

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Danijel Grah

Sistem za podporo grafološke analize

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJ RAČUNALNIŠTVA IN
INFORMATIKE

MENTOR: prof. dr. Igor Kononenko

Ljubljana 2012

Rezultati diplomskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil \LaTeX .



Št. naloge: 01810/2012

Datum: 15.03.2012

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **DANIJEL GRAH**

Naslov: **SISTEM ZA PODPORO GRAFOLOŠKE ANALIZE
A SYSTEM FOR GRAPHOLOGY ANALYSIS SUPPORT**

Vrsta naloge: Diplomsko delo univerzitetnega študija

Tematika naloge:

Grafologija je veda, ki je razvila principe za analizo ročno zapisanega teksta in razbrati psihološke tipe ljudi, ki so tekst pisali. Naloga je implementirati sistem za podporo dela grafologov. Na voljo je baza analiziranih in klasificiranih rokopisov, ki so jih opravili profesionalni grafologi, ter program, ki so ga razvili študenti v okviru seminarske naloge pri predmetu Tehnologija znanja. Naloga je dopolniti obstoječi program z izračunavanjem parametrov, ki še niso implementirani, razviti grafologu prijazen grafični vmesnik, povezati program z grafološko tabelo interpretacij ter narediti nekaj poskusov s strojnim učenjem dodatnih interpretacij, ki povezujejo več izmerjenih parametrov. Diplomsko delo se bo izvajalo v sodelovanju strokovnjaka grafologa.

Mentor:

prof. dr. Igor Kononenko

Dekan:

prof. dr. Nikolaj Zimic



IZJAVA O AVTORSTVU DIPLOMSKEGA DELA

Spodaj podpisani Danijel Grah, z vpisno številko **63040041**, sem avtor diplomskega dela z naslovom:

Sistem za podporo grafološke analize

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom prof. dr. Igorja Kononenka,
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 3. september 2012

Podpis avtorja:

Zahvalil bi se mentorju prof. dr. Igorju Kononeku za vodenje, odzivnost in strokovne nasvete pri izdelavi diplomske naloge. Prav tako bi se zahvalil Magdaleni Metelko za nešteto nasvetov, čas in strokovnost. Posebna zahvala gre vsem, ki so me podpirali v času študija.

“The day science begins to study non-physical phenomena, it will make more progress in one decade than in all the previous centuries of its existence.”

-Nikola Tesla

Kazalo

Povzetek

Abstract

1	Uvod	1
1.1	Opis problema in ciljev	1
1.2	Pregled vsebine	2
2	Grafološka analiza	3
2.1	Grafologija in pisava	3
2.2	Umeščenost grafološke analize	4
2.3	Meritve grafološke analize	5
3	Izbrane metode strojnega učenja	13
3.1	Opis in priprava podatkov	13
3.2	Testiranje in rezultati	14
3.3	Izbrane metode	15
4	Razvoj in uporaba sistema	17
4.1	Podobni sistemi	17
4.2	Orodja in knjižnice	22
4.3	Načrtovanje	22
4.4	Implementacija	24
4.5	Namestitev in struktura datotek	28
4.6	Uporaba	29

KAZALO

4.7	Rezultati opravljenih meritev	40
4.8	Rezultati napovedovanja	42
5	Sklepne ugotovitve	43

Povzetek

V diplomski nalogi smo se osredotočili na razvoj sistema, ki sodi med aplikativne rešitve na področju grafologije. V sodelovanju s strokovnjakom grafologom smo izdelali uporabniško prijazen sistem, ki omogoča opravljanje, urejanje in shranjevanje meritev. Na voljo smo imeli podatke na katerih smo testirali metode strojnega učenja. Odločili smo se vključiti metodo kNN ter implementirati preprost vmesnik za vnos poljubnega števila odločitvenih pravil. S pomočjo strokovnjaka grafologa smo definirali in vnesli 88 odločitvenih pravil, ki v sklepnem delu vsebujejo splošne opise osebnostnih značilnosti ali poklicni tip.

Natančnost sistema smo testirali na 10-tih rokopisih. Opravljene meritve smo primerjali z originalnimi podatki in opazili večja odstopanja. Razlog za napake vidimo v pomanjkanju strokovnega znanja pri opravljanju meritev. Pregledali smo tudi interpretacije, ki jih sistem poda na podlagi meritev. Ugotavljamo, da na podlagi napovedi ne moremo enolično sklepati na poklicni tip, niti ne moremo natančno določiti osebnostnih značilnosti. Vzrok za nenatančnost pripisujemo preveč splošno definiranim sklepnem delu odločitvenih pravil.

Ugotavljamo, da nam je uspelo razviti sistem, ki nadomešča standardna orodja, ki jih grafolog uporablja pri svojem delu. Prav tako menimo, da je sistem primeren za demonstracijske namene in poučevanje grafologije.

Ključne besede

sistem, grafologija, grafološka analiza, strojno učenje

Abstract

In the diploma thesis we focused on the development of the system that belongs to software solutions in the field of graphology. In cooperation with an expert in the field of graphology we created a user friendly system that enables managing and saving the measurements. The Machine learning methods were tested on the data that we had at disposal. We decided to include the kNN method. A simple interface for entering any number of decision rules was also implemented. With the help of the expert graphologist 88 decision rules, which combine general predictions of personal features or professional types in the final part of the decision rules, were defined and entered.

We tested the accuracy of the system with 10 manuscripts and compared measurements made with the original data. Larger deviations were noticed. In our opinion the cause for these mistakes is lack of professional knowledge in making the measurements. We established that we cannot uniquely conclude on a professional type based on prediction, nor can we thoroughly determine personal features and we think that the cause for inaccuracy lays in too generally defined final part of the decision rules.

We have determined that we succeeded to create a system that substitutes standard tools which are used in a graphologists work. We also think that the system is suitable for demonstration purposes in teaching Graphology.

Keywords

system, graphology, graphoanalysis, machine learning

Poglavje 1

Uvod

Vse od kar je Michon postavil temelje za grafologijo s svojo znamenito zvezo *Signes fixes*, se je njena uporaba razširila in izpopolnila. Kljub razdvojenosti pri priznavanju grafologije kot znanstvene discipline, jo danes večina psihologov obravnava kot psihodiagnostično sredstvo. Grafologija kot znanost o človeškem izrazu [11] pozna več različnih sistemov analize pisave. Eden izmed priznanih sistemov je grafološka analiza.

1.1 Opis problema in ciljev

Grafolog pri svojem delu uporablja orodja, kot so: šestilo, ravnilo, kotomer in svinčnik. S pomočjo orodij opravlja meritve rokopisov, ki so ponavadi na belem papirju formata A4. Opravljene meritve zapisuje, išče najbolj pogosto vrednost, povprečno vrednost in na podlagi izmerjenih vrednosti poizkuša razbrati psihološki tip ljudi. Samo opravljanje meritev je zamudni in vendar rutinski proces. Napovedovanje od grafologa zahteva veliko izkušenj in strokovnih znanj, nekateri avtorji trdijo, da tudi intuicije. Da bi olajšali in pohitrili delo grafologa, smo si pri razvoju sistema za podporo grafološki analizi zadali naslednje cilje:

- omogočiti natančno in hitro izvajanje meritev,
- omogočiti avtomatično preslikavo meritev v rezultate,

- povezati rezultate z grafološko ustreznimi interpretacijami,
- narediti nekaj poskusov s strojnim učenjem dodatnih interpretacij, ki povezujejo več meritev.

Nekatere zamisli za doseg ciljev in motivacijo za reševanje opisanega problema smo našli v seminarski nalogi [7], ki je nastala v okviru predmeta Tehnologija znanja. Naloga rešuje opisano problematiko in nam je služila kot izhodišče pri razvoju sistema za podporo grafološke analize.

1.2 Pregled vsebine

Diplomsko delo obsega 5 poglavij. Prvo poglavje vsebuje kratek uvod v vsebino diplomskega dela. V naslednjem poglavju opišemo grafologijo in opredelimo grafološko analizo ter predstavimo meritve, ki jih grafolog pri svojem delu opravlja. V tretjem poglavju opišemo pridobljeno tabelo podatkov in prikažemo rezultate testiranja izbranih metod strojnega učenja. Poglavje zaključimo z argumentacijo izbranih metod strojnega učenja, ki jih kasneje vključimo v sistem. V naslednjem poglavju najprej predstavimo podobne sisteme in nato podrobno opišemo načrtovanje, implementacijo in uporabo sistema. Načrtovanje in implementacijo razdelimo na funkcionalne bloke in skladno z razdelitvijo opišemo, kako smo funkcionalnost načrtovali oziroma kako smo jo razvili. Uporabo sistema smo razdelili glede na funkcionalnosti, ki jih sistem pri delu ponuja in jih po tem vzoru tudi opisali. Zaključimo z ugotovitvami o prednostih in pomanjkljivostih sistema ter podamo nekaj idej za nadaljnji razvoj sistema.

Poglavje 2

Grafološka analiza

2.1 Grafologija in pisava

Iznajdbo pisave štejemo kot začetek človeške zgodovine. Prvi začetki pisave v obliki jamskih slikarij segajo v obdobje mlajšega paleolitika. Že takrat je človek čutil potrebno, da zapiše govorjeno besedo oziroma izrazi svoj miselni proces. Potreba po pismenosti je naraščala s časom in skladno s tem je naraščalo tudi zanimanje za simboličnost človeške besede. Vse od kar so priznana imena, kot so B.J. Huxley, W. Heinz, N. Chomsky, I.Pavlov, dodelili znanosti o simboličnosti človeške besede pomembno mesto v znanstveni psihologiji, je simbolična razlaga pisave dobila nove pobude in trdnejšo znanstveno podlago [11].

