je pri šestih primerih manjša od 6%, pri enem primeru (primer 436) pa je ta verjetnost 13%.



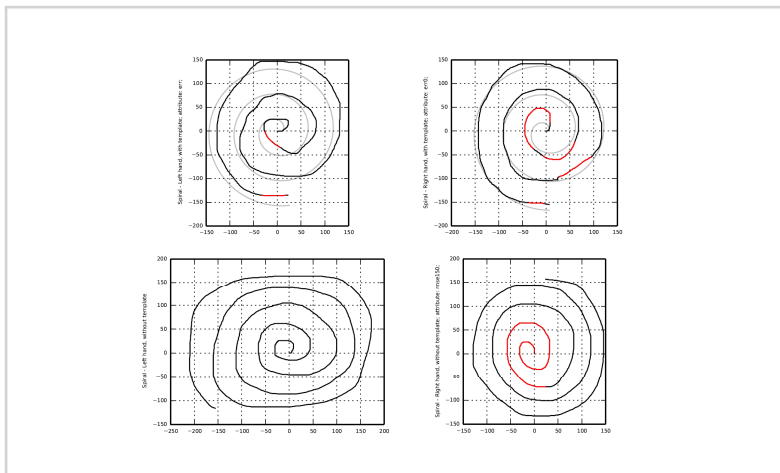
Slika 5.20: Primer 165 z označenimi najpomembnejšimi odseki spiral

Na podlagi dobljenih rezultatov lahko trdimo, da ima vizualizacija močan vpliv na odločitve ekspertov pri klasificiranju primerov v razrede. Kljub temu, da so bili eksperti seznanjeni z dejstvom, da so v prvem in drugem koraku predstavljeni isti primeri, se je kar nekajkrat zgodilo, da so v drugem koraku spremenili svojo odločitev. To nam dokazuje, da so eksperti zaupali vizualizaciji in našemu odločitvenemu modelu, ki je bil podlaga za vizualizacijo. Morda celo preveč, saj je jasno, da je pri primerih, ki so sicer zdravi, vizualizacija ocenjevalce zavedla in jih pripeljala do napačne odločitve. Če sistemu zaupaš in vidiš, da je označil nekatere dele na spiralah, lahko podvomíš v svojo odločitev do te mere, da podaš napačno mnenje. S tem je postalo jasno, da na

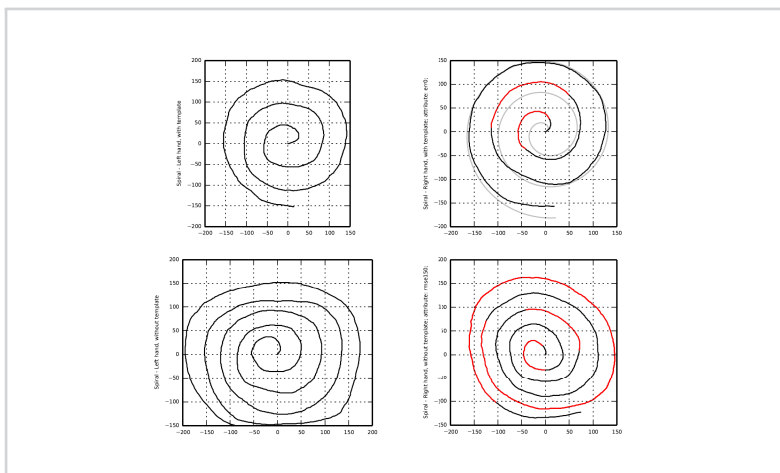
Slika 5.21: Primer 173 z označnimi najpomembnejšimi odseki spiral



spiralah primerov, ki jih sistem klasificira kot zdrave, ne bi smeli označevati atributov, ki so prožili, saj s tem zavedemo uporabnika sistema. V dejanskem sistemu za podporo odločanju bi bilo tako smotrno označevati attribute le na tistih spiralah, pri katerih je bil primer klasificiran v razred PB ali ET.

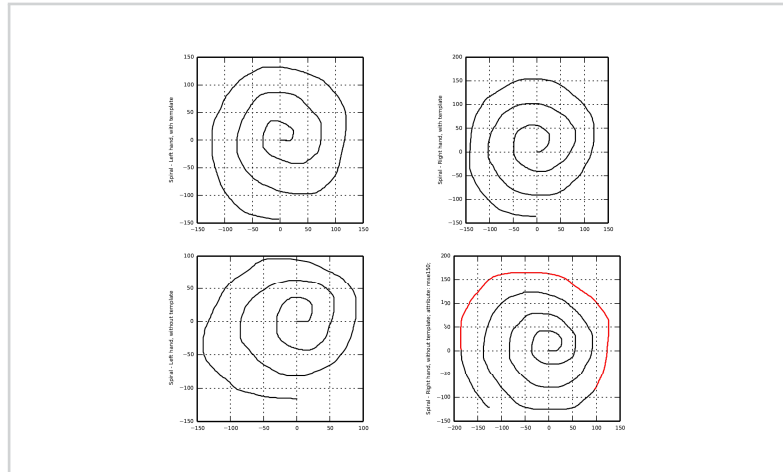


Slika 5.22: Primer 236 z označnimi najpomembnejšimi odseki spiral

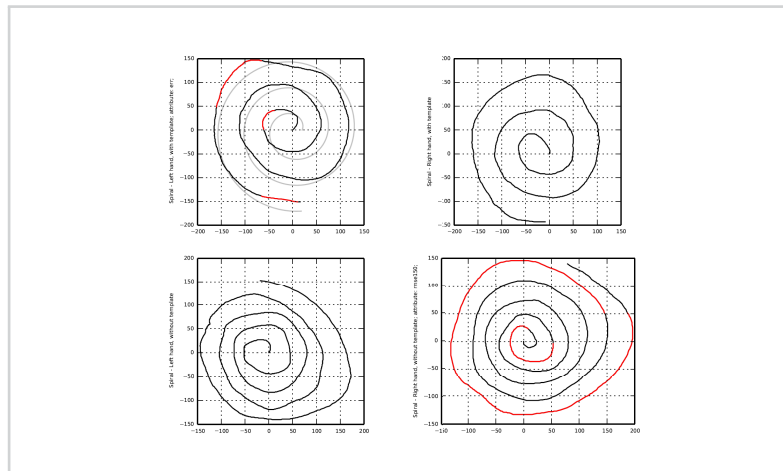


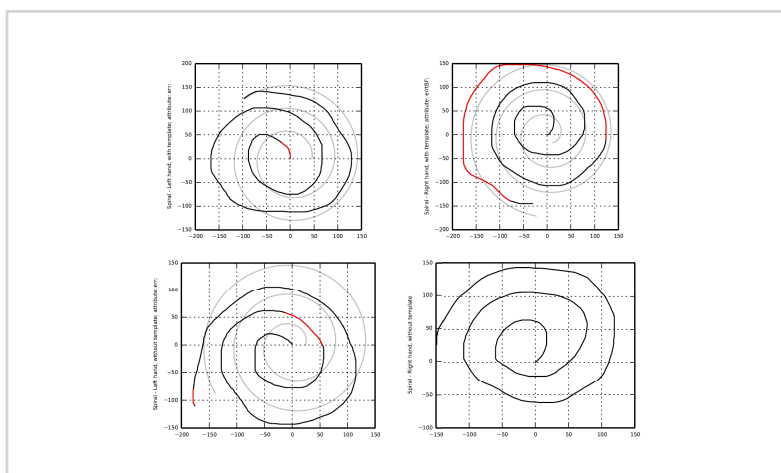
Slika 5.23: Primer 246 z označnimi najpomembnejšimi odseki spiral

Slika 5.24: Pri-
mer 276 z označ-
nimi najpomemb-
nejšimi odseki spi-
ral



Slika 5.25: Pri-
mer 401 z označ-
nimi najpomemb-
nejšimi odseki spi-
ral





Slika 5.26: Primer 436 z označnimi najpomembnejšimi odseki spiral





Zaključki

6

6.1 Zaključki

V disertaciji smo se ukvarjali z izgradnjo diagnostičnih modelov in metode za vizualno razlago diagnoz za različne tipe tresavic (tremorjev). Osredotočili smo se na tri najpogostejše tipe tremorjev - esencialni tremor, parkinsonski tremor in mešani tip tremorja. Na začetku smo se lotili gradnje modela za razločevanje tremorjev na podlagi kliničnih in spirografskih podatkov. Za izgradnjo modela smo uporabili argumentirano strojno učenje (ABML), ki na podlagi argumentov, ki jih tekom izvajanja postopka podaja ekspert, zgradi ekspertu razumljiv odločitveni model. Ker je bila naša naloga razločevati med tremi različnimi tremorji, algoritem pa v splošnem deluje na dvorazrednih problemih, smo naš problem prevedli na dva dvorazredna problema in primerno prilagodili algoritem, da je na koncu podal klasifikacijo primera v enega izmed treh razredov. Gradnja in testiranje modela sta potekala na podatkih 122 preiskovancev (z uporabo metode prečnega preverjanja), s čimer smo dobili model z visoko klasifikacijsko točnostjo (0.91). Prednost pristopa, ki smo ga uporabili, je v boljši klasifikacijski točnosti modela (klasifikacijska točnost je narasla za skoraj deset odstotkov od točnosti modela, ki smo ga dobili zgolj z uporabo algoritma CN2), razumljivosti končnega modela in medicinski smiselnosti elicitanega znanja skozi cikel algoritma ABML. Prepričani smo, da je razumljivost ključnega pomena za oceno primernosti končne množice pravil za uporabo v sistemih za podporo odločanju. Le-ta je pomembna za razlago odločitve, saj pripomore k večjemu zaupanju zdravnika v predlagane diagnoze oz. pomaga zdravnikom odkriti potencialne napake v diagnozi.

Kljub začetnim pomislekom ekspertov, da bi lahko klasifikacija preiskovancev temeljila le na podatkih spirografije, se je izkazalo, da je v končnem modelu kar devet pravil (od trinajstih) vsebovalo attribute, dobljene iz spirografije. To nas je vzpodbudilo k temu, da smo se lotili problema diagnostike izključno na podatkih spirografije. V ta namen smo naredili namensko aplikacijo za mobilne telefone s pomočjo katere smo zbrali podatke 143 preiskovancev. Med njimi so bili takšni, ki so imeli parkinsonski ali esencialni tremor in zdrave kontrole. Pred samo izgradnjo modela smo jih, po posvetovanju z eksperti, 19 izločili, saj so prejeli zdravilo (levodopa) manj kot tri ure pred testiranjem, kar je lahko vplivalo na sam rezultat spirografije (zmanjšanje tremorja). Za razliko od prejšnjega problema, ko so bili podatki o spirografiji podani v kvalitativni obliki, smo tokrat obravnavali kvantitativne podatke (časovne vrste). Za izgradnjo odločitvenega modela smo podatke skušali predstaviti v obliki atributov, ki

smo jih zgradili ročno, predvsem z uporabo predznanja, ki smo ga pridobili od ekspertov v prejšnjem odločitvenem problemu (npr. asimetričnost spirale, nazobčanost spirale, spremembe v hitrosti risanja ipd.). Ker smo tudi tokrat imeli trirazreden problem, smo se klasifikacije lotili v dveh korakih. V prvem koraku je model določil, ali je tremor prisoten ali ne. V primeru, da je bila zaznana prisotnost tremorja, je v drugem koraku sledilo razločevanje med esencialnim in parkinsonskim tremorjem. Attribute smo uporabili za gradnjo modela z različnimi metodami, na koncu pa smo zaradi potrebe po razumljivosti modela izbrali model, ki je bil dobljen z logistično regresijo. V splošnem bi bil pristop za raševanje klasifikacijskih problemov na časovnih vrstah uporaba metode k-najbližjih sosedov, ki uporablja metodo dinamičnega ukrivljanja časa (kNN-DTW), a so naši poskusi pokazali, da je ta metoda na naših podatkih delovala slabše od metode za izgradnjo modela z uporabo zgrajenih atributov. Sklepamo lahko, da predznanje, ki je zakodirano v attribute doprinese več informacije učnim metodam kot same časovne vrste. Odločitveni model smo implementirali v mobilno aplikacijo ParkinsonCheck, ki je prosto dostopna in deluje na različnih mobilnih platformah.

Modeli, dobljeni z uporabo logistične regresije, so v splošnem razumljivi, seveda ob predpogoju, da uporabnik razume attribute, ki so zastopani v modelu. Ker so naši potencialni uporabniki zdravniki, ki modelov niso gradili, je za njih ključnega pomena, da jim podamo razlago odločitve. V ta namen smo razvili algoritem, ki poišče tiste točke v vrstah, ki so pomembno prispevale k vrednosti posameznega atributa v odločitvenem modelu. Pri tem smo se držali tudi pravila, da si ne želimo najti točk, ki so popolnoma razpršene v vrsti, ampak da preferiramo povezane točke (torej daljše intervale). Smiselnost naše vizualizacije smo preverili z eksperimentom, v katerem smo trem različnim nevrologom prikazali spirale, na katerih na začetku nismo označili pomembnih odsekov. Eksperti so se morali odločiti ali je oseba, ki je spirale narisala, zdrava ali ima prisoten tremor. V drugem koraku smo na spiralah označili odseke, ki so pomembno prispevali h klasifikaciji preiskovanca. Na podlagi rezultatov ekspertov lahko trdimo, da ima vizualizacija močan vpliv na odločitve ekspertov pri klasifikaciji primerov v razrede, saj so le-ti v drugem koraku velikokrat spremenili svojo prvotno klasifikacijo primera. V dejanskem sistemu za podporo odločanju bi tako označili attribute le na spiralah, ki so bile klasificirane v razred esencialni ali parkinsonski tremor.

Naš algoritem za vizualizacijo je v splošnem požrešen algoritem, saj je zaradi velikosti problema algoritem, ki bi poiskal optimalno rešitev, neuporaben oz. ga ni moč uporabiti. Tako smo uspešnost delovanja našega algoritma primerjali z rezultati, ki

bi jih dobili z uporabo optimalnega algoritma na manjšem številu točk. Rezultati so pokazali, da je delovanje požrešnega algoritma primerljivo z delovanjem optimalnega algoritma. Rezultati, ki smo jih na manjši množici podatkov dobili z enim ali drugim algoritmom, se v povprečju razlikujejo za okoli pet odstotkov. Za naše potrebe pa takšna razlika nima bistvenega pomena.

6.2 Prispevki k znanosti

Ključna prispevka k znanosti, ki sta predstavljena v disertaciji, sta:

1. računalniški diagnostični model za ločevanje med esencialnim, parkinsonskim in mešanim tremorjem;
2. metoda za vizualno razlago diagnoz.



*Začetni atributi, uporabljeni
za izgradnjo modela za
razločevanje tremorjev na
podlagi kliničnih podatkov in
digitalne spirografije*

A

Tabela A.1: Začetni atributi, ki smo jih uporabili za gradnjo modela s podanim opisom in deležem manjkajočih podatkov.

<i>Atribut</i>	<i>Kratek opis</i>	<i>Delež manjkajočih vrednosti (%)</i>
Age	Starost preiskovanca	5,26
Alcohol.response	Odzivanje ob zaužitju alkohola - ET se zmanjša ali popolnoma izzveni v 60% primerov, za PT to ne velja	77,19
Bare.left.freq.harmonics	Frekvence harmonikov v spektralni analizi ob risanju z levo roko brez predloge	7,89
Bare.left.freq.maxamp	Maksimalna amplituda frekvence tremorja ob risanju z levo roko brez predloge	28,95
Bare.left.freq.range	Frekvenčno območje tremorja ob risanju z levo roko brez predloge	21,05
Bare.left.radius.angle	Transformacija radij-kot ob risanju z levo roko brez predloge	17,54
Bare.left.speed.time	Transformacija hitrost-čas ob risanju z levo roko brez predloge	16,67
Bare.right.freq.harmonics	Frekvence harmonikov v spektralni analizi ob risanju z desno roko brez predloge	9,65
Bare.right.freq.maxamp	Maksimalna amplituda frekvence tremorja ob risanju z desno roko brez predloge	33,33
Bare.right.freq.range	Frekvenčno območje tremorja ob risanju z desno roko brez predloge	14,04
Bare.right.radius.angle	Transformacija radij-kot ob risanju z desno roko brez predloge	9,65
Bare.right.speed.time	Transformacija hitrost-čas ob risanju z desno roko brez predloge	7,02
Nadaljevanje na naslednji strani –		

– nadaljevanje s prejšnje strani.		
<i>Atribut</i>	<i>Kratek opis</i>	<i>Delež manjkajočih vrednosti (%)</i>
Bradykinesia.left	Upočasnjeno gibanje leve strani	4,39
Bradykinesia.right	Upočasnjeno gibanje desne strani	4,39
Diagnosis	Postavljena diagnoza	0,00
Disease.duration	Trajanje bolezni	46,49
Education	Dosežena stopnja izobrazbe	9,65
Gait	Klinična nevrološka ocena hoje; pri osebah z ET je hoja normalna, pri osebah s PT in nekaterimi drugimi nevrološkimi motnjami pa ima določene specifične	70,18
History	Družinska anamneza - ali ima še kdo od sorodnikov podobno motnjo	25,44
Hypokinesia.left	Zmanjšana amplituda giba rok na levi strani	44,74
Hypokinesia.right	Zmanjšana amplituda giba rok na desni strani	44,74
Kinetic.tremor.up.left	Tremor zgornje leve okončine ob gibanju	10,53
Kinetic.tremor.up.right	Tremor zgornje desne okončine ob gibanju	10,53
Postural.tremor.up.left	Tremor leve roke ob držanju iztegnjenih rok pred seboj	8,77
Postural.tremor.up.right	Tremor desne roke ob držanju iztegnjenih rok pred seboj	8,77
Qualitative.spiral	Kvalitativna ocena spiral	8,77
Resting.tremor.up.left	Tremor leve roke v mirovanju	5,26
Resting.tremor.up.right	Tremor desne roke v mirovanju	5,26
Rigidity.low.left	Povečan mišični tonus (rigidnost) spodnje leve okončine	80,70
Nadaljevanje na naslednji strani –		

– nadaljevanje s prejšnje strani.		
<i>Atribut</i>	<i>Kratek opis</i>	<i>Delež manjkajočih vrednosti (%)</i>
Rigidity.low.right	Povečan mišični tonus (rigidnost) spodnje desne okončine	80,70
Rigidity.neck	Povečan mišični tonus (rigidnost) vratu	85,09
Rigidity.up.left	Povečan mišični tonus (rigidnost) leve roke	6,14
Rigidity.up.right	Povečan mišični tonus (rigidnost) desne roke	6,14
Sex	Spol preiskovanca	0,00
Template.left.freq.harmonics	Frekvence harmonikov v spektralni analizi ob risanju z levo roko po predlogi	7,89
Template.left.freq.maxamp	Maksimalna amplituda frekvence tremorja ob risanju z levo roko po predlogi	48,25
Template.left.freq.range	Frekvenčno območje tremorja ob risanju z levo roko po predlogi	42,98
Template.left.radius.angle	Transformacija radij-kot ob risanju z levo roko po predlogi	38,60
Template.left.speed.time	Transformacija hitrost-čas ob risanju z levo roko po predlogi	37,72
Template.right.freq.harmonics	Frekvence harmonikov v spektralni analizi ob risanju z desno roko po predlogi	8,77
Template.right.freq.maxamp	Maksimalna amplituda frekvence tremorja ob risanju z desno roko po predlogi	46,49
Template.right.freq.range	Frekvenčno območje tremorja ob risanju z desno roko po predlogi	46,49
Template.right.radius.angle	Transformacija radij-kot ob risanju z desno roko po predlogi	40,35
Nadaljevanje na naslednji strani –		

– nadaljevanje s prejšnje strani.

<i>Atribut</i>	<i>Kratek opis</i>	<i>Delež manjkajočih vrednosti (%)</i>
Template.right.speed.time	Transformacija hitrost-čas ob risanju z desno roko po predlogi	47,37
Tremor.duration	Trajanje tremorja	12,28
Tremor.neck	Tremor vratu	8,77
Tremor.start	Na kateri strani telesa se je pričel tremor?	38,60





*Primer sledi izvajanja
algoritma*

B

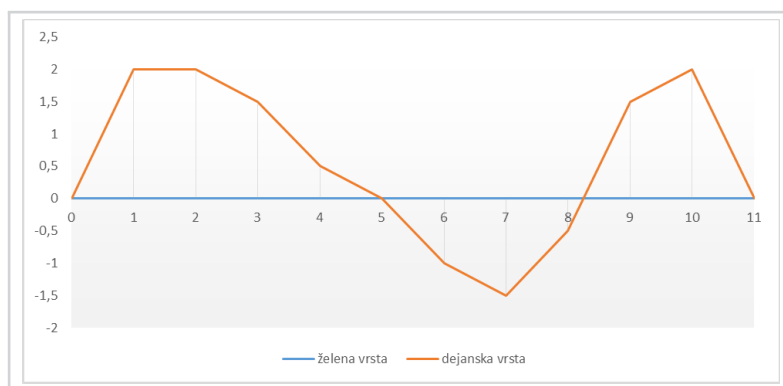
Podano imamo vrsto, ki vsebuje 12 točk. Atribut, ki ga bomo uporabili za prikaz delovanja algoritma, je RMSE (Root mean squared error) in je za dano vrsto enak 1,299. Vrednosti vrste in nevtralne vrste so podane spodaj. Vrednosti točk so prikazane na sliki B.1.

$$Vrsta = [0.0, 2.0, 2.0, 1.5, 0.5, 0.0, -1.0, -1.5, -0.5, 1.5, 2.0, 0.0]$$

$$NeutralnaVrsta = [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]$$

$$RMSE = 1,299$$

Slika B.1: Slika prikazuje vrednosti točk v vrsti (oranžna lomljena črta povezuje posamezne točke) in vrednosti točk nevtralne vrste (ki so enake 0 in so povezane z modro črto).



Izračun prispevkov intervalov

Algoritem najprej izračuna prispevke posamezne točke, kot je podano v opisu algoritma. Vrednost točke iz seznama $Vrsta$ nadomesti z vrednostjo istoležne točke v seznamu $NeutralnaVrsta$. Tako spremenjeno vrsto uporabi za nov izračun vrednosti atributa $RMSE_{nov}$. Prispevek posamezne točke je razlika med prvotno vrednostjo atributa RMSE in vrednostjo atributa $RMSE_{nov}$.

Za vsako točko so v nadaljevanju podane vrednosti spremenjenega seznama $Vrsta$, na novo izračunana vrednost $RMSE$ in *prispevek* vsake točke. V primeru, ko iščemo le eno točko, je prispevek enak koristnosti točke. Spremenjene vrednosti v seznamu $Vrsta$ so tudi grafično prikazane na slikah.

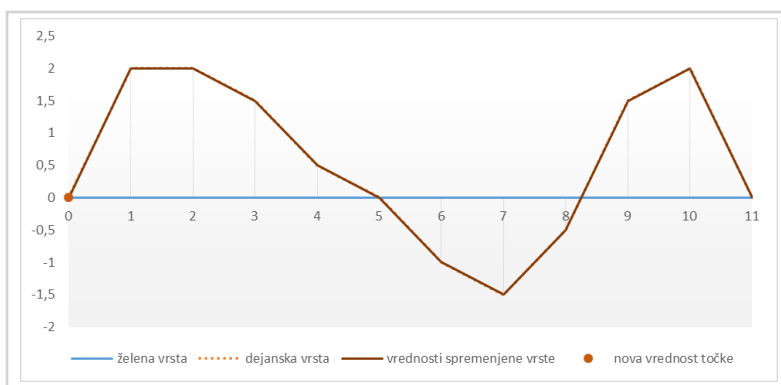
Ko algoritem izbere prvo točko, ki ima največjo koristnost, je ta točka ob naslednjem obhodu že podana v rešitvi, ki ji dodajamo nove točke. Vrednosti vseh točk v seznamu *Vrsta*, ki so v trenutno računani rešitvi, nadomesti z vrednostmi istoležnih točk v seznamu *NeutralnaVrsta* in tako spremenjeno vrsto uporabi za izračun vrednosti atributa $RMSE_{nov}$ (vrednost atributa se torej izračuna na novo vsakokrat, ko dodamo v potencialno rešitev novo točko). Prispevek rešitve je razlika med vrednostjo atributa $RMSE$ in $RMSE_{nov}$.