Proces pisanja poteka pod nadzorom centralno živčnega sistema, ki upravlja z mišicami in slednje z okončinami, ki se v samem procesu uporabljajo. Gibanje, ki je produkt sodelovanja mišic in centralno živčnega sistema, večji delež časa v procesu pisanja ni pod zavestnim nadzorom [13]. Informacije, ki jih izda pisana beseda, so lahko odraz zavesti ali pa izhajajo iz podzavesti subjekta. Pisana beseda torej izraža stanje subjekta v času pisanja [1]. Michon, ki ga v [4] omenjajo kot začetnika znanstvenega gibanja grafologije, je trdil, da je fiziološko in psihološko stanje subjekta v času pisanja konstantno. Iz tega izhaja, da so tudi grafični simboli konstantni. Vendar Trstenjak

v [11] opozarja, da direktnih zvez med grafičnimi simboli in osebnostnimi potezami ne moremo dobiti. Pisalni znaki tvorijo celoto in šele v kombinaciji jim lahko pripisujemo pomen. L. Klages v svojem delu, kot navaja Trstenjak, vpelje dvoumnost znakov. Posameznemu znaku lahko pripišemo pozitiven ali negativen pomen. Njegovo naravo razkrije grafolog s pomočjo dopolnilnih vidikov, ki jih tvorijo drugi znaki. Pri tem se ne moremo izogniti subjektivnemu odločanju, ki je prisotno pri vseh psihodiagnostičnih tehnikah. Grafologu, kot pravi Pfanne, pri odločitvi pomaga dobro razvit kinestetski čut. Z izrazom se Pfanne izogiba domnevi o intuiciji, ki jo grafolog potrebuje pri svojem delu. Namesto tega postavi v ospredje potrebo po nadarjenosti. Tako kot ostali poklici poleg strokovnosti in izkušenj v določenih primerih potrebujejo posebne značilnosti človeka, tudi grafolog svoje delo opravi boljše in lažje, če poseduje kinestetski čut.

Grafologijo lahko po zaslugi J.H.Michona, J. Crepiex-Jamina, W.Langenburga, H. Busseja, W.Preyerja, G.Meyerja in L. Klagesa opredelimo kot posebno panogo znanosti o človeškem izrazu. Zaradi številnih strogo psihometrično usmerjenih psihologov težko govorimo o grafologiji kot o razlagi človeške osebnosti na podlagi pisave. Očitki se nanašajo predvsem na to, da je grafologija premalo eksaktna. Če pisavo definiramo kot psihofizično kretnjo organizma, ki se odraža v obliki potez in krivulj na pisalni površini, potem lahko grafologijo opredelimo kot nauk o psihofizičnih kretnjah posameznika. Ne glede na opredelitev in številne očitke psihologi danes grafologijo uporabljajo kot psihodiagnostično sredstvo na raznih področjih: pri poklicnem in osebnem svetovanju, gospodarstvu in industriji, pri izbiri zakonskega partnerja in ugotavljanju karakterističnih značilnosti obtožencev na sodiščih [11].

2.2 Umeščenenost grafološke analize

V grafologiji poznamo tri pristope:

- Integrativna (angl. *Integrative*) grafologija: Povezuje strukture, ki jih

tvori pisana linija, z osebnostnimi značilnostmi. Pristop je usmerjen od znotraj navzven.

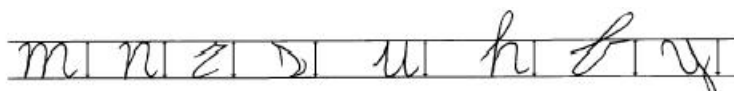
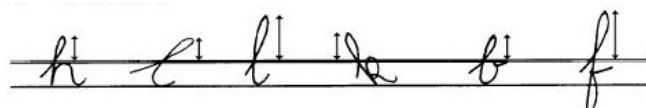
- Holistična (angl. *Holistic*) grafologija: Osebnostni profil se določi na podlagi oblike, gibanja in prostora. Pristop je usmerjen od zunaj navznoter.
- Simbolična (angl. *Symbolic*) analiza: Pri tem pristopu grafolog išče simbole, ki se skrivajo v pisavi.

Vsak pristop je razvil več sistemov analize pisav. **Grafološka analiza** je sistem, ki ga uvrščamo v Intergrativno grafologijo [13].

2.3 Meritve grafološke analize

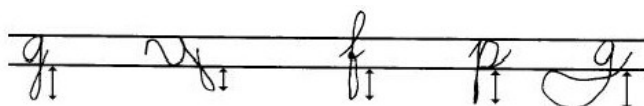
Grafološka analiza kot sistem analize pisav temelji na meritvah. Različne šole uporabljajo različne skupine meritev in njim pripadajočih metrik. Naše delo je upoštevalo meritve po zgledu [5]. Na podlagi zgleda in po posvetu z grafologom smo prišli do končne skupine meritev, ki smo jih kasneje implementirali v sistemu. Skupino meritev prikazujemo v tem razdelku. Če ni drugače navedeno, smo pri posamezni meritvi upoštevali najbolj pogosto, najnižjo, najvišjo in povprečno vrednost ter uporabili mersko enoto *mm*. Prav tako smo (zaradi preglednosti) posamezne meritve, ki izhajajo iz skupnega tipa, zapisali skupaj. V skladu z definicijami, kot jih podajajo v [5], v nadaljevanju predstavljamo izbrane meritve grafološke analize:

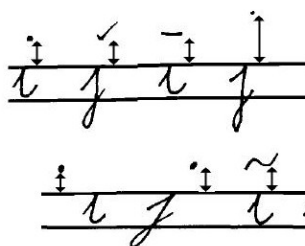
- *Višina in širina srednje cone*: Meritev opravimo tako, da ravnilo položimo na osnovno črto in merimo celotno srednjo cono. Višino srednje cone predstavlja razdalja med skrajnima točkama srednje cone v vertikalni smeri, širino pa v horizontalni smeri. Primer črk s srednjo cono, kjer je bila opravljena meritev *višina srednje cone*, prikazuje slika 2.1.
- *Višina in širina zgornje cone*: Merimo tako, da ravnilo postavimo na navidezno črto med srednjo in zgornjo cono, to je na višičke črk *m*,

Slika 2.1: Meritev *višina srednje cone*Slika 2.2: Meritev *višina zgornje cone*

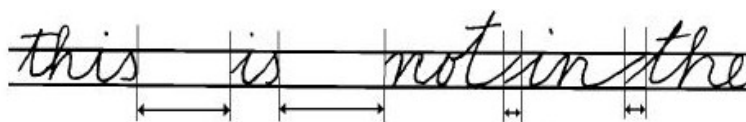
a, e itd. in izmerimo višino oziroma širino. Pri tem ne upoštevamo naklona. Primer črk z zgornjo cono, kjer je bila opravljena meritev *višina zgornje cone*, prikazuje slika 2.2

- *Dolžina in širina spodnje cone*: Spodnja cona se nadaljuje iz srednje cone navzdol. Meritev opravimo tako, da ravnilo postavimo na osnovno črto in izmerimo razdaljo med skrajnima točkama spodnje cone v horizontalni smeri, če gre za širino, in v vertikalni smeri, če gre za dolžino. Pri opravljanju meritve se ne oziramo na naklon. Primer črk s spodnjo cono, ki jim je bil opravljena meritev *dolžina spodnje cone*, prikazuje slika 2.3.
- *Višina in širina ovalov*: Postopek merjenja višine in širine ovalov je enak postopku merjenja *višine in širine srednje cone*. Pri tem upoštevamo samo ovale.
- *Višina in širina velikih začetnic*: Pri opravljanju meritve *višina in*

Slika 2.3: Meritev *dolžina spodnje cone*



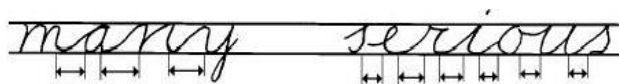
Slika 2.4: Meritev višina diakritičnih znakov



Slika 2.5: Meritev presledki med besedami

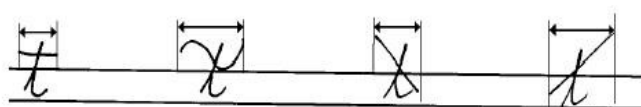
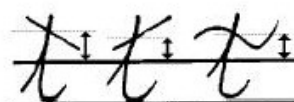
širina velikih začetnic postopamo enako kot pri opravljanju meritve *višina in širina srednje cone*. Upoštevamo samo velike začetnice.

- *Višina diakritičnih znakov*: Meritev opravimo tako, da izmerimo razdaljo med skrajno točko srednje cone in diakritičnim znakom v smeri vertikale. Primer opravljenih meritev prikazuje slika 2.4.
- *Presledki med besedami*: Meritev opravimo tako, da izmerimo razdaljo med zadnjo potezo ene besede in prvo potezo naslednje besede. Če se potezi prepletata, imamo opravka z negativno meritvijo. Primer opravljene meritve *presledki med besedami* prikazuje slika 2.5.
- *Presledki med črkami*: Pri opravljanju meritve *presledki med črkami* postopamo enako, kot pri opravljanju meritve *presledki med besedami*. Osnovna enota pri opravljanju meritve je črka. Primer prikazuje slika 2.6.
- *Presledki med vrsticami*: Postopek pri opravljanju meritve je enak kot pri meritvah *presledki med črkami* in *presledki med besedami*. Upoštevamo razdaljo med vrsticami.

Slika 2.6: Meritev *presledki med črkami*Slika 2.7: Meritev *presledki med podaljški*

- *Presledki med podaljški*: Meritev opravimo enako kot pri ostalih meritvah presledkov. Osnovna enota je podaljšek črke. Primer opravljanje meritve nam prikazuje slika 2.7. Na sliki vidimo pozivno in negativno meritev.
- *20 horizontalnih presledkov*: Meritev opravimo tako, da izmerimo 20 črk v smeri horizontale (vključno s presledki). Opravljeno meritev prikazuje slika 2.8.
- *Zgornji, spodnji, levi in desni rob*: Skupino meritev opravljamo vsako posebej. Meritev *levi rob* opravimo tako, da izmerimo razdaljo od levega roba do prve poteze pisalne črte v horizontalni smeri. Postopek za opravljanje meritev ostalih robov je enak.
- *Vsi štirje robovi*: Meritve *vsil štirje robovi* ne opravljamo, ampak je neposredna izpeljanka posameznih meritev robov. Pri izračunu najbolj

Slika 2.8: Meritev *20 horizontalnih presledkov*

Slika 2.9: Meritev *dolžina črtice na t*Slika 2.10: Meritev *višina črtice na t*

pogoste, povprečne, največje in najmanjše vrednosti potem upoštevamo vse štiri robove.

- *Višina in dolžina črtice na t*: Višino črtice na *t* dobimo tako, da izmerimo razdaljo črtice do srednje cone črke *t*. Pri merjenju dolžine črtice na *t* ne upoštevamo naklona. Primer meritev prikazujeta sliki 2.9 in 2.10.
- *Višina osebnega zaimka*: Pri opravljanju meritve *višina osebnega zaimka* upoštevamo osebni zaimek, v kolikor je prisoten (osebni zaimek “I” v angleških tekstih ali “Ich” v nemških tekstih). Meritev opravimo kot ostale meritve, kjer merimo višino.
- *Širina duktusa*: Merimo širino pisalne črte.
- *Širjenje in oženje spodnjega, zgornjega, levega in desnega roba*: Merimo odmik na začetku in na koncu (upoštevamo prvo in zadnjo meritev za vsako vrsto roba posebej) ter ugotavljamo, ali se rob spreminja.
- *Poravnava osnovne linije*: Merimo spremembo osnovne linije v horizontalni smeri. Meritev opravimo na sredini in na koncu. Prav tako ugotavljamo, če se osnovna linije spreminja.

- *Naklon*: Rezultat meritve je v procentih. Pri naklonu ne upoštevamo povprečne vrednosti. Ravno tako kot pri poravnavi osnovne linije in robovih tudi pri naklonu ugotavljamo, ali se spreminja.
- *Rezmerje črnilo-belina*: Merimo samo odstotek črnine, ki je v procentih.
- *Figura*: Meritev *figura* ni podobna ostalim meritvam. Figura, ki jo dobimo na papirju velikosti A4, je črna risba človeka. Na risbi izmerimo višino in širno celotne figure, višino in širino glave, ter ostalih delov telesa. Potreba po meritvi je nastala po posvetu z grafologom, ki meritve na figuri opravlja za pomoč pri poklicnem svetovanju.

Tabela 2.1 prikazuje meritve, ki se opravljajo pri posameznih črkah.

Črka	Zgornja cona	Srednja cona	Spodnja cona	Oval	Drugo
a				x	
b	x	x			
c		x			
č		x			višina diakritičnega znaka
d	x			x	
e		x			
f	x	x	x		
g			x	x	
h	x	x			
i		x			višina diakritičnega znaka
j		x	x		višina diakritičnega znaka
k	x	x			
l	x	x			
m		x			
n		x			
o				x	
p		x	x		
r		x			
s		x			
š		x			višina diakritičnega znaka
t	x	x			višina in dolžina črtice na t
u		x			
v		x			
z		x			
ž		x			višina diakritičnega znaka
A-Ž					višina in širina velikih začetnic

Tabela 2.1: Prikaz meritev, ki se opravljajo pri posameznih črkah

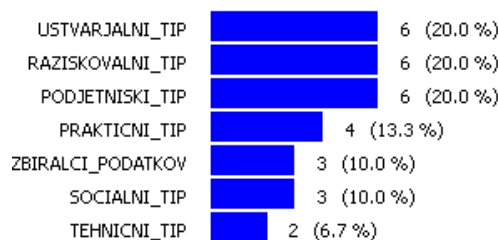
Poglavje 3

Izbrane metode strojnega učenja

3.1 Opis in priprava podatkov

Metode strojnega učenja, ki predstavljajo vejo umetne inteligence, so v zadnjih dvajsetih letih doživele veliko izboljšav. Posledično jih vključujejo številni komercialni sistemi [3]. V ta namen smo pridobili opravljene meritve za 30 rokopisov. Vsak rokopis ima 168 atributov, ki so neposredna izpeljanka tabele meritev, ki je prikazana v Dodatku 1. Od tega je 161 atributov zveznih in 7 diskretnih. Rokopisi so uvrščeni v sedem razredov: *Podjetniški tip*, *raziskovalni tip*, *zbiralci podatkov*, *ustvarjalni tip*, *tehnični tip*, *praktični* in *socialni tip*. Distribucijo razredov v podatkih prikazuje slika 3.1. Opazimo, da imajo razredi *ustvarjalni*, *raziskovalni* in *umetniški tip* največjo zastopanost v podatkih. Večinski razred z verjetnostjo 0.2 torej ni enolično določen.

Po posvetu z grafologom smo se odločiti iz tabele izpustiti meritve, ki se končajo z besedo *nepraviln-a/i*. Prav tako smo se odločili izpustiti (za vsako meritev) kategorijo *nihanje*. Naslednji posegi v tabeli so bili: Odstranitev kategorije *mm na začetku* pri meritvi *poravnava osnovne linije*, odstranitev kategorije *povprečna vrednost* pri meritvi *naklon* in odstranitev vseh kategorij razen kategorije *najbolj pogosta meritev* pri meritvi *razmerje črnilo:belina*.



Slika 3.1: Distribucija razredov

Slednja pravzaprav pomeni samo eno opravljeno meritev. Do končne strukture podatkov in skupine meritev smo prišli po posvetu z grafologom in kasnejši analizi. Pokazala se je potreba po vključitvi novih meritev. Dodali smo meritve: *višina črtice na t*, *dolžina črtice na t*, *razdalja prepletenosti* in kategorijo *oblika* pri meritvi *poravnava osnovne linije*. Prav tako smo dodali skupino meritev *figura*. Posledično je bilo potrebno podatkom dodati attribute, ki so neposredna izpeljanka dodanih meritev. Vsi omenjeni posegi so privedli do podatkov s skupaj 138 atributi. Od tega je 130 zveznih in 8 diskretnih. Testiranje je potekalo na tako spremenjeni strukturi podatkov. Pri tem novi dodani atributi niso imeli vpliva na testiranje, saj noben rokopis ni imel podatkov za omenjene attribute.

3.2 Testiranje in rezultati

Za namene testiranja metod strojnega učenja smo uporabili programski paket *Orange*. Primerjali smo tri metode:

- Klasifikacijsko drevo,
- Naivni Bayes,
- kNN.

Testirali smo po metodi izloči enega (angl. *leave one out*) ter primerjali klasifikacijsko točnost posameznih metod. Rezultate prikazuje tabela 2.1.

Metoda	Klasifikacijska točnost	Ploščina pod krivulju ROC
Naivni Bayes	0.2	0.4788
kNN	0.1667	0.5212
Klasifikacijsko drevo	0.2	0.3926

Tabela 3.1: Klasifikacijska točnost izbranih metod

Ploščino pod krivuljo ROC smo računali po privzetih nastavitvah programskega paketa *Orange*. Privzeta metoda pri večrazrednih problemih, ki se uporablja za računanju ploščine pod krivuljo ROC, se imenuje *ByWeighted-Pairs*. Metoda računa ploščino pod krivuljo ROC za vsak par razredov. Pri tem ne upošteva primerov, ki pripadajo ostalim razredom. Delni rezultat, ki ga metoda dobi za posamezni par, je obtežen s produktom verjetnosti teh razredov. Seštevek delnih rezultatov nam da končni rezultat ploščine pod krivuljo ROC.

3.3 Izbrane metode

Rezultati, prikazani v tabeli 2.1, so pričakovani. Klasifikacijska točnost z najboljšim rezultatom 0.2 ni dovolj visoka, da bi se odločili, za katero od metod. Prav tako ploščina pod krivuljo ROC, ki se pri vseh metodah giblje okrog 0.5, dodatno okrepi našo odločitev. Po [3] je ploščina pod krivuljo ROC enaka verjetnosti, da bo klasifikator razločil med pozitivnim in negativnim primerom. Iz povedanega lahko sklepamo, da se metode obnašajo podobno kot, če bi napovedi prepustili naključni izbiri oziroma še slabše. Vendar lahko rezultate upravičimo z razmerjem števila atributov proti številu primerov. Imamo 168 atributov in 30 primerov. Potrebovali bi bistveno več primerov, da bi za omenjeno število atributov dobili rezultate z zadostno stopnjo zaupanja.