Interval 0 (I0):

$$Vrsta_{I0} = [0.0, 2.0, 2.0, 1.5, 0.5, 0.0, -1.0, -1.5, -0.5, 1.5, 2.0, 0.0]$$

$$RMSE_{I0} = 1,299$$

$$Prispevek_{I0} = 0,0$$



Slika B.2: Slika prikazuje vrednosti spremenjene vrste (v tem primeru je vrednost točke ostala na 0).

Interval 1 (I1):

$$Vrsta_{I1} = [0.0, 0.0, 2.0, 1.5, 0.5, 0.0, -1.0, -1.5, -0.5, 1.5, 2.0, 0.0]$$

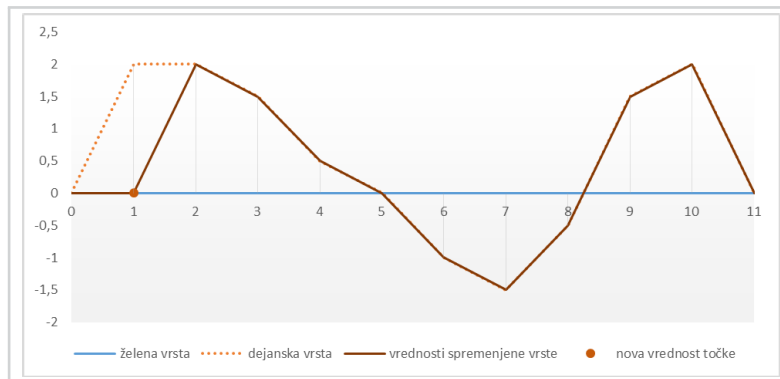
$$RMSE_{I1} = 1,163687$$

$$Prispevek_{I1} = 0,135351$$

Interval 2 (I2):

$$Vrsta_{I2} = [0.0, 2.0, 0.0, 1.5, 0.5, 0.0, -1.0, -1.5, -0.5, 1.5, 2.0, 0.0]$$

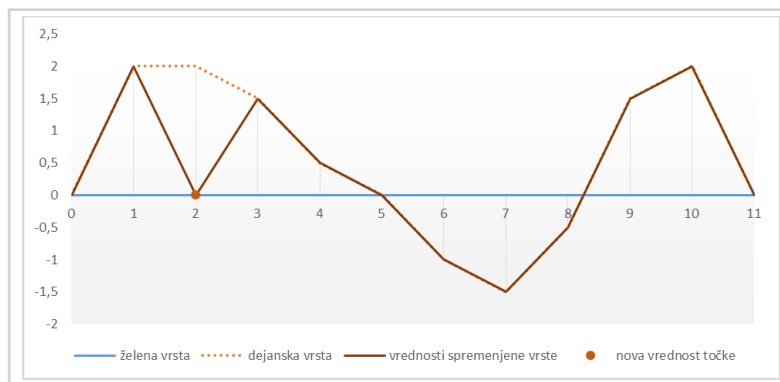
Slika B.3: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 1.



$$RMSE_{12} = 1,163687$$

$$Prispevek_{12} = 0,135351$$

Slika B.4: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 2.



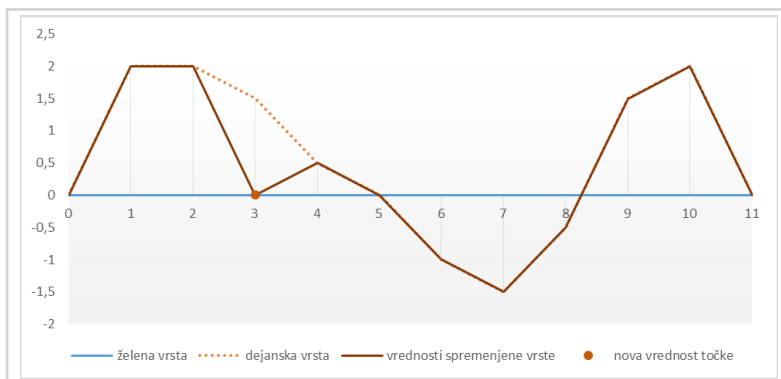
Interval 3 (I3):

$$Vrsta_{13} = [0.0, 2.0, 2.0, 0.0, 0.5, 0.0, -1.0, -1.5, -0.5, 1.5, 2.0, 0.0]$$

$$RMSE_{13} = 1,224745$$

$$Prispevek_{13} = 0,074293$$

Interval 4 (I4):

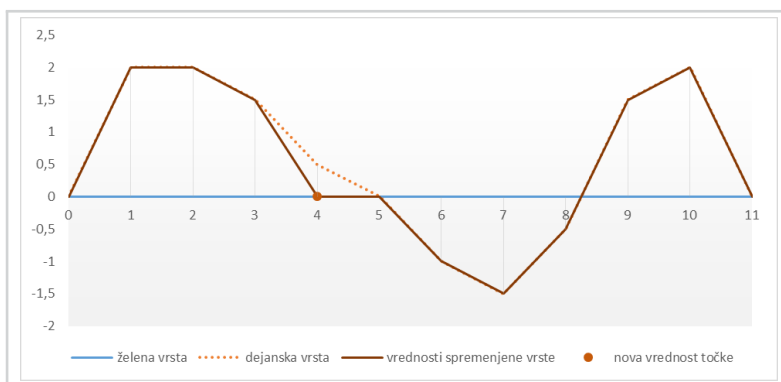


Slika B.5: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 3.

$$Vrsta_{14} = [0.0, 2.0, 2.0, 1.5, 0.0, 0.0, -1.0, -1.5, -0.5, 1.5, 2.0, 0.0]$$

$$RMSE_{14} = 1,290994$$

$$Prispevek_{14} = 0,008044$$



Slika B.6: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 4.

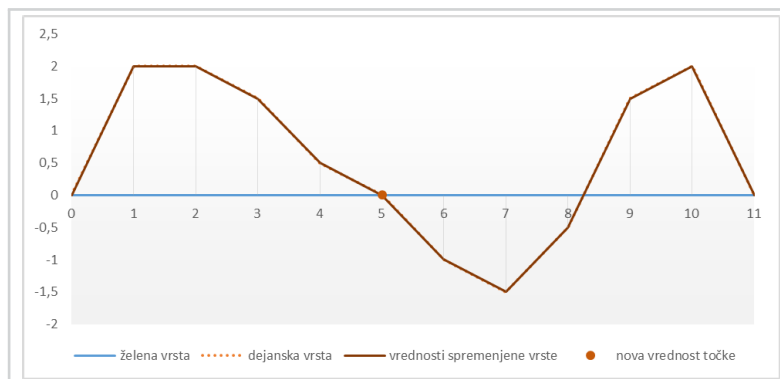
Interval 5 (I5):

$$Vrsta_{15} = [0.0, 2.0, 2.0, 1.5, 0.5, 0.0, -1.0, -1.5, -0.5, 1.5, 2.0, 0.0]$$

$$RMSE_{15} = 1,299038$$

$$Prispevek_{15} = 0,0$$

Slika B.7: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 5.



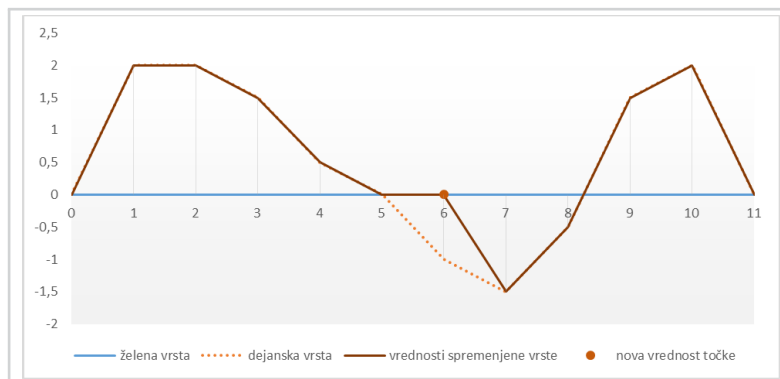
Interval 6 (I6):

$$\mathbf{Vrsta}_{I6} = [0.0, 2.0, 2.0, 1.5, 0.5, 0.0, 0.0, -1.5, -0.5, 1.5, 2.0, 0.0]$$

$$RMSE_{I6} = 1,266557$$

$$Prispevek_{I6} = 0,032481$$

Slika B.8: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 6.

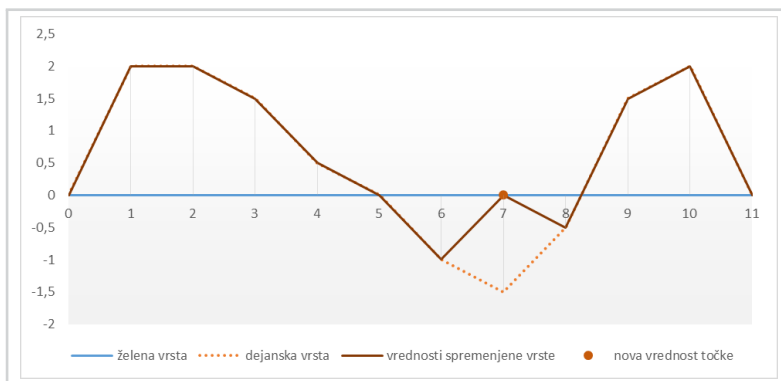


Interval 7 (I7):

$$\mathbf{Vrsta}_{I7} = [0.0, 2.0, 2.0, 1.5, 0.5, 0.0, -1.0, 0.0, -0.5, 1.5, 2.0, 0.0]$$

$$RMSE_{I7} = 1,224745$$

$$Prispevek_{17} = 0,074293$$



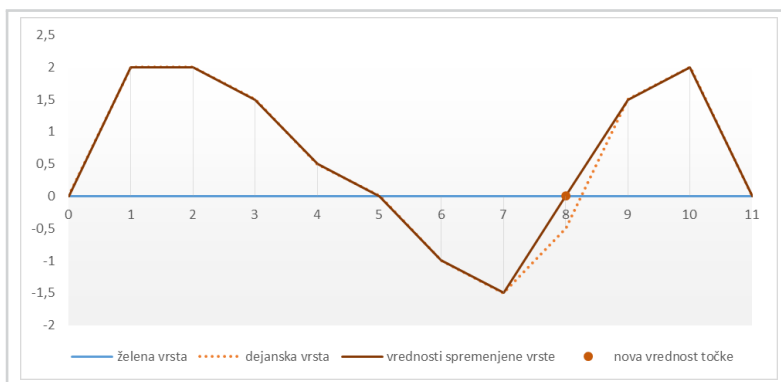
Slika B.9: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 7.

Interval 8 (I8):

$$Vrsta_{I8} = [0,0, 2,0, 2,0, 1,5, 0,5, 0,0, -1,0, -1,5, 0,0, 1,5, 2,0, 0,0]$$

$$RMSE_{I8} = 1,290994$$

$$Prispevek_{I8} = 0,008044$$



Slika B.10: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 8.

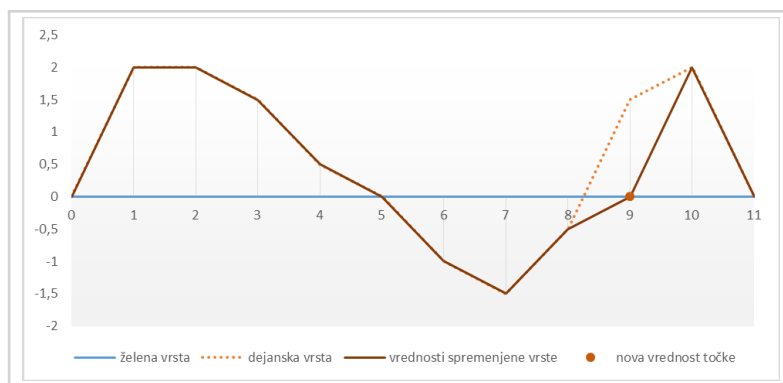
Interval 9 (I9):

$$Vrsta_{I9} = [0,0, 2,0, 2,0, 1,5, 0,5, 0,0, -1,0, -1,5, -0,5, 0,0, 2,0, 0,0]$$

$$RMSE_{I_9} = 1,224745$$

$$Prispevek_{I_9} = 0,074293$$

Slika B.11: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 9.