Ne glede na rezultate smo se odločili, da bomo v sistem vključili metodo kNN, razlog zato pa je v zasnovi sistema. Sistem je bil načrtovan tako, da omogoča shranjevanje opravljenih meritev za posamezne primere, ki so v

našem primeru rokopisi. Če bo sistem v uporabi dovolj časa, se bo število primerov povečevalo. Zaradi značilnosti kNN, da klasificira testni primer glede na razrede najbližjih sosedov, predvidevamo, da se bo klasifikacijska točnost metode kNN povečavala. Koristna lastnost kNN metode je tudi, da omogoča vpogled v primere, ki so testnemu najbližji. Grafologu bo tako omogočena neposredna primerjava testnega primera z najbližjimi že analiziranimi primeri.

Delo grafologa velikokrat poteka intuitivno in neposredno ne upošteva meritev, ki so na voljo [14]. Intuicija temelji na znanju in izkušnjah, znanje pa lahko predstavimo na več načinov, pri čemer so odločitvena pravila eden izmed načinov [3]. Predstavitev pravil smo posplošili tako, da smo v pogojnem delu dovolili poljubne logične operatorje. Na ta način smo povečali izrazno moč pravil. V sklepnem delu odločitvenih pravil se nismo omejili na določeno zalogo vrednosti. Tako lahko v sklepnem delu nastopa poljuben simbol, številka ali tekst. Omenjeno omogoča predstavitev širokega obsega strokovnega znanja, ki ga grafolog potrebuje pri svojem delu. Odločitvena pravila imajo še eno za nas koristno lastnost - modularnost. Pravila so torej zamenljiva, kar omogoči preprosto vzdrževanje. Opisane lastnosti odločitvenih pravil so zadostovale za našo odločitev za vključitev v sistem.

Poglavje 4

Razvoj in uporaba sistema

4.1 Podobni sistemi

Cilj sistema za podporo grafološki analizi je popolna avtonomnost. Kot popolno avtonomnost v tem delu razumemo sposobnost samostojnega delovanja sistema po korakih, ki kot vhod prejmejo rokopis in kot izhod vrnejo končno oceno obravnavanega rokopisa. Korake bi lahko strnili v:

- segmentacija,
- razpoznavanje,
- merjenje,
- interpretacija.

V zadnjih letih so razvili nekaj komercialnih in tudi prosto dostopnih sistemov, ki se ukvarjajo s problematiko naštetih področij. V [6] navajajo pet sistemov za računalniško podprto grafološko analizo: Jerral Sapienza's Self-Analysis, Andy Hunt's Graphonomizer, Handwriting University's Handwriting Wizard, Sheila Lowe's Handwriting Analyzer in Garth Michaels' Handwriting Analyst. Sistemi po naši opredelitvi pokrivajo področje interpretacije. Skupen jim je princip delovanja. Sistemi zahtevajo interakcijo uporabnika, Sheila Lowe's Handwriting Analyzer tudi grafologa ali ustrezno

usposobljenega uporabnika. Po seriji vprašanj ali postavljenih zahtev generirajo končno oceno v obliki profila osebnosti.

V [15] zasledimo primere sistemov, ki pokrivajo področje razpoznavanja in segmentacije. V grobem lahko sisteme, ki se ukvarjajo z omenjeno problematiko, razdelimo glede na način razpoznavanja. Prva vrsta je sposobna razpoznavanja v realnem času (angl. *on line*). V to kategorijo spadajo sistemi, ki temeljijo na napravah z zasloni, občutljivimi na dotik, in vsi ostali, ki opravljajo razpoznavanje v času pisanja. Druga vrsta sistemov kot vir za razpoznavanje uporablja optično prebrane rokopise ali druge oblike dokumentov. Spadajo v kategorijo sistemov, ki nalogo razpoznavanja opravljajo na že napisanem primerku (angl. *off line*). Ne glede na vrsto, sistemi težijo k učinkoviti segmentaciji, ki je predpogoj za uspešno razpoznavanje. Na področju razpoznavanja poznamo več pristopov. Najbolj razširjena sta: nevronske mreže in procesiranje značilk. Nevronske mreže omogočajo hitro vpeljavo, vendar ponujajo le malo kontrole pri sledljivosti rezultatov. Njihova učinkovitost je odvisna od kvalitete naučenega. Značilke po drugi strani omogočajo večji nadzor v procesu razpoznavanja. Ker lastnosti značilk niso naučene avtomatično, posledično razvoj razpoznavanja traja več časa. Poznamo več komercialnih sistemov, ki sodijo v področje razpoznavanja pisave v realnem času. V sklopu strojne opreme je bil njihov namen ponuditi alternativo tipkovnici kot standardni vhod računalniškega sistema. Primer modernega sistema za razpoznavanje pisave so operacijski sistemi podjetja Microsoft, ki so nameščeni na tabličnih računalnikih. Temeljijo na t.i. Klasifikatorju TDNN (Time delay neural network). V grobem lahko rečemo, da sta sistemom skupna strojna oprema, ki predstavlja medij pri posredovanju pisave, in programska oprema za razpoznavanje napisanega. Pri tem se nekateri poslužujejo tudi učenja karekteristik iz pisav posameznih oseb. V interakciji z uporabnikom sistem poizkuša ugotoviti za uporabnika značilne poteze v pisavi. Slednje izboljša zanesljivost razpoznavanja [15].

S problemom razpoznavanja se ukvarjajo tudi v [2]. Opisana je vpeljavo LCS (angl. Learning classifier system), ki je v bližnji povezavi z ge-

netičnimi algoritmi in učenjem z nagrajevanjem [16]. Za namene testiranja so ustvarili 20.000 različnih primerov črk, ki pripadajo 20 različnim oblikam računalniških pisav (angl. *font*). Primer generiranih črk prikazuje slika 4.1. Vsako črko so naključno degenerirali, da bi bila bolj podobna črki, ki je nastala pod človeško roko. Množico so razdelili na učno in testno. Od tega je bilo 16.000 primerov učnih. Po 5 iteracijah učenja skozi 16.000 primerov je sledilo uvrščanje 4.000 primerov testne množice. V postopku učenja so testirali več metod predstavitve črk, generiranja pravil oziroma klasifikatorjev ter njihovega točkovanja. Najuspešnejša kombinacija metod in vrednosti parametrov je dosegla na testnih primerih klasifikacijsko točnost 80,8%.

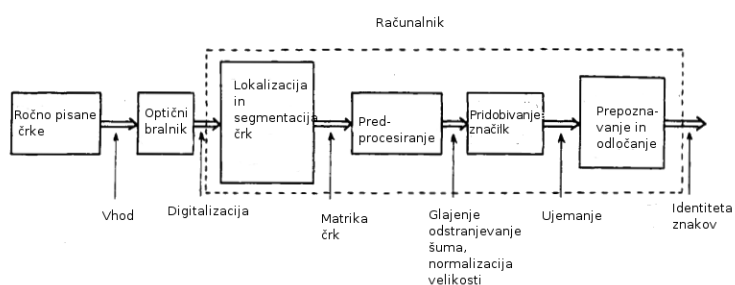
V [10] zasledimo obsežno primerjavo metod in tehnik razpoznavanja. Ugotavljajo, da z ustrezno velikostjo učne množice in z odstranitvijo dvoumnosti zapisanih znakov metode dosegajo klasifikacijsko točnost tudi nad 90%. Sam proces razpoznavanja opisujejo, kot prikazuje slika 4.2. Po njihovi opredelitvi na težavnost problema razpoznavanja vplivajo naslednji dejavniki:

- šum,
- izkrivljanje,
- različni stili,
- translacija,
- rotacija.

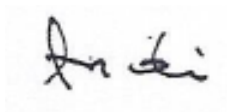
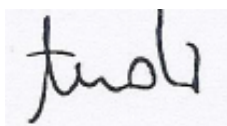
Človek neposredno vpliva na vse dejavnike razen šuma. Omenjeno in trditev ameriškega statistika, ki jo navajajo v [1], da je verjetnost identičnosti dveh rokopisov 1 proti $68 * 10^{12}$, problem segmentacije in razpoznavanja, uvrščajo med zahtevne probleme. Trditev lahko argumentiramo tudi izkustveno. Slika 4.3 in slika 4.4 prikazujeta dva različna rokopisa za isto besedo. Opazimo, da le s težavo ugotovimo pomen zapisane besede na sliki 4.4. Šele na podlagi konteksta lahko sklepamo na pomen besede. Prav tako bi težko razločili črko *i* na sliki 4.3, če ne bi sklepali, da je črka del besede *tudi*.



Slika 4.1: Prikaz primerov generiranih črk



Slika 4.2: Proces razpoznavanja

Slika 4.3: Beseda *tudi* v prvem rokopisuSlika 4.4: Beseda *tudi* v drugem rokopisu

Sistem, ki ga predlagajo v [8], po naši opredelitvi pokriva področje merjenja in interpretacije. Delovanje sistema razdelijo na več korakov. Med temi so: Predobdealava rokopisa, segmentacija, pridobivanje značilk, učenje in testiranje s pomočjo metode podpornih vektorjev (SVM) ter rezultati. Segmentacija pozna tri osnovne segmente: črke, besede in vrstice. Na podlagi dobljenih segmentov sistem avtomatično pridobi značilke, ki jih kot vhod uporabi metoda SVM. Pri tem moramo omeniti, da sistem segmentacijo ne izvede avtomatično. Sistem so testirali na množici 100 rokopisov, na katerih je bila opravljena grafološka analiza. Dve tretjini rokopisov in pripadajočih grafoloških analiz so uporabili za učenje sistema in ostalo tretjino za testiranje. Sistem na podlagi kriterijev, ki so jih dobili s pomočjo jedrne funkcije RBF (Gaussian radial basis function) in metode SVM, posamezne značilke uvrsti v določen razred, ki mu pripada psihološki pomen. Navajajo 90.3% točnost pri uvrščanju.