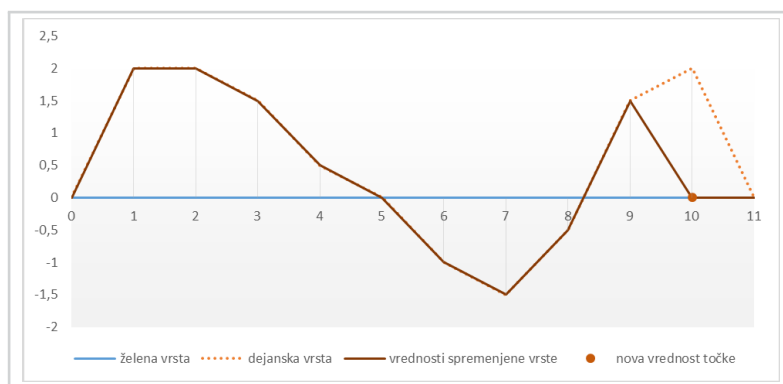
Interval 10 (I10):

$$Vrsta_{I_{10}} = [0.0, 2.0, 2.0, 1.5, 0.5, 0.0, -1.0, -1.5, -0.5, 1.5, 0.0, 0.0]$$

$$RMSE_{I_{10}} = 1,163687$$

$$Prispevek_{I_{10}} = 0,135351$$

Slika B.12: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 10.

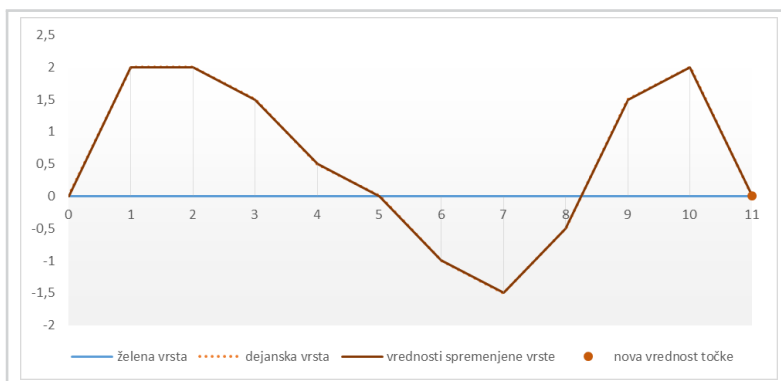


Interval 11 (I11):

$$Vrsta_{111} = [0,0, 2,0, 2,0, 1,5, 0,5, 0,0, -1,0, -1,5, -0,5, 1,5, 2,0, 0,0]$$

$$RMSE_{111} = 1,299038$$

$$Prispevek_{111} = 0,0$$



Slika B.13: Slika prikazuje vrednosti spremenjene vrste ob spremembi vrednosti intervala 11.

Uporaba funkcije *poiščiIntervale()*

Ustavitveni kriterij: z obhodom zanke prenehaj, ko je označenih več kot 50% točk

Rešitev = []

$\omega = 0,5$

najboljšaKoristnost = -inf

trenutnaRešitev = []

1. OBHOD ZANKE:

kandidatZaRešitev = [10]

prispevek = 0,0

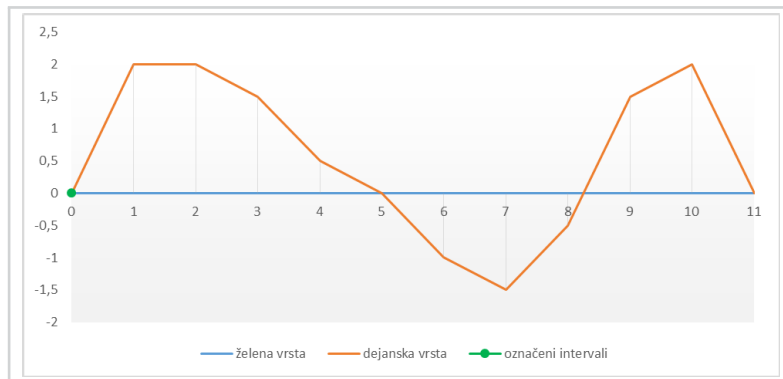
cena = 1,0

koristnost = 0,0

najboljšaKoristnost = 0,0

trenutnaRešitev = [10]

kandidatZaRešitev = [11]



Slika B.14: Slika prikazuje trenutno rešitev - označitev točke 10.

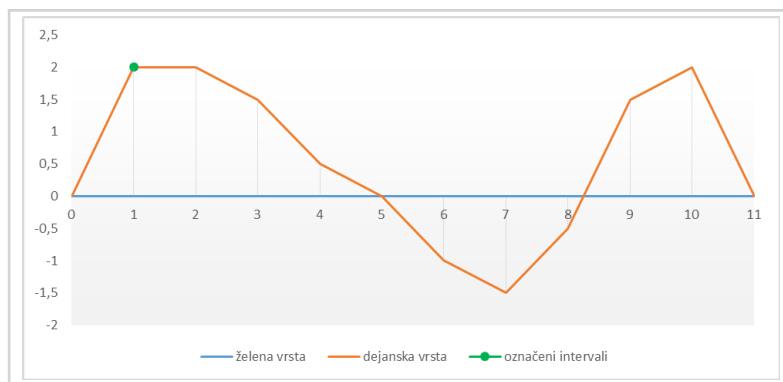
prispevek = 0,135351435

cena = 1,0

koristnost = 0,135351435

najboljšaKoristnost = 0,135351435

trenutnaRešitev = [11]



Slika B.15: Slika prikazuje trenutno rešitev - označitev točke 11.

kandidatZaRešitev = [12]

prispevek = 0,135351435

cena = 1,0

koristnost = 0,135351435
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [I1]

kandidatZaRešitev = [I3]
prispevek = 0,074293234
cena = 1,0
koristnost = 0,074293234
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [I1]

kandidatZaRešitev = [I4]
prispevek = 0,008043657
cena = 1,0
koristnost = 0,008043657
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [I1]

kandidatZaRešitev = [I5]
prispevek = 0,0
cena = 1,0
koristnost = 0,0
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [I1]

kandidatZaRešitev = [I6]
prispevek = 0,032481093
cena = 1,0
koristnost = 0,032481093
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [I1]

kandidatZaRešitev = [I7]
prispevek = 0,074293234



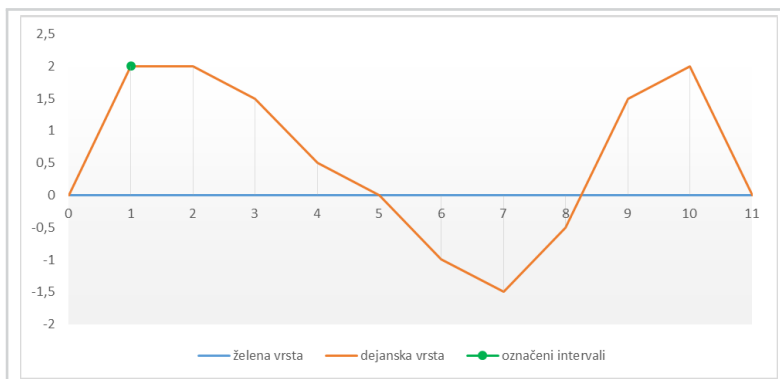
cena = 1,0
koristnost = 0,074293234
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [11]

kandidatZaRešitev = [18]
prispevek = 0,008043657
cena = 1,0
koristnost = 0,008043657
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [11]

kandidatZaRešitev = [19]
prispevek = 0,074293234
cena = 1,0
koristnost = 0,074293234
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [11]

kandidatZaRešitev = [110]
prispevek = 0,135351435
cena = 1,0
koristnost = 0,135351435
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [11]

kandidatZaRešitev = [111]
prispevek = 0,0
cena = 1,0
koristnost = 0,0
najboljšaKoristnost = 0,135351435
trenutnaRešitev = [11]



Slika B.16: Označena rešitev po prvem obhodu zanke.

Po prvem obhodu zanke bi označili interval I1, kot je prikazano na sliki B.16. Ker še nismo presegli ustavitvenega kriterija, nadaljujemo z novim obhodom zanke, pri čemer imamo v trenutni rešitvi že eno točko. V naslednjem obhodu zanke bomo torej obstoječi točki v rešitvi dodajali nove in preverjali, če se s katero izmed dodanih točk koristnost rešitve poveča.

2. OBHOD ZANKE

kandidatZaRešitev = [I1–I0] /* Intervala sta sosedna, zato ju združimo */

prispevek = 0,135351435

cena = 1,79248125

koristnost = 0,075510656

najboljšaKoristnost = 0,135351435

trenutnaRešitev = [I1]

kandidatZaRešitev = [I1–I2] /* Intervala sta sosedna, zato ju združimo */

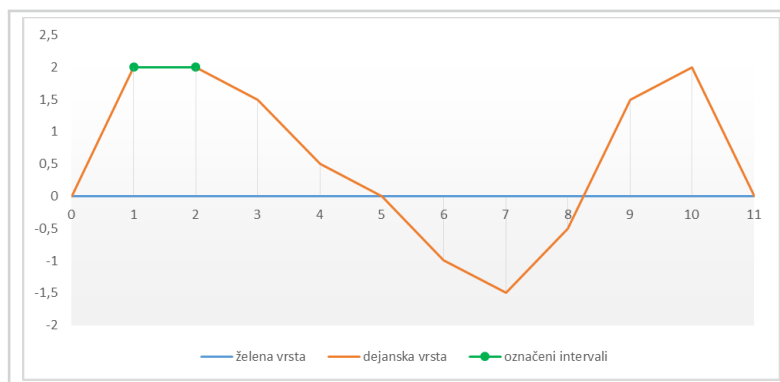
prispevek = 0,288675135

cena = 1,79248125

koristnost = 0,161047785

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1–I2]



Slika B.17: Trenutna rešitev - označitev intervala I1–I2.

kandidatZaRešitev = [I1, I3]
 prispevek = 0,218914656
 cena = 2.0
 koristnost = 0,109457328
 najboljšaKoristnost = 0,161047785
 trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1, I4]
 prispevek = 0,144337567
 cena = 2.0
 koristnost = 0,072168784
 najboljšaKoristnost = 0,161047785
 trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1, I5]
 prispevek = 0,135351435
 cena = 2.0
 koristnost = 0,067675718
 najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1, I6]

prispevek = 0,171725667

cena = 2.0

koristnost = 0,085862834

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1, I7]

prispevek = 0,218914656

cena = 2.0

koristnost = 0,109457328

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1, I8]

prispevek = 0,144337567

cena = 2.0

koristnost = 0,072168784

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1, I9]

prispevek = 0,218914656

cena = 2.0

koristnost = 0,109457328

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1, I10]

prispevek = 0,288675135

cena = 2.0

koristnost = 0,144337567



najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1, I11]

prispevek = 0,135351435

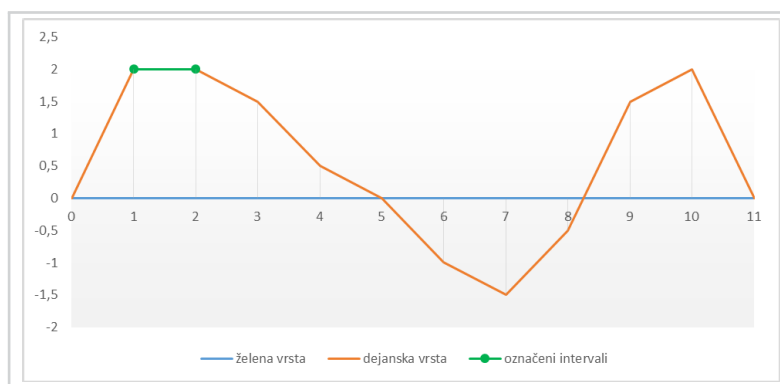
cena = 2.0

koristnost = 0,067675718

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

Po drugem obhodu zanke smo označili točki I1 in I2, ki smo ju združili v skupen interval I1—I2. Rešitev je prikazana na sliki B.18.