V [9] navajajo sistem za računalniško podprto grafologijo, ki je sposoben pridobiti značilke in analizirati obravnavan rokopis. Glede na našo opredelitev avtonomnosti opisan sistem pokriva vsa področja. Vendar ga obravnavamo kot nepopoln za naše potrebe. Razlog je v številu meritev, ki jih je sistem sposoben sam opraviti, ter razlikah v metrikah posameznih meritev. Sistem v fazi pridobivanja značilk rokopis najprej segmentira v vrstice. Na

posameznih vrsticah potem pridobi naslednje značilke, ki jih v našem delu imenujemo meritve: levi, desni, zgornji in spodnji rob, presledki med vrsticami, poravnava osnovne linije, naklon ter zgornja, srednja in spodnja cona. V [9] nismo zasledili podrobno opisanih postopkov, ki jih sistem izvede pri opravljanju posamezne meritve.

Trenutno stanje na opisanih področjih in težavnost problema segmentacije in razpoznavanja je vplivalo na našo odločitev pri razvoju sistema.

4.2 Orodja in knjižnice

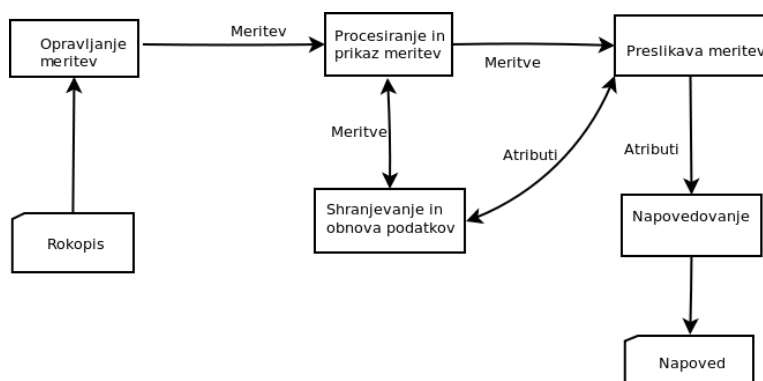
Razvoj sistema je potekal v programskem jeziku *Java* s pomočjo razvojnega orodja *Eclipse*. Za označevanje rokopisov smo uporabili prosto dostopno programsko opremo *ImageJ*, ki smo jo kot paket vključili v sistem. Čeprav smo potrebovali le del funkcionalnosti, ki ga ponuja *ImageJ*, smo se odločili vključiti celotni paket. Glavni argument za našo odločitev je fleksibilnost in širok nabor funkcionalnosti. Povedano omogoča lažje vzdrževanje in dopolnjevanje sistema z novimi metodami.

4.3 Načrtovanje

Načrtovanje sistema ni potekalo po priznanih metodologijah razvoja informacijskih sistemov. Sistem smo razvijali namenu primerno (angl. *Ad-Hoc*). Pri tem smo se zgledovali po [9]. Funkcionalnost sistema prikazuje diagram na sliki 4.5.

4.3.1 Opravljanje meritev

Funkcionalnost, kot vhod prejme rokopis. V procesu opravljanja meritev poteka merjenje in označevanje rokopisa. Pri tem lahko rokopis označuje stroj ali grafolog. Opravljanje meritev smo načrtovali ohlapno. To pomeni, da smo najprej predvideli, da meritve opravlja grafolog. Sistem grafologu tako nudi funkcionalnost, ki nadomešča standardna orodja, katera grafolog pri svojem



Slika 4.5: Funkcionalnost sistema

delu uporablja. Nismo izključili možnosti, da bo meritve opravljal stroj. Z vključitvijo nekaterih metod iz področja strojnega vida bi bistveno olajšali delo grafologa. V ta namen smo funkcionalnost načrtovali tako, da omogoča vključiti meritve, ki jih je opravil stroj, pri tem pa dopušča grafologu, da te meritve pregleda in poljubno spreminja. Opisan pristop omogoča preverjanje kakovosti in pravilnosti meritev ter nudi dobro uporabniško izkušnjo.

4.3.2 Procesiranje in prikaz meritev

Predvidena funkcionalnost skrbi za prepoznavanje vhodnih podatkov kot meritev, njihovo ustrezno razvrščanje in prikazovanje v obliki rezultatov. Podatke pridobiva neposredno iz funkcionalnosti, ki skrbi za opravljanje meritev. Prikazani rezultati izhajajo iz skupine meritev in podatkov, ki so opisani v razdelku 3.1.

4.3.3 Shranjevanje in obnova podatkov

V sistemu smo predvideli funkcionalnost shranjevanja in obnove podatkov. Omenjen sklop funkcij skrbi za obnovo in trajnost dokončanih kot tudi nedokončanih meritev. Z dokončanimi meritvami razumemo rokopis, na katerem so bile opravljene vse meritve, ki so po mnenju grafologa bile potrebne. Opi-

san del sistema izvršuje ukaze, ki jih dobi od funkcionalnosti procesiranja in prikaza meritev.

4.3.4 Preslikava meritev

Sklop funkcij iz funkcionalnosti preslikava meritev skrbi za ustrezno preslikavo meritev v domeno atributov. Preslikava je skladna z domeno atributov v podatkih, ki smo jih pridobili v sodelovanju z grafologom.

4.3.5 Napovedovanje

Funkcionalnost napovedovanja vključuje metode strojnega učenja. Načrtovali smo vključitev metod, ki smo jih izbrali v razdelku 3.3. Metode strojnega učenja za svoje napovedovanje in učenje uporabljajo končno podobo podatkov, kot so opisani v razdelku 3.1.

4.4 Implementacija

4.4.1 Opravljanje meritev

Za implementacijo načrtovane funkcionalnosti smo uporabili prosto dostopno knjižnico *ImageJ*. V orodjarno smo vključili nekatera standardna orodja, ki jih *ImageJ* ponuja. Poleg metod in razredov knjižnice smo uporabili tudi vtič za upravljanje meritev (*ROIManager*). Vtič nam je omogočil implementacijo prisotnosti več meritev naenkrat, prav tako pa tudi njihovo upravljanje.

4.4.2 Procesiranje in prikaz meritev

Funkcionalnost procesiranja in prikaza meritev izhaja iz standardnih objektov in metod programskega jezika Java. V ozadju deluje logika, ki iz enega okna (angl. *Frame*) preslika podatke in jih prikaže v obliki meritev na drugem. V nadaljevanju prikazujemo enačbo, s pomočjo katere smo preslikali velikostne razrede meritev na računalniškem zaslonu v okvire papirja A4.

$$v = \frac{v_{izmerjena} * v_{A4}}{v_{rokopisa}} \quad (4.1)$$

kjer $v_{izmerjena}$ pomeni višino/širino opravljene meritve na računalniškem zaslonu, v_{A4} dejansko višino/širino papirja A4 in $v_{rokopisa}$ višino/širino na računalniškem zaslonu prikazanega rokopisa.

4.4.3 Shranjevanje in obnova podatkov

Sklop funkcij shranjevanja in obnove podatkov smo implementirali s standardnimi metodami in objekti programskega jezika *Java*. Odločili smo se za shranjevanje podatkov v datoteko tipa *csv*. Našo odločitev argumentiramo z načinom uporabe sistema in viri, ki so bili na razpolago v času implementacije. Sistem je bil načrtovan kot namizna aplikacija in kot taka grafologu omogoča opravljanje meritev na poljubnem rokopisu. Če bi se odločili za centralizirano implementacijo sistema in enotno podatkovno bazo, bi prav tako morali centralizirati shranjevanje in nalaganje rokopisov. V nasprotnem primeru bi prišlo do podvajanje meritev za različno poimenovane a vendar identične rokopise. V našem primeru ima vsak rokopis, na katerem so bile opravljene meritve, svojo datoteko tipa *csv*. Prav tako obstaja ena datoteka tipa *csv*, ki hrani preslikane vrednosti meritev v domeno atributov. Datoteko uporabljajo metode strojnega učenja. Standardna oblika datotek za shranjevanje podatkov omogoča uvoz in izvoz podatkov v primerih morebitne centralizacije sistema.

4.4.4 Preslikava meritev

Implementirana funkcionalnost preslika meritve v domeno atributov. Sklop funkcij smo implementirali s pomočjo standardnega nabora metod in objektov programskega jezika *Java*.

4.4.5 Napovedovanje

Sklop funkcij, ki jih zajema napovedovanje, vključuje implementacijo metode kNN in odločitvenih pravil. Odločitvena pravila smo implementirali s pomočjo prosto dostopne programske knjižice *mvel*, ki je dostopna na [12]. Omenjena knjižica nam je omogočila preprosto vpeljavo in evaluacijo operatorjev algebre *Bool*. Pri implementaciji kNN smo morali rešiti nekaj problemov, ki so prikazani v nadaljevanju.

- **Manjkajoče vrednosti atributov**

Odločili smo se, da algoritem ne upošteva atributov ob direktni primerjavi učnega in testnega primera, če oba primera za dani atribut nimata ustrezne vrednosti. Prvotna zamisel je bila, da manjkajoče vrednosti zapolnimo s povprečno vrednostjo obravnavanega atributa. Vendar smo ugotovili, da učni podatki vsebujejo veliko manjkajočih vrednosti. Pravzaprav ni nobenega primera, ki bi imel prav za vsak atribut vrednost. Dodatni argument k naši odločitvi smo našli v načinu dela grafologa. Ni nujno, da grafolog opravi vse meritve pri analizi trenutnega rokopisa. Zato tudi trenutni primer, ki mu iščemo najbližje sosedo, vsebuje poljubno število manjkajočih vrednosti. Omenjeno smo upoštevali pri računanju razdalje med primeri.

- **Normalizacija vrednosti atributov**

Kot omenjeno v razdelku 3.1, imamo opravka tako z zvezni kot diskretnimi atributi. Zaradi različne zaloge vrednosti zveznih atributov smo njihove vrednosti normalizirali s pomočjo enačbe (4.2). Za zvezni atribut A_i velja:

$$v_n = \frac{v - v_{min}}{v_{max} - v_{min}} \quad (4.2)$$

kjer v_n pomeni normalizirano vrednost trenutnega primera za atribut A_i , v_{min} in v_{max} pa predstavljata najmanjšo in največjo vrednost atributa A_i . Pri diskretnih atributih normalizacija vrednosti ni bila potrebna.

- **Razdalja med primeri**

Metrika, ki smo jo izbrali pri računanju razdalje med trenutnim primerom in učnimi primeri, je evklidska razdalja. Enačbe za računanje razdalje med dvema primeroma in uvrščanje smo povzeli po [3]. Pri tem smo razdaljo normalizirali glede na število skupnih atributov dveh primerov. Skupni atributi so določeni kot atributi, pri katerih oba primera vsebujeta ustrezno vrednost.

Tako je razdalja podana z:

$$D(u_l, u_j) = \frac{m}{a} * \sqrt{\sum_{i=1}^a d(v^{(i,l)}, v^{(i,j)})^2} \quad (4.3)$$

kjer je m število skupnih atributov in a število vseh atributov, pri čemer za zvezni atribut A_i velja:

$$d(v^{(i,l)}, v^{(i,j)}) = |v^{(i,l)} - v^{(i,j)}| \quad (4.4)$$

in za diskretnega

$$d(v^{(i,l)}, v^{(i,j)}) = \begin{cases} 0, & v^{(i,l)} = v^{(i,j)} \\ 1, & v^{(i,l)} \neq v^{(i,j)} \end{cases} \quad (4.5)$$

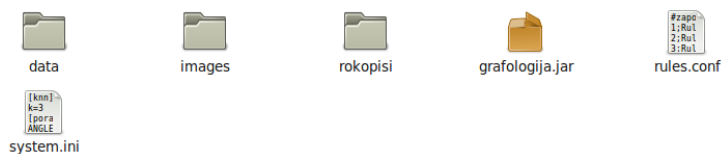
- **Izbira parametra k in verjetnostna distribucija**

Končna ocena metode kNN je verjetnost, da izbrani primer pripada določenemu razredu. Vsakemu razredu, ki je zastopan v množici k najbližjih učnih primerov, pripada verjetnost p_r , ki jo računamo po enačbi (4.6).

$$p_r = \frac{1}{k} * \sum_{i=1}^k \sigma(r, r^{(i)}) \quad (4.6)$$

kjer je

$$\sigma(a, b) = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases} \quad (4.7)$$

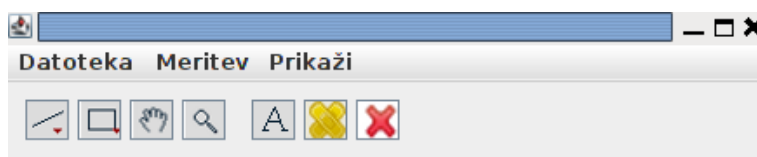


Slika 4.6: Datotečna struktura programa

Množica vseh verjetnosti p_r in njim pripadajočih razredov predstavlja verjetnostno porazdelitev po razredih, ki jo dobimo pri obravnavi trenutnega primera. Število najbližjih primerov k smo določili empirično. Privzeta vrednost ob zagonu sistema je 3, ki pa jo lahko spreminjamo v nastavitvah.

4.5 Namestitev in struktura datotek

Pogoj za delovanje sistema je nameščena različica *jre* (*Java Virtual Machine*) različice 1.5 ali višje. Za uspešno namestitev moramo kopirati celotno mapo *grafoloski sistem* v ciljno mapo, kjer imamo pravico pisanja. Sistem, ki je bil zasnovan kot namizna aplikacija, ima preprosto datotečno strukturo, ki je prikazana na sliki 4.6. Vidimo lahko prisotnost map *rokopisi*, *data* in *images*. Mapa *rokopisi* hrani rokopise, mapa *data* pa podatke o meritvah za posamezni rokopis, podatke o meritvah figur, shranjene opombe ter skupne podatke, ki jih potrebujejo metode strojnega učenja. Mapa *images* vsebuje slike, ki so del grafičnega vmesnika sistema. Poleg omenjenih map sta prisotni še nastavitvena datoteka *system.ini* in datoteka *rules.conf*, ki vsebuje odločitvena pravila.



Slika 4.7: Osnovno okno

4.6 Uporaba

4.6.1 Osnovno okno

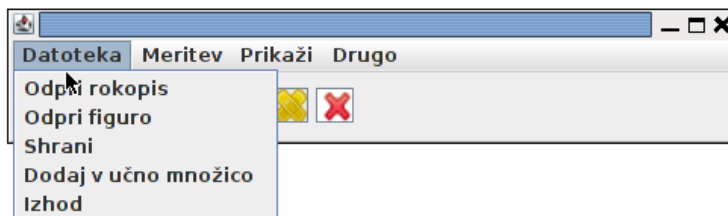
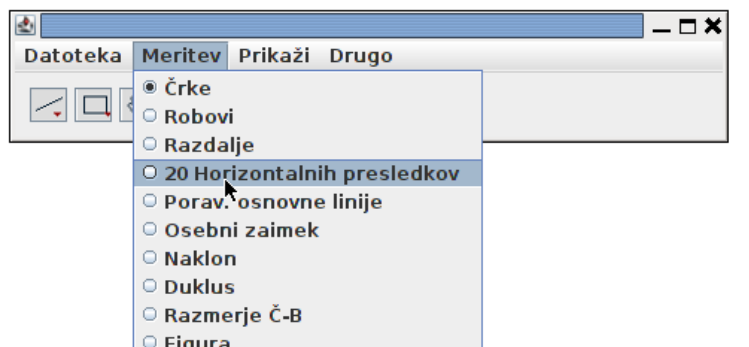
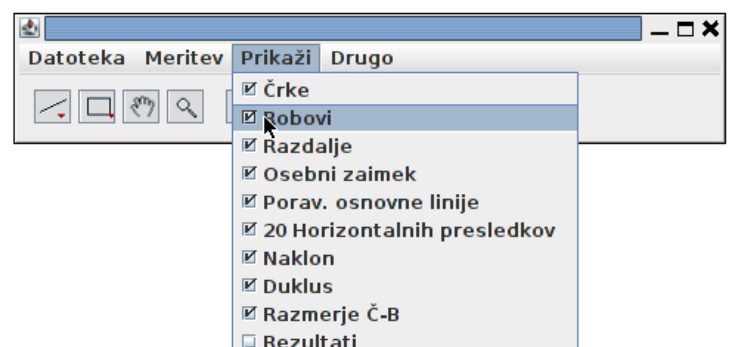
Posledica zagona programa je osnovno okno, kot prikazano na sliki 4.7. Osnovno okno sestavljata orodna vrstica in meni, ki sta izhodišče za nadaljnje delo.

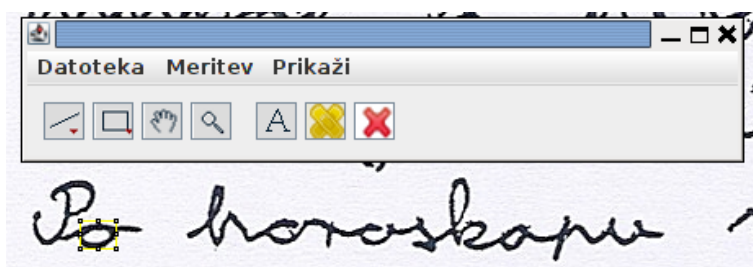
Orodna vrstica

Orodna vrstica je dinamična. Vsebuje elemente za opravljanje meritev. Orodja, ki sestavljajo orodno vrstico in sodijo v skupino orodij za opravljanje meritev, so: *Pravokotnik*, *Ravna črta*, *Povečava*, *Roka* in *Prelomljena črta*. Orodna vrstica prav tako vsebuje orodja, ki omogočajo dodajanje, spreminjanje in brisanje meritve. Slednje spadajo v drugo skupino orodij. Prikaz orodij obeh skupin se spreminja glede na izbiro meritve, ki jo želimo opraviti.