Slika B.18: Označena rešitev po drugem obhodu zanke.

3. OBHOD ZANKE

kandidatZaRešitev = [I0—I1—I2] /* I0, ki ga dodajamo rešitvi, je sosedni intervalu v rešitvi, zato ju združimo v en skupen interval. */

prispevek = 0,288675135

cena = 2,5

koristnost = 0,115470054

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1—I2—I3] /* I3, ki ga dodajamo rešitvi, je sosedni intervalu v rešitvi, zato ju združimo v en skupen interval. */

prispevek = 0,386167177

cena = 2,5

koristnost = 0,154466871

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1—I2, I4]

prispevek = 0,299038106

cena = 2,79248125

koristnost = 0,10708688

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1—I2, I5]

prispevek = 0,288675135

cena = 2,79248125

koristnost = 0,103375854

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1—I2, I6]

prispevek = 0,330792269

cena = 2,79248125

koristnost = 0,118458188

najboljšaKoristnost = 0,161047785

trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1—I2, I7]

prispevek = 0,386167177

cena = 2,79248125

koristnost = 0,13828819
najboljšaKoristnost = 0,161047785
trenutnaRešitev = [I1—I2]

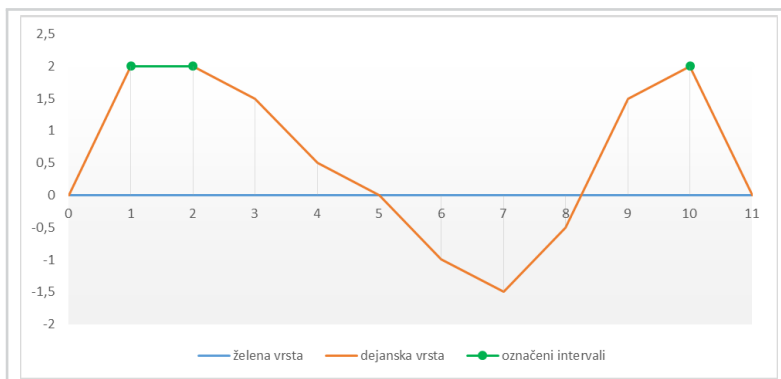
kandidatZaRešitev = [I1—I2, I8]
prispevek = 0,299038106
cena = 2,79248125
koristnost = 0,10708688
najboljšaKoristnost = 0,161047785
trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1—I2, I9]
prispevek = 0,386167177
cena = 2,79248125
koristnost = 0,13828819
najboljšaKoristnost = 0,161047785
trenutnaRešitev = [I1—I2]

kandidatZaRešitev = [I1—I2, I10]
prispevek = 0,469881908
cena = 2,79248125
koristnost = 0,168266809
najboljšaKoristnost = 0,168266809
trenutnaRešitev = [I1—I2, I10]

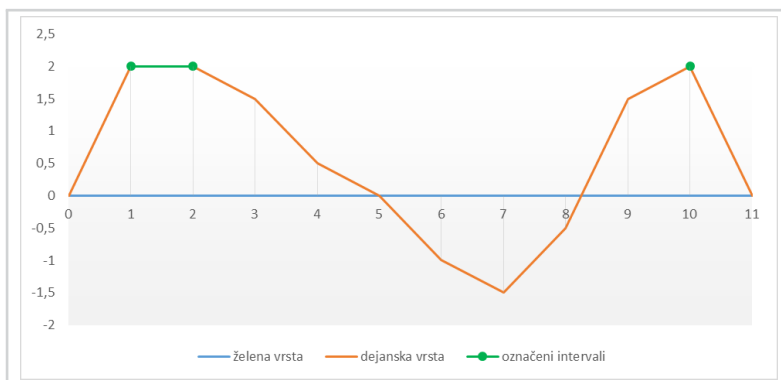
kandidatZaRešitev = [I1—I2, I11]
prispevek = 0,288675135
cena = 2,79248125
koristnost = 0,103375854
najboljšaKoristnost = 0,168266809
trenutnaRešitev = [I1—I2, I10]

Po tretjem obhodu zanke smo označili interval I1—I2 ter točko I10. Rešitev je prikazana na sliki B.20. Ustavitvenega kriterija še nismo presegli, zato nadaljujemo z



Slika B.19: Trenutna rešitev - označitev intervala I1–I2 in točke I10.

novim obhodom zanke.



Slika B.20: Označena rešitev po tretjem obhodu zanke.

4. OBHOD ZANKE

kandidatZaRešitev = [I0–I1–I2, I10] /* Točka I0 je sosedna intervalu I1–I2, zato ju združimo */

prispevek = 0,469881908

cena = 3,5

koristnost = 0,134251974

najboljšaKoristnost = 0,168266809

trenutnaRešitev = [I1—I2, I10]

kandidatZaRešitev = [I1—I2—I3, I10] /* Točka I3 je sosedna intervalu I1—I2, zato ju združimo */

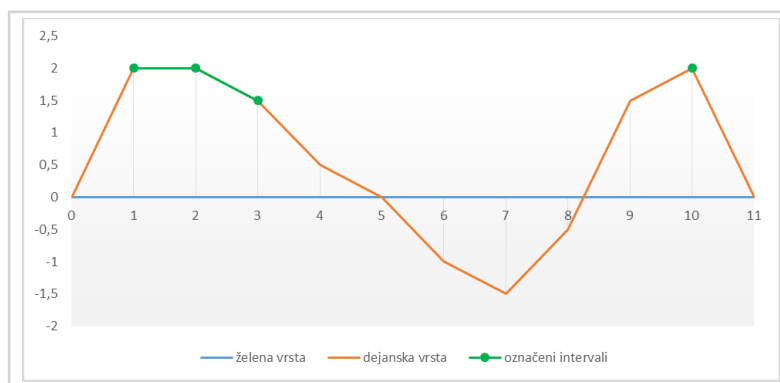
prispevek = 0,591931324

cena = 3,5

koristnost = 0,169123236

najboljšaKoristnost = 0,169123236

trenutnaRešitev = [I1—I2—I3, I10]



Slika B.21: Trenutna rešitev - označitev intervala I1—I2—I3 in točke I10.

kandidatZaRešitev = [I1—I2, I10, I4]

prispevek = 0,482541525

cena = 3,79248125

koristnost = 0,127236364

najboljšaKoristnost = 0,169123236

trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2, I10, I5]

prispevek = 0,469881908

cena = 3,79248125

koristnost = 0,123898281

najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2, I10, I6]
prispevek = 0,521756518
cena = 3,79248125
koristnost = 0,137576558
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2, I10, I7]
prispevek = 0,591931324
cena = 3,79248125
koristnost = 0,156080224
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

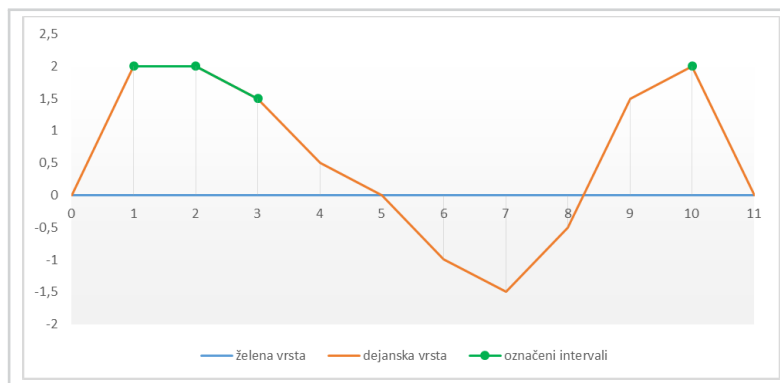
kandidatZaRešitev = [I1—I2, I10, I8]
prispevek = 0,482541525
cena = 3,79248125
koristnost = 0,127236364
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2, I9—I10]
prispevek = 0,591931324
cena = 3,584962501
koristnost = 0,165115067
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2, I10—I11]
prispevek = 0,469881908
cena = 3,584962501

koristnost = 0,131070244
 najboljšaKoristnost = 0,169123236
 trenutnaRešitev = [I1—I2—I3, I10]

Po četrtem obhodu zanke smo označili interval I1—I2—I3 ter točko I10. Rešitev je prikazana na sliki B.22. Ustavitvenega kriterija še nismo presegli, zato nadaljujemo z novim obhodom zanke.



Slika B.22: Označena rešitev po četrtem obhodu zanke.

5. OBHOD ZANKE

kandidatZaRešitev = [I0—I1—I2—I3, I10] /* Točka I0 je sosedna intervalu I1—I2—I3, zato ju združimo v skupen interval. */

prispevek = 0,591931324
 cena = 4,160964047
 koristnost = 0,142258217
 najboljšaKoristnost = 0,169123236
 trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2—I3—I4, I10] /* Točka I4 je sosedna intervalu I1—I2—I3, zato ju združimo v skupen interval. */

prispevek = 0,60681945
 cena = 4,160964047

koristnost = 0,145836264
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2—I3, I10, I5]
 prispevek = 0,591931324
 cena = 4,5
 koristnost = 0,131540294
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

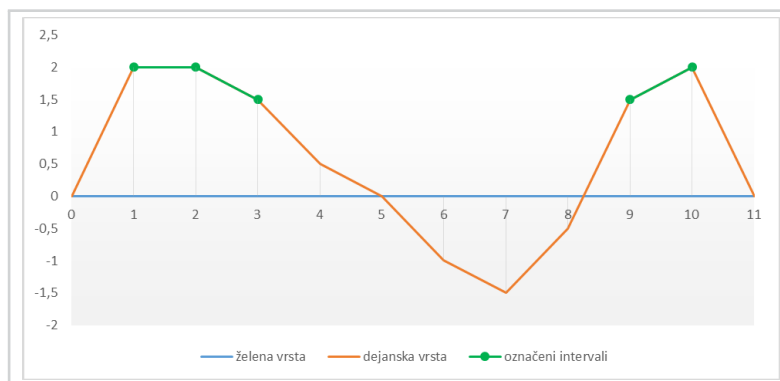
kandidatZaRešitev = [I1—I2—I3, I10, I6]
 prispevek = 0,653540881
 cena = 4,5
 koristnost = 0,145231307
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2—I3, I10, I7]
 prispevek = 0,740021111
 cena = 4,5
 koristnost = 0,164449136
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2—I3, I10, I8]
 prispevek = 0,60681945
 cena = 4,5
 koristnost = 0,134848767
najboljšaKoristnost = 0,169123236
trenutnaRešitev = [I1—I2—I3, I10]

kandidatZaRešitev = [I1—I2—I3, I9–I10] /* Točka I9 je sosedna točki I10, zato ju združimo. */

prispevek = 0,740021111
 cena = 4,29248125
 koristnost = 0,172399381
 najboljšaKoristnost = 0,172399381
 trenutnaRešitev = [I1—I2—I3, I9—I10]



Slika B.23: Trenutna rešitev - označitev intervalov I1—I2—I3 in I9—I10.

kandidatZaRešitev = [I1—I2—I3, I10—I11] /* Točka I11 je sosedna točki I10, zato ju združimo. */

prispevek = 0,591931324
 cena = 4,29248125
 koristnost = 0,137899571
 najboljšaKoristnost = 0,172399381
 trenutnaRešitev = [I1—I2—I3, I9—I10]

Po petem obhodu zanke smo označili intervala I1—I2—I3 ter I9-I10. Rešitev je prikazana na sliki B.23. Ustavitvenega kriterija še nismo presegli, zato nadaljujemo z novim obhodom zanke.

6. OBHOD ZANKE

kandidatZaRešitev = [I0—I1—I2—I3, I9—I10] /* Intervali so sosedni, zato jih združimo */

prispevek = 0,740021111
 cena = 4,953445298
 koristnost = 0,149395232
najboljšaKoristnost = 0,172399381
trenutnaRešitev = [I1—I2—I3, I9–I10]

kandidatZaRešitev = [I1—I2—I3—I4, I9–I10] /* Intervali so sosedni, zato jih združimo */

prispevek = 0,758976381
 cena = 4,953445298
 koristnost = 0,153221916
najboljšaKoristnost = 0,172399381
trenutnaRešitev = [I1—I2—I3, I9–I10]

kandidatZaRešitev = [I1—I2—I3, I9–I10, I5]

prispevek = 0,740021111
 cena = 5,29248125
 koristnost = 0,139824985
najboljšaKoristnost = 0,172399381
trenutnaRešitev = [I1—I2—I3, I9–I10]

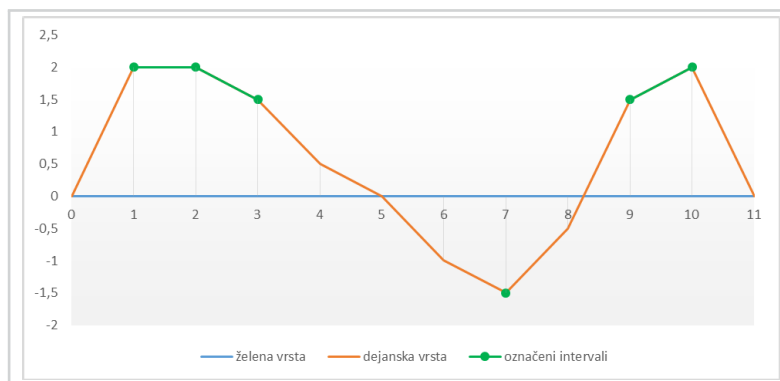
kandidatZaRešitev = [I1—I2—I3, I9–I10, I6]

prispevek = 0,820324552
 cena = 5,29248125
 koristnost = 0,154998103
najboljšaKoristnost = 0,172399381
trenutnaRešitev = [I1—I2—I3, I9–I10]

kandidatZaRešitev = [I1—I2—I3, I9–I10, I7]

prispevek = 0,945484715
 cena = 5,29248125
 koristnost = 0,178646777
najboljšaKoristnost = 0,178646777

trenutnaRešitev = [I1—I2—I3, I7, I9—I10]



Slika B.24: Trenutna rešitev - označitev intervalov I1—I2—I3 in I9—I10 ter točke I7.

kandidatZaRešitev = [I1—I2—I3, I8—I9—I10] /* Intervali so sosedni, zato jih združimo */

prispevek = 0,758976381

cena = 5,0

koristnost = 0,151795276

najboljšaKoristnost = 0,178646777

trenutnaRešitev = [I1—I2—I3, I7, I9—I10]

kandidatZaRešitev = [I1—I2—I3, I9—I10—I11] /* Intervali so sosedni, zato jih združimo */

prispevek = 0,740021111

cena = 5,0

koristnost = 0,148004222

najboljšaKoristnost = 0,178646777

trenutnaRešitev = [I1—I2—I3, I7, I9—I10]

Po šestem obhodu zanke smo označili intervala I1—I2—I3 ter I9—I10 in dodatno še točko I7. Označitev je prikazana na sliki B.24. Tudi po tem obhodu ustavitveni kriterij še ni bil presežen, zato nadaljujemo z novim obhodom zanke.

7. OBHOD ZANKE

kandidatZaRešitev = [I0—I1—I2—I3, I7, I9—I10] /* Intervali so sosedni, zato jih združimo */

prispevek = 0,945484715

cena = 5,953445298

koristnost = 0,158813035

najboljšaKoristnost = 0,178646777

trenutnaRešitev = [I1—I2—I3, I7, I9—I10]

kandidatZaRešitev = [I1—I2—I3—I4, I7, I9—I10] /* Intervali so sosedni, zato jih združimo. */

prispevek = 0,976289493

cena = 5,953445298

koristnost = 0,163987312

najboljšaKoristnost = 0,178646777

trenutnaRešitev = [I1—I2—I3, I7, I9—I10]

kandidatZaRešitev = [I1—I2—I3, I7, I9—I10, I5]

prispevek = 0,945484715

cena = 6,29248125

koristnost = 0,150256263

najboljšaKoristnost = 0,178646777

trenutnaRešitev = [I1—I2—I3, I7, I9—I10]

kandidatZaRešitev = [I1—I2—I3, I6—I7, I9—I10] /* Intervali so sosedni, zato jih združimo. */

prispevek = 1,09491396

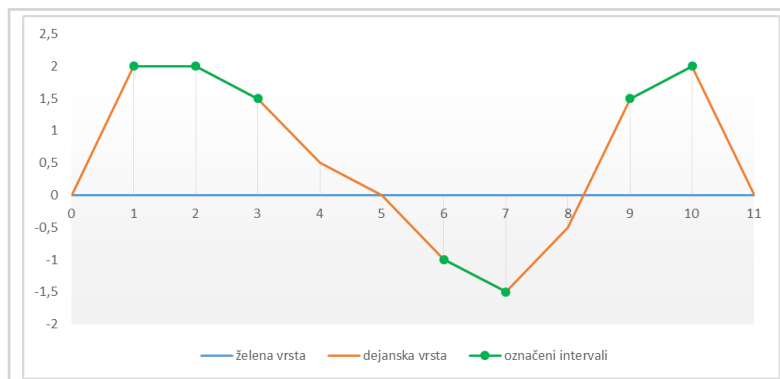
cena = 6,084962501

koristnost = 0,179937668

najboljšaKoristnost = 0,179937668

trenutnaRešitev = [I1—I2—I3, I6—I7, I9—I10]

kandidatZaRešitev = [I1—I2—I3, I7—I8—I9—I10]



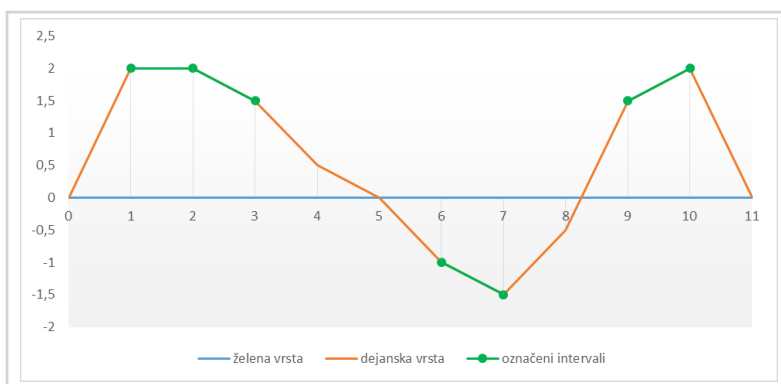
Slika B.25: Trenutna rešitev - označitev intervalov I1–I2–I3, I6–I7 in I9–I10.

prispevek = 0,976289493
 cena = 5,660964047
 koristnost = 0,172459935
 najboljšaKoristnost = 0,179937668
 trenutnaRešitev = [I1—I2—I3, I6–I7, I9–I10]

kandidatZaRešitev = [I1—I2—I3, I7, I9–I10—I11]
 prispevek = 0,945484715
 cena = 6,0
 koristnost = 0,157580786
 najboljšaKoristnost = 0,05194353
 trenutnaRešitev = [I1—I2—I3, I6–I7, I9–I10]

Po sedmem obhodu zanke imamo označenih sedem točk, ki so združene v tri intervale: I1—I2—I3, I6—I7 ter I9—I10. Ker smo s tem obhodom presegli vrednost ustavitvenega kriterija (obarvali smo več kot 50% točk), predstavlja rešitev tega obhoda tudi končno rešitev za označitev točk na grafu.

REŠITEV = [I1—I2—I3, I6–I7, I9–I10]



Slika B.26: Končna rešitev označitev točk - označimo tri intervale: I1–I2–I3, I6–I7 in I9–I10.



*Navodila za izvedbo
eksperimenta*

C

V nadaljevanju so navodila, ki so bila posredovana ekspertom s področja spirografije, v katerih smo podrobneje predstavili način izvedbe eksperimenta in spletno aplikacijo, ki so jo uporabljali za ocenjevanje spiral. Navodila so v angleškem jeziku, saj so bili vsi eksperti/ocenjevalci tujci.

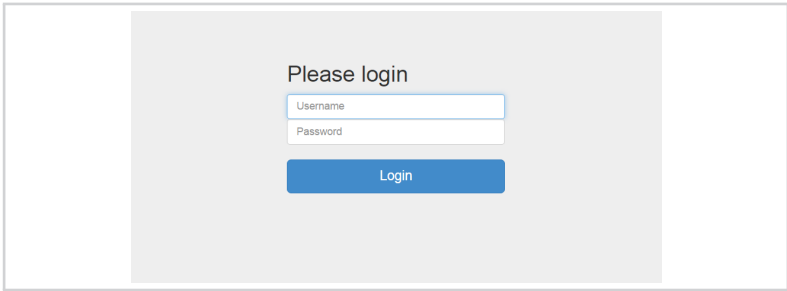
Instructions for the experiment

We have developed the ParkinsonCheck application which has a built-in decision model that can differentiate between spirals drawn by people with essential tremor, Parkinsonian tremor and healthy subjects. The goal of our project is to visually present the most meaningful parts of the spirals based on which the computer gave its prediction. We are basically showing what the computer “found” in these spirals. We would like to check whether the visualisation of the computer’s decision could be helpful for the neurologists when faced with the task to diagnose a person based on the spirals only.

We have set up a web application with a survey to determine the usefulness of our visualisations. We would be pleased if you could spare some of your time and participate in the survey. The whole survey should take at most two hours and it does not have to be done in one run, however, it is advised to.

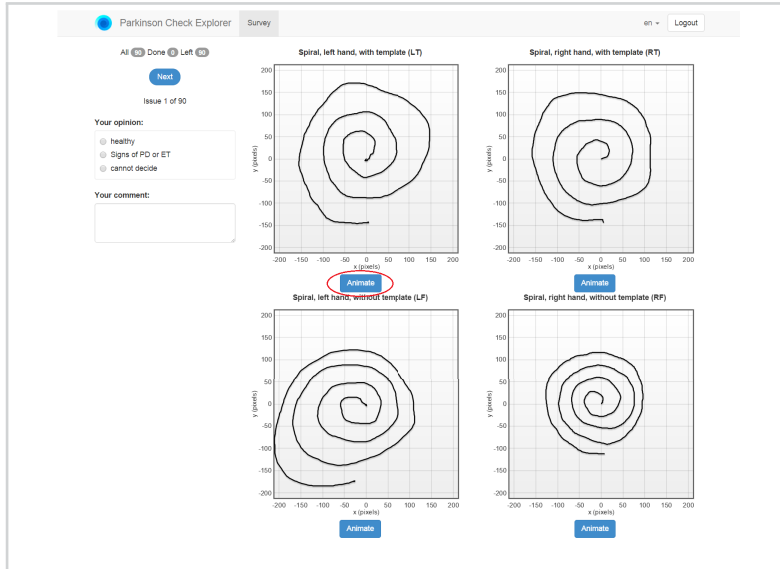
The survey is available at: <https://www.parkinsoncheck.net/vidadoctor/>

When accessing the application, you should see the following login screen (Figure C.1) where you enter your user credentials.



Slika C.1: Login screen for the survey.

After the successful login you will automatically enter the survey. In the first step, you are asked to rate 45 test subjects based on four spirals shown on the screen (Figure C.2).