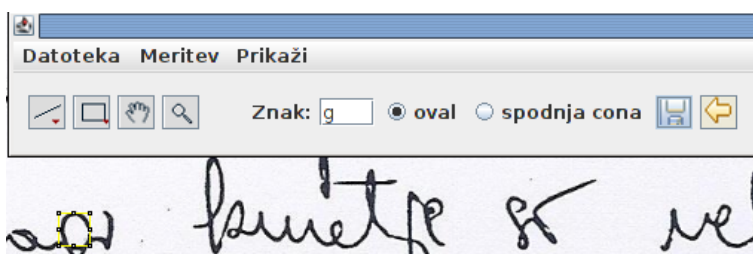
Meni

Meni sistema je sestavljen iz štirih delov: *Datoteka*, ki vsebuje ukaze za odpiranje in nalaganje rokopisov ter shranjevanje opravljenih meritev, *Meritev*, ki vsebuje razvrščene meritve in pripadajoče ukaze za določanje načina delovanja sistema, *Prikaži*, ki skrbi za prikaz posamezne skupine meritev, in *Drugo*, kjer lahko dodamo rezultate za meritve na figuri. Podrobno sestavo menija prikazuje zaporedje slik 4.8, 4.9, 4.10

Slika 4.8: Prikaz elementov menija *Datoteka*Slika 4.9: Prikaz elementov menija *Meritve*Slika 4.10: Prikaz elementov menija *Prikaži*



Slika 4.11: Meritev črke



Slika 4.12: Dodajanje meritve črke

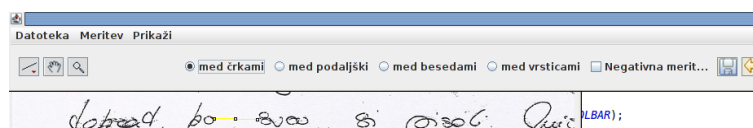
4.6.2 Opravljanje meritev

Meritve se opravljajo skladno z navodili, ki smo jih podali v razdelku 2.3.

Črke

Opravljanje meritev črk pričnemo tako, da naložimo rokopis ter v meniju izberemo skupino meritev *Črke*. V orodni vrstici izberemo orodje *pravokotnik* in označimo poljuben del ali celotno črko. Rezultat opisane operacije prikazuje slika 4.11, kjer smo označili črko *o*.

Naslednji korak je dodajanje meritve. Operacijo izvedemo, tako da kliknemo gumb *Dodaj*. Prikaže se nam vnosno polje, kjer vnesemo črko, ki smo jo označili. Nato nam sitem predlaga meritev, ki jo lahko za izbrano črko opravimo. Korak zaključimo s klikom na gumb *Shrani*.



Slika 4.13: Merjenje razdalje

Razdalje

Opravljanje meritev razdalje pričnemo tako, da v meniju pod *Meritve* izberemo skupino meritev *Razdalje*. Za merjenje razdalj potrebujemo orodje *Črta*. Izberemo orodje in označimo del rokopisa, ki ga želimo izmeriti. Sistem nam ob kliku na gumb *Dodaj* predlaga meritve, ki so na voljo. Omeniti je še potrebno, da so razdalje med črkami in podaljški lahko tudi negativne. V ta namen služi opsijsko polje *Negativna meritve*, ki jo lahko izberemo, če gre za meritev, ki je negativna. Korak zaključimo s klikom na gumb *Shrani*.

Robovi

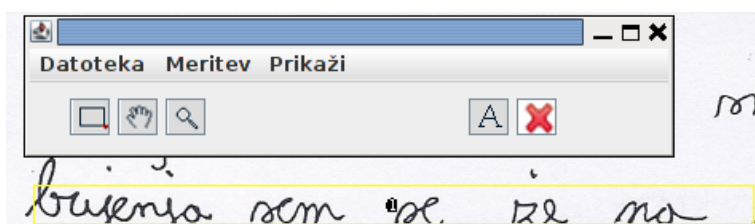
Podobno kot pri razdaljah imamo pri robovih na voljo orodje *Črta*. Preden začnemo z opravljanjem meritve, moramo v meniju pod *Meritve* izbrati skupino meritev *Robovi*. Postopek opravljanja meritve roba je podoben postopku pri opravljanju meritve razdalje. Razlika je le v predlaganih meritvah sistema. Sistem nam ob kliku na gumb *Dodaj* predlaga meritve levi, desni, zgornji ali spodnji rob.

Poravnava osnovne linije

Meritve *poravnava osnovne linije* pričnemo tako, da v meniju pod *Meritve* izberemo skupino meritev *Poravnava osnovne linije*. Za opravljanje meritve uporabimo orodje *Prelomljena črta*. Lik, ki ga nariše orodje, se od navadne črte razlikuje po dodatni gibljivi točki in ima tako tri točke, ki so povezane in premične. Primer opravljene meritve prikazuje slika 4.14

Koko se pripravi za preslovy
v hili Bled o. Ge Ham metodi.

Slika 4.14: Meritev poravnava osnovne linije



Slika 4.15: Meritev 20 horizontalnih presledkov

20 horizontalnih presledkov

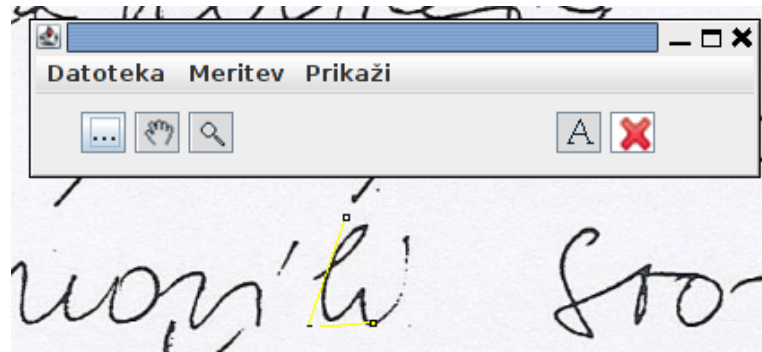
Meritev *20 horizontalnih presledkov* opravimo tako, da v meniju pod *Meritve* izberemo skupino meritev *20 horizontalnih presledkov*. Za opravljanje meritve potrebujemo orodje *Pravokotnik*. Rezultat opravljene meritve prikazuje slika 4.15

Naklon

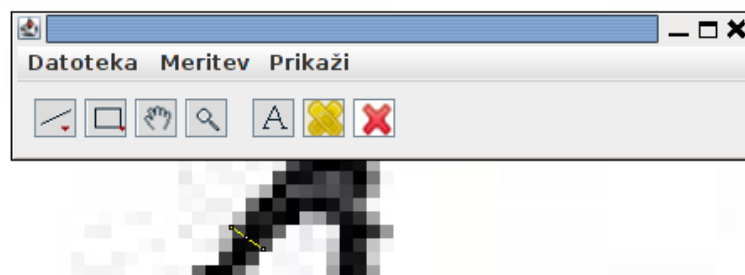
Podobno kot pri meritvi *poravnava osnovne linije* tudi pri opravljanju meritve *naklon* uporabimo orodje *Prelomljena črta*. Sam postopek opravljanje meritve se bistveno ne razlikuje od opravljanja ostalih meritev. Opravljeno meritev prikazuje slika 4.16.

Duktus

Za opravljanje meritve *duktus* v meniju pod *Meritve* izberemo skupino meritev *Duktus*. Uporabimo orodje *Črta* in *Povečava*. Najprej rokopis povečamo do ustrezne povečave, nato izberemo orodje *Črta* in označimo del rokopisa, ki nas zanima. Primer opravljane meritve prikazuje slika 4.17



Slika 4.16: Meritev naklon



Slika 4.17: Meritev širina duktusa

Razmerje črnilo-belina

Podobno kot pri ostalih meritvah tudi pri opravljanju meritve *razmerje črnilo-belina* izberemo ustrezno skupino meritev v meniju. Sam postopek opravljanje meritve se bistveno ne razlikuje od opravljanja ostalih meritev. Uporabimo orodje *Pravokotnik*.

Figura

Za opravljanje meritve *figura* v meniju pod *Datoteka* izberemo možnost *Odprti figuro* in naložimo sliko figure. Nato v meniju pod *Meritve* izberemo opcijsko polje *Figura*. Orodji, ki jih imamo na razpolago pri opravljanju meritve sta *Ravna črta* in *Pravokotnik*. Ko opravimo meritev in kliknemo na gumb *Dodaj*, nam sistem predlaga meritve *Celota*, *Glava* ali *Deli*. Izberemo ustrezno meritev in končamo s klikom na gumb *Shrani*.

4.6.3 Preslikava meritev v rezultate

Meritve, ki so vidne kot označbe na rokopisu, niso ustreznega velikostnega razreda. Preslikava je potekala s pomočjo enačbe (4.8) za meritve z vodoravno orientacijo in enačbe (4.9) za meritve s horizontalno orientacijo. S pomočjo omenjenih enačb smo dobili normalizirane meritve, ki ustrezajo realni velikosti meritve na papirju formata A4. Izjema je meritev širina duktusa, kjer enačbi (4.8) in (4.9) nista zadostovali za uspešno preslikavo. Meritev širina duktusa nima neposredne horizontalne ali vertikalne orientacije, ampak je v veliko primerih kombinacija obeh. S pomočjo trigonometričnih funkcij smo izpeljali niz enačb, s katerimi smo meritev uspešno preslikali. Enačbe so prikazane v nadaljevanju.

Najprej izračunamo projekcijo opravljene meritve na os X .

$$x_{projekcija} = \frac{x_{izmerjena} * SA4}{S_{rokopisa}} \quad (4.8)$$

kjer je $x_{izmerjena}$ vrednost, ki jo dobimo kot razliko skrajnih točk daljice v komponenti X ($x_2 - x_1$).