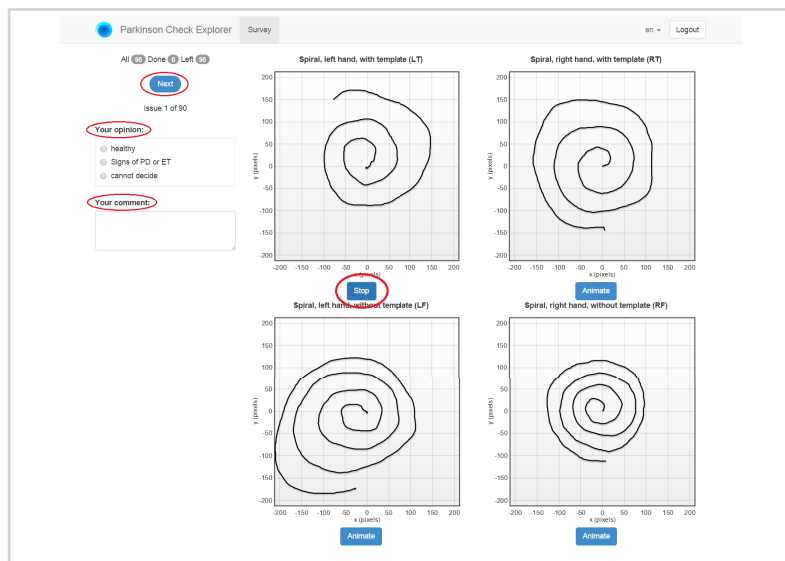


Slika C.2: Login screen for the survey.

The top left spiral was drawn with the non-dominant hand (denoted as left hand in the survey) on the template (i.e. a person had to follow the line of an already drawn spiral – template), the top right spiral was drawn with the dominant hand (denoted as right hand in the survey) on the template, the bottom left spiral was drawn with the non-dominant hand without template (i.e. a person had to draw a spiral on a blank screen – no spiral template) and the bottom right spiral was drawn with a dominant hand without template. All the spirals can be animated by clicking on the “Animate” button below the specific spiral. If at any point you would like to stop the animation, you can do that with a click on the “Stop” button (shown in Figure C.3).

After the examination of the spirals, please answer the question on the left side regarding your opinion whether the spirals were drawn by a healthy person, the spirals show signs of Parkinsonian tremor or essential tremor or that you cannot decide.

Please enter your comment on the shown spirals or on your opinion in the space below your opinion. Your comments will be most valuable to our research. Please briefly explain your classification in terms of the spiral features that you consider most important for your classification.

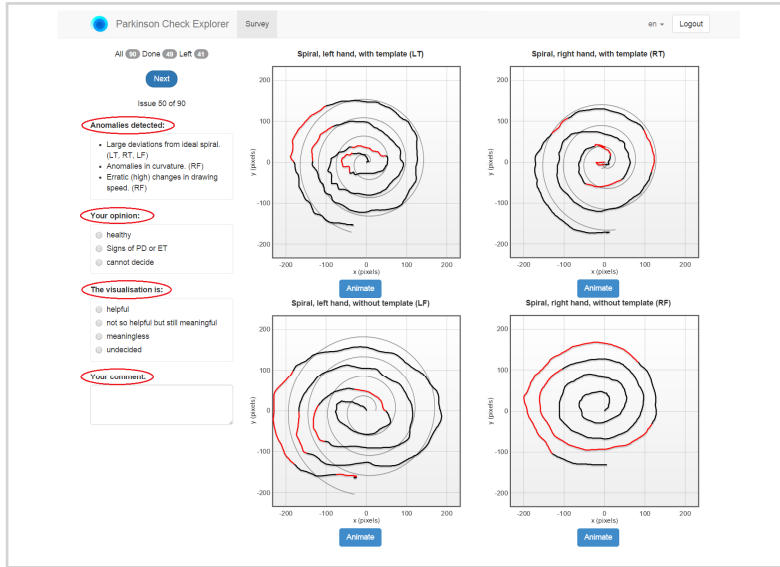


Slika C.3: Login screen for the survey.

After answering the question, you will be able to proceed to the next test subject by clicking on the “next” button on the top left side. Once you click on this button, you will not be able to go back to examine the previous spirals or change the opinion!

After examining 45 subjects, you will automatically enter step 2 of the survey (Figure C.4). In step 2, the same subjects’ spirals will be shown in the same order as in step 1. Some (or all) of the four spirals will have parts of the spiral marked with red colour (when static) and blue colour (when animated). These parts were detected by our algorithm as the “most critical” parts of the spiral with respect to different attributes in our ParkinsonCheck decision model. The static and animated visualisation may highlight different parts of the spiral as the highlighted attributes are different. Some of the attributes measure the deviation of the drawn spiral from the “ideal” Archimedean spiral. In the cases where we are taking into account these attributes, there will be two spirals shown on the same picture. The light grey spiral will be representing the “ideal” spiral. This is not the same as the before mentioned template. The “ideal” spiral is calculated based on the drawn spiral and represents the best-fit version of it.

On the left side, under “Anomalies detected”, you will find computer’s explanation



Slika C.4: An example of visualisation and computer's explanation of the visualisation.

of the visualisation. It includes a short explanation of the detected anomalies that are actually visualised in the spirals and references (e.g. LT – left template) to the spirals on which the anomaly was detected. Explanations will include five different interpretations of the visualised anomalies:

Large deviations from ideal spiral denotes a difference between the drawn spiral and its “ideal” spiral, as it was already explained in the previous paragraph.

Significant proportion of the spiral drawn towards the centre – the radius of the spiral should be monotonically increasing. We measure the proportion of the spiral where the radius decreased.

Anomalies in curvature means the computer detected some anomalies in the curvature, e.g. sharp angles in the spiral.

Erratic (high) changes in drawing speed means the speed of the drawing was not constant but it was significantly changing (increasing – decreasing) on some parts of the spiral.

Erratic (high) changes in radial speed means that the radial speed (radial speed measures speed towards or away from the starting point of the spiral) was not constant but it was significantly changing (increasing – decreasing) on some parts of the spiral.

Note that some of these anomalies can only be seen in animated spirals.

You are now asked to give once more your opinion on classification of the drawn spirals, and answer the added question about your opinion on the visualisation. Below you can find some guiding points for each answer. We would like to know whether the visualisation is:

helpful - it helped you detect the anomalies in the spirals you weren't aware (or were less aware) of without the visualisation,

not so helpful but still meaningful – you did not need any help detecting the anomalies in the spirals, but the visualisations mainly showed the crucial anomalies,

meaningless – the visualisation did not show the crucial anomalies, or

undecided – you cannot decide whether the visualisation was meaningful or not.

If you have any specific comments on your classification or visualisation, please enter them in the space below the questions. Even a short remark (a few words) is valuable.

We appreciate you taking the time to participate in our survey.



LITERATURA

- [1] Vida Groznik, Matej Guid, Aleksander Sadikov, Martin Možina, Dejan Georgiev, Veronika Kragelj, Samo Ribarič, Zvezdan Pirtošek, and Ivan Bratko. Elicitation of neurological knowledge with abml. In Mor Peleg, Nada Lavrač, and Carlo Combi, editors, *Artificial Intelligence in Medicine*, volume 6747 of *Lecture Notes in Computer Science*, pages 14–23. Springer Berlin Heidelberg, 2011. doi: [10.1007/978-3-642-22218-4_3](https://doi.org/10.1007/978-3-642-22218-4_3).
- [2] Matej Guid, Martin Možina, Vida Groznik, Dejan Georgiev, Aleksander Sadikov, Zvezdan Pirtošek, and Ivan Bratko. *ABML Knowledge Refinement Loop: A Case Study*, pages 41–50. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012. doi: [10.1007/978-3-642-34624-8_5](https://doi.org/10.1007/978-3-642-34624-8_5).
- [3] D Georgiev, V Groznik, A Sadikov, M Možina, M Guid, V Kragelj, I Bratko, S Ribarič, and Z Pirtošek. Digitalised spirography and clinical examination based decision support system for differentiating between tremors. *European Journal of Neurology*, 19:298, 2012.
- [4] V. Groznik, A. Sadikov, M. Možina, J. Žabkar, D. Georgiev, and I. Bratko. Attribute visualisation for computer-aided diagnosis: A case study. In *2014 IEEE International Conference on Healthcare Informatics*, pages 294–299, Sept 2014. doi: [10.1109/I CHI.2014.47](https://doi.org/10.1109/I CHI.2014.47).
- [5] Aleksander Sadikov, Vida Groznik, Jure Žabkar, Martin Možina, Dejan Georgiev, Zvezdan Pirtošek, and Ivan Bratko. Parkinsoncheck smart phone app. 263: 1213–1214, 01 2014.
- [6] Aleksander Sadikov, Jure Žabkar, Martin Možina, Vida Groznik, Dag Nyholm, and Mevludin Memedi. *Feasibility of Spirography Features for Objective Assessment of Motor Symptoms in Parkinson's Disease*, pages 267–276. Springer International Publishing, Cham, 2015. ISBN 978-3-319-19551-3. doi: [10.1007/978-3-319-19551-3_35](https://doi.org/10.1007/978-3-319-19551-3_35).
- [7] Vida Groznik, Matej Guid, Aleksander Sadikov, Martin Možina, Dejan Georgiev, Veronika Kragelj, Samo Ribarič, Zvezdan Pirtošek, and Ivan Bratko. Elicitation of neurological knowledge with argument-based machine learning. *Artificial Intelligence in Medicine*, 57(2):133 – 144, 2013. ISSN 0933-3657. doi: <https://doi.org/10.1016/j.artmed.2012.08.003>.
- [8] Aleksander Sadikov, Vida Groznik, Martin Možina, Jure Žabkar, Dag Nyholm, Mevludin Memedi, Ivan Bratko, and Dejan Georgiev. Feasibility of spirography features for objective assessment of motor function in parkinson's disease. *Artificial Intelligence in Medicine*, 2017. ISSN 0933-3657. doi: <https://doi.org/10.1016/j.artmed.2017.03.011>.
- [9] A Sadikov, J Žabkar, M Možina, V Groznik, D Georgiev, and I Bratko. Parkinsoncheck: A decision support system for spirographic testing. *University of Ljubljana, Faculty of Computer and Information Science, Tech. Rep.*, 2014.
- [10] Vida Groznik, Martin Možina, Jure Žabkar, Dejan Georgiev, Ivan Bratko, and Aleksander Sadikov. *Development, Debugging, and Assessment of Parkinson Check Attributes Through Visualisation*, pages 47–71. Springer International Publishing, Cham, 2015. ISBN 978-3-319-17963-6. doi: [10.1007/978-3-319-17963-6_4](https://doi.org/10.1007/978-3-319-17963-6_4).
- [11] Günther Deuschl, Peter Bain, and Mitchell Brin. Consensus statement of the movement disorder society on tremor. *Movement Disorders*, 13(S3):2–23, 1998. ISSN 1531-8257. doi: [10.1002/mds.870131303](https://doi.org/10.1002/mds.870131303).
- [12] P.G. Bain. Tremor. *Parkinsonism & Related Disorders*, 13, Supplement 3(0):S369 – S374, 2007. ISSN 1353-8020. doi: [https://dx.doi.org/10.1016/S1353-8020\(08\)70032-4](https://dx.doi.org/10.1016/S1353-8020(08)70032-4). Proceedings of the (XVII) (WFN) World Congress on Parkinson's Disease and Related Disorders.
- [13] G. Grimaldi and M. Manto. *Tremor: From Pathogenesis to Treatment*. Synthesis lectures on biomedical engineering. Morgan & Claypool Publishers, 2008. ISBN 9781598296389.
- [14] Veronika Kragelj, Dejan Georgiev, Zvezdan Pirtošek, and Samo Ribarič. Tremor. *Zdravniški Vestnik*, 81(1), 2012.

- [15] Abdul Qayyum Rana and Haris M Vaid. A review of primary writing tremor. *International Journal of Neuroscience*, 122(3):114–118, 2012.
- [16] Rachel Saunders-Pullman, Carol Derby, Kaili Stanley, Alicia Floyd, Susan Bressman, Richard B. Lipton, Amanda Deligdisch, Lawrence Severt, Qiping Yu, Mónica Kurtis, and Seth L. Pullman. Validity of spiral analysis in early Parkinson's disease. *Movement Disorders*, 23(4):531–537, 2008. doi: [10.1002/mds.21874](https://doi.org/10.1002/mds.21874).
- [17] M. Filipova, V. Filip, Z. Macek, S. Mullerova, J. Markova, S. Kas, B. Žižkova, J. Krivka, M. Votavova, and H. Krejčova. Terguride in parkinsonism a multicenter trial. *European archives of psychiatry and neurological sciences*, 237(5):298–303, 1988. doi: [10.1007/BF00450549](https://doi.org/10.1007/BF00450549).
- [18] Kaili Stanley, Johann Hagenah, Norbert Brüggenmann, Kathrin Reetz, Lawrence Severt, Christine Klein, Qiping Yu, Carol Derby, Seth Pullman, and Rachel Saunders-Pullman. Digitized spiral analysis is a promising early motor marker for parkinson disease. *Parkinsonism and Related Disorders*, 16(3):233–234, 2010. doi: [10.1016/j.parkreldis.2009.12.007](https://doi.org/10.1016/j.parkreldis.2009.12.007).
- [19] Jerker Westin, Samira Ghiamati, Mevludin Memedi, Dag Nyholm, Anders Johansson, Mark Dougherty, and Torgny Groth. A new computer method for assessing drawing impairment in parkinson's disease. *Journal of Neuroscience Methods*, 190(1):143–148, 2010. doi: [10.1016/j.jneumeth.2010.04.027](https://doi.org/10.1016/j.jneumeth.2010.04.027).
- [20] Xuguang Liu, Camille B. Carroll, Shou-Yan Wang, John Zajicek, and Peter G. Bain. Quantifying drug-induced dyskinesias in the arms using digitised spiral-drawing tasks. *Journal of Neuroscience Methods*, 144(1):47–52, 2005. doi: [10.1016/j.jneumeth.2004.10.005](https://doi.org/10.1016/j.jneumeth.2004.10.005).
- [21] S. H. Alusi, J. Worthington, S Glickman, L. J. Findley, and P. G. Bain. Evaluation of three different ways of assessing tremor in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 68(6):756–760, 2000. doi: [10.1136/jnnp.68.6.756](https://doi.org/10.1136/jnnp.68.6.756).
- [22] Peter Feys, Werner Helsen, Ann Prinsmel, Stephan Ilsbrouckx, Shouyan Wang, and Xuguang Liu. Digitised spirography as an evaluation tool for intention tremor in multiple sclerosis. *Journal of Neuroscience Methods*, 160(2):309–316, 2007. doi: [10.1016/j.jneumeth.2006.09.019](https://doi.org/10.1016/j.jneumeth.2006.09.019).
- [23] M.G. Longstaff and R.A. Heath. Spiral drawing performance as an indicator of fine motor function in people with multiple sclerosis. *Human Movement Science*, 25(4–5):474–491, 2006. doi: [10.1016/j.humov.2006.05.005](https://doi.org/10.1016/j.humov.2006.05.005). Advances in Graphonomics: Studies on Fine Motor Control, Its Development and Disorders.
- [24] Rodger J. Elble, Mikhail Brilliant, Keith Leffler, and Constance Higgins. Quantification of essential tremor in writing and drawing. *Movement Disorders*, 11(1):70–78, 1996. doi: [10.1002/mds.870110113](https://doi.org/10.1002/mds.870110113).
- [25] Elan D. Louis, Q. Yu, A. G. Floyd, Carol Moskowitz, and Seth L. Pullman. Axis is a feature of handwritten spirals in essential tremor. *Movement Disorders*, 21(8):1294–1295, 2006. doi: [10.1002/mds.20915](https://doi.org/10.1002/mds.20915).
- [26] J. P. Hubble, K.L. Busenbark, S. Wilkinson, R. D. Penn, K. Lyons, and W.C. Koller. Deep brain stimulation for essential tremor. *Neurology*, 46(4):1150–1153, 1996. doi: [10.1212/WNL.46.4.1150](https://doi.org/10.1212/WNL.46.4.1150).
- [27] P.G. Bain and L.J. Findley. *Assessing Tremor Severity: A Clinical Handbook*. Standards in neurology. Smith-Gordon, 1993.
- [28] F. Miralles, S. Tarongí, and A. Espino. Quantification of the drawing of an archimedes spiral through the analysis of its digitized picture. *Journal of Neuroscience Methods*, 152(1–2):18–31, 2006. doi: [10.1016/j.jneumeth.2005.08.007](https://doi.org/10.1016/j.jneumeth.2005.08.007).
- [29] Peter H. Kraus and Arndt Hoffmann. Spiralometry: Computerized assessment of tremor amplitude on the basis of spiral drawing. *Movement Disorders*, 25(13):2164–2170, 2010. ISSN 1531-8257. doi: [10.1002/mds.23193](https://doi.org/10.1002/mds.23193).
- [30] R J Elble, R Sinha, and C Higgins. Quantification of tremor with a digitizing tablet. *J Neurosci Methods*, 32(3):193–8, 1990. ISSN 0165-0270.
- [31] Seth L. Pullman. Spiral analysis: A new technique for measuring tremor with a digitizing tablet. *Movement Disorders*, 13(53):85–89, 1998. ISSN 1531-8257. doi: [10.1002/mds.870131315](https://doi.org/10.1002/mds.870131315).
- [32] Martin Možina, Jure Žabkar, and Ivan Bratko. Argument based machine learning. *Artificial Intelligence*, 171(10/15):922–937, 2007.
- [33] Peter Clark and Robin Boswell. Rule induction with CN2: Some recent improvements. In Yves Kodratoff, editor, *EWVL*, volume 482 of *Lecture Notes in Computer Science*, pages 151–163. Springer, 1991.
- [34] Martin Možina. *Argument Based Machine Learning*. Ph.d. thesis, University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia, 2009.
- [35] Steven T. Moore, Hamish G. MacDougall, Jean-Michel Gracies, Helen S. Cohen, and William G. Ondo. Long-term monitoring of gait in parkinson's disease. *Gait & Posture*, 26(2):200–207, 2007. doi: [10.1016/j.gaitpost.2006.09.011](https://doi.org/10.1016/j.gaitpost.2006.09.011).

- [36] Bogdan Pogorelec, Zoran Bosnic, and Matjaž Gams. Automatic recognition of gait-related health problems in the elderly using machine learning. *Multimedia Tools and Applications*, 58:333–354, 2012. doi: [10.1007/s11042-011-0786-1](https://doi.org/10.1007/s11042-011-0786-1).
- [37] Taha Khan, Jerker Westin, and Mark Dougherty. Cepstral separation difference: A novel approach for speech impairment quantification in parkinson's disease. *Biocybernetics and Biomedical Engineering*, 34(1):25–34, 2014. doi: [10.1016/j.bbe.2013.06.001](https://doi.org/10.1016/j.bbe.2013.06.001).
- [38] Taha Khan, Jerker Westin, and Mark Dougherty. Classification of speech intelligibility in parkinson's disease. *Biocybernetics and Biomedical Engineering*, 34(1):35–45, 2014. doi: [10.1016/j.bbe.2013.10.003](https://doi.org/10.1016/j.bbe.2013.10.003).
- [39] M.A. Little, P.E. McSharry, E.J. Hunter, J. Spielman, and L.O. Ramig. Suitability of dysphonia measurements for telemonitoring of parkinson's disease. *Biomedical Engineering, IEEE Transactions on*, 56(4):1015–1022, 2009. doi: [10.1109/TBME.2008.2005954](https://doi.org/10.1109/TBME.2008.2005954).
- [40] A. Tsanas, M.A. Little, P.E. McSharry, and L.O. Ramig. Accurate telemonitoring of parkinson's disease progression by noninvasive speech tests. *Biomedical Engineering, IEEE Transactions on*, 57(4):884–893, 2010. doi: [10.1109/TBME.2009.2036000](https://doi.org/10.1109/TBME.2009.2036000).
- [41] Mehmet Engin. A recording and analysis system for human tremor. *Measurement*, 40(3):288–293, 2007. doi: [10.1016/j.measurement.2006.05.015](https://doi.org/10.1016/j.measurement.2006.05.015).
- [42] A. Salarian, H. Russmann, C. Wider, P.R. Burkhard, F.J.G. Vingerhoets, and K. Aminian. Quantification of tremor and bradykinesia in parkinson's disease using a novel ambulatory monitoring system. *Biomedical Engineering, IEEE Transactions on*, 54(2):313–322, 2007. doi: [10.1109/TBME.2006.886670](https://doi.org/10.1109/TBME.2006.886670).
- [43] Mario C Barroso, Guilherme P Esteves, Thiago P Nunes, Lucia MG Silva, Alvaro CD Faria, and Pedro L Melo. A telemedicine instrument for remote evaluation of tremor: design and initial applications in fatigue and patients with parkinson's disease. *BioMedical Engineering OnLine*, 10(1):1–17, 2011. doi: [10.1186/1475-925X.10.14](https://doi.org/10.1186/1475-925X.10.14).
- [44] M Memedi, J Westin, and D Nyholm. Spiral drawing during self-rated dyskinesia is more impaired than during self-rated off. *Parkinsonism Relat Disord*, 19(5):533–536, 2013. doi: [10.1016/j.parkreldis.2013.01.011](https://doi.org/10.1016/j.parkreldis.2013.01.011).
- [45] Meng-Hsiang Yang, Yung-Hoh Sheu, Yuan-Hsing Shih, and M.S. Young. Portable tremor monitor system for real-time full-wave monitoring and analysis. *Review of Scientific Instruments*, 74(3):1303–1309, 2003. doi: [10.1063/1.1536259](https://doi.org/10.1063/1.1536259).
- [46] S. Patel, K. Lorincz, R. Hughes, N. Huggins, J. Growdon, D. Standaert, M. Akay, J. Dy, M. Welsh, and P. Bonato. Monitoring motor fluctuations in patients with parkinson's disease using wearable sensors. *Information Technology in Biomedicine, IEEE Transactions on*, 13(6):864–873, 2009. doi: [10.1109/ITTB.2009.2033471](https://doi.org/10.1109/ITTB.2009.2033471).
- [47] B Chen, S. Patel, T. Buckley, R. Rednic, D.J. McClure, L. Shih, D. Tarsy, M. Welsh, and P. Bonato. A web-based system for home monitoring of patients with parkinson's disease using wearable sensors. *Biomedical Engineering, IEEE Transactions on*, 58(3):831–836, 2011. doi: [10.1109/TBME.2010.2090044](https://doi.org/10.1109/TBME.2010.2090044).
- [48] Daniel Rodríguez-Martin, Carlos Perez-Lopez, Albert Sama, Joan Cabestany, and Andreu Catala. A wearable inertial measurement unit for long-term monitoring in the dependency care area. *Sensors*, 13(10):14079–14104, 2013. doi: [10.3390/s131014079](https://doi.org/10.3390/s131014079).
- [49] Janez Demšar, Tomaž Curk, Aleš Erjavec, Črt Gorup, Tomaž Hočevar, Mitar Milutinović, Martin Možina, Matija Polajnar, Marko Toplak, Anže Starič, Miha Štajdohar, Lan Umek, Lan Zagar, Jure Žbontar, Marinka Žitnik, and Blaž Zupan. Orange: Data mining toolbox in python. *Journal of Machine Learning Research*, 14:2349–2353, 2013.
- [50] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874, 2008. doi: [10.1145/1390681.1442794](https://doi.org/10.1145/1390681.1442794).
- [51] Bojan Cestnik. Estimating probabilities: A crucial task in machine learning. In *ECAL*, pages 147–149, 1990.
- [52] Usama M. Fayyad and Keki B. Irani. Multi-interval discretization of continuous-valued attributes for classification learning. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence. Chambéry, France, August 28 - September 3, 1993*, pages 1022–1029, 1993.
- [53] Xiaopeng Xi, Eamonn J. Keogh, Christian R. Shelton, Li Wei, and Chotirat Ann Ratanamahatana. Fast time series classification using numerosity reduction. In *Machine Learning, Proceedings of the Twenty-Third International Conference (ICML 2006), Pittsburgh, Pennsylvania, USA, June 25-29, 2006*, pages 1033–1040, 2006. doi: [10.1145/1143844.1143974](https://doi.org/10.1145/1143844.1143974).