Po istem postopku izračunamo projekcijo opravljene meritve na os Y

$$y_{projekcija} = \frac{y_{izmerjena} * v_{A4}}{v_{rokopisa}} \quad (4.9)$$

Na koncu po *Pitagorovem* izreku izračunamo dejansko vrednost opravljene meritve.

$$v = \sqrt{x_{projekcija}^2 + y_{projekcija}^2} \quad (4.10)$$

Naslednji korak pri preslikavi meritev v rezultate je zaokroževanje. Meritev naklon smo zaokrožili na 5 stopinj natančno. Ostale meritve smo zaokrožili na eno decimalo natančno. Posebnost sistema je tudi določanje oblike osnovne linije, ki spada pod meritev *poravnava osnovne linije*. Sistem kot vhod prejme tri točke, na podlagi katerih izračuna kot v sredinski točki. Če kotu, ki ga tvorijo točke, prištejemo vnaprej določeno vrednost *EPS* in ta ne presega 180 stopinj, sistem predvideva, da imamo opravka z osnovno linijo, ki ni ravna. Nato po enačbi (4.11) ugotovi, če je oblika konkavna ali konveksna. V nasprotnem primeru je osnovna linija ravna. Pri določanju, če osnovna linija pada, narašča ali ostaja ravna, sistem upošteva Y komponento prve in zadnje točke. Če je razlika vrednosti večja od nekega vnaprej določenega *EPS*, sistem predvideva, da imamo opravka z osnovno linijo, ki narašča ali pada. V nasprotnem primeru je osnovna linija ravna. Negativen predznak razlike pomeni, da osnovna linija narašča, pozitiven pa, da pada.

$$r = (x_3 - x_1) * (y_2 - y_1) - (y_3 - y_1) * (x_2 - x_1) \quad (4.11)$$

Rezultat za meritev *razmerje črnilo-belina* dobimo po enačbi, ki jo prikazujemo v nadaljevanju.

$$crnina = \frac{n_{crnih}}{n_{belih} + n_{crnih}} \quad (4.12)$$

kjer n_{crnih} pomeni število slikovnih pik (angl. *pixels*) črne barve, n_{belih} pa število slikovnih pik bele barve. Pri tem moramo omeniti, da sitem izbrano območje najprej pretvori v sivinsko sliko s pomočjo vnaprej določenega praga.

MERITEV	Najbolj pogosta (mm)	Najnižja (mm)	Najvišja (mm)	Povprečna (mm)
Višina ovalov	2.1 Uredi	1.2 Uredi	7.1 Uredi	2.3 Uredi
Širina ovalov	4.1 Uredi	2.5 Uredi	6.3 Uredi	4.3 Uredi
Višina srednje cone	2.8 Uredi	1.1 Uredi	4.8 Uredi	2.8 Uredi
Širina srednje cone	2.3 Uredi	0.7 Uredi	6.1 Uredi	2.3 Uredi
Višina zgornje cone	7.3 Uredi	5.7 Uredi	13 Uredi	8.5 Uredi
Širina zgornje cone	2.4 Uredi	0.9 Uredi	3.3 Uredi	2.1 Uredi
Dolžina spodnje cone	6.3 Uredi	4.2 Uredi	8.1 Uredi	6.2 Uredi
Širina spodnje cone	1 Uredi	1 Uredi	3.3 Uredi	1.9 Uredi
Višina velikih začetnic	7 Uredi	7 Uredi	13.6 Uredi	10.7 Uredi
Širina velikih začetnic	18.1 Uredi	4.8 Uredi	18.1 Uredi	10.1 Uredi
Višina diakritičnih znakov	Uredi	Uredi	Uredi	Uredi
Presledki med besedami	5.4 Uredi	5.4 Uredi	6.5 Uredi	5.8 Uredi
Presledki med črkami	Uredi	Uredi	Uredi	Uredi
Presledki med podaljški	Uredi	Uredi	Uredi	Uredi
Presledki med vrsticami	10.6 Uredi	0 Uredi	10.6 Uredi	10.0 Uredi

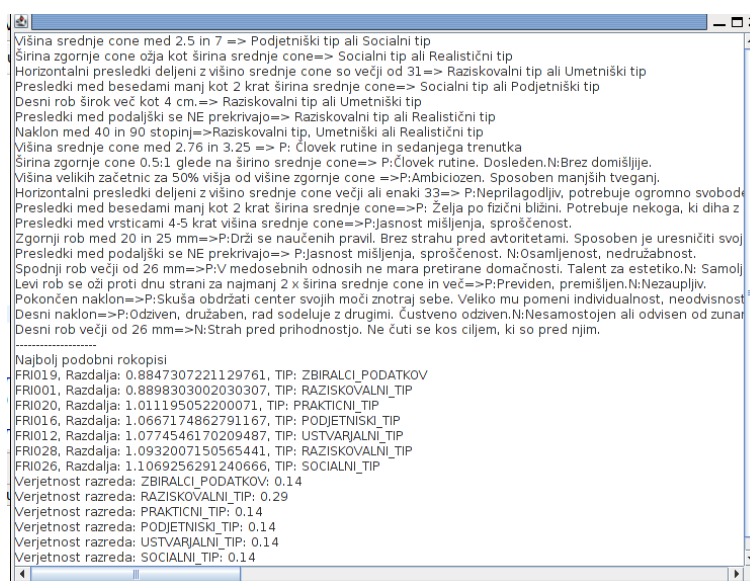
Tip: Nedoločen Raziskovalni tip Praktični tip Zbiralci podatk... Ustvarjalni tip Podjetniški tip

Slika 4.18: Prikaz rezultatov

4.6.4 Prikaz in urejanje rezultatov

V kolikor želimo videti rezultate za opravljene meritve, v meniju pod *Prikaži* označimo opcijsko polje *Rezultati*. Pojavi se okno z rezultati, kot ga prikazuje slika 4.18.

Sistem omogoča tudi neposredno urejanje rezultatov. Če želimo urejati posamezen rezultat, označimo izbrano opcijsko polje *Uredi* in vnesemo ustrezno vrednost. Sistem naredi validacijo vnosa in ob neustrezno vneseni vrednosti zahteva ponovitev postopka. Ustrezno vnesena vrednost ni rezultat opravljenih meritev na rokopisu. Torej sistem ne prikazuje rezultatov opravljenih meritev (na naloženem rokopisu), za katere se je grafolog odločil, da jih bo vnesel ročno. Metode strojnega učenja upoštevajo rezultate, ki so prikazani v oknu *Rezultati* ne glede na to, ali jih je vnesel grafolog ali so rezultat opravljenih meritev.



Slika 4.19: Prikaz dejstev

4.6.5 Prikaz dejstev

Rezultati metod odločitvena pravila in kNN so neposredno dosegljivi, če v meniju pod *Prikaži* označimo opcijsko polje *Dejstva*. Pojavi se okno, kot ga prikazuje slika 4.19

Kot razvidno iz slike 4.19, se v okno najprej izpišejo izpolnjena pravila, za tem najbolj podobni rokopisi in na koncu verjetnostna distribucija po razredih, ki jo predlaga metoda kNN.

4.6.6 Prikaz opomb

Če v meniju pod *Prikaži* označimo opcijsko polje *Opombe*, sistem grafologu omogoča zapisovanje poljubnega teksta. Tekst je lahko opomba, analiza, končna ocena ali kaj drugega. Pri shranjevanju opravljenega dela se, če v meniju pod *Datoteka* izberemo možnost *Shrani*, shrani tudi zapisan tekst, ki je tako dosegljiv ob ponovnem nalaganju obravnavanega rokopisa.

4.6.7 Dodajanje in urejanje figure

Po opravljeni meritvi figure, kot smo to opisali v razdelku 4.6.2, shranimo opravljeno meritvev tako, da v meniju pod *Datoteka* izberemo možnost *Shrani*. Sistem shrani meritve v datoteko z imenom slike, podaljškom *figure* in končnico *csv*. Opravljene meritve na sliki z imenom *slika* sistem shrani kot podatkovno datoteko z imenom *slika_figure.csv*. Podatkovna datoteka se shrani v mapo *Data*. Sistem omogoča tudi ponovno nalaganje in urejanje meritev, opravljenih na figuri. V ta namen postopamo enako kot pri novih meritvah na figuri. Najprej naložimo sliko s figuro nato urejamo in dodajamo opravljene meritve. Meritve na figuri so ločene od ostalih meritev rokopisov. Če želimo posameznemu rokopisu dodati opravljene meritve figure, moramo v meniju pod *Drugo* → *Dodaj figuro* izbrati ustrezno podatkovno datoteko, ki smo jo pred tem shranili. Na ta način smo rezultatom, ki so odraz opravljenih meritev na rokopisu, dodali rezultate opravljenih meritev na figuri.

4.6.8 Nastavitve in odločitvena pravila

Nastavitveno datoteko *system.ini* in datoteko z odločitvenimi pravili *rules.conf* prvič omenjamo v razdelku 4.5. Nastavitvena datoteka v trenutni različici sistema vsebuje tri parametre. S parametrom *k* povemo, koliko najbližjih sosedov upošteva metoda *kNN* pri uvrščanju novih primerov. Naslednja dva parametra *ANGLE_EPS* in *HEIGHT_EPS* uporablja sistem pri določanju oblike osnovne linije. Parametra se uporabljata, kot je opisano v razdelku 4.6.3. Pri tem se parameter *ANGLE_EPS* uporabi za določanje, če je osnovna linija ravna, konkavna ali konveksna. S pomočjo parametra *HEIGHT_EPS* sistem ugotavlja, če osnovna linija narašča, pada oziroma ostaja ravna.

Slika 4.20 prikazuje del vsebine datoteke *rules.conf*, kjer vrstica s posameznim pravilom vsebuje več parametrov: zaporedna številka pravila, opis, pogojni del pravila, sklepni del pravila in seznam oznak atributov (ločenih z vejico), ki jih uporabljamo v pravilu. Pri zapisu pravila lahko uporabljamo

```
#zaporedna številka; opis; pravilo; posledica ob izpolnjenem pogoju; seznam atributov v pravilu ločenih z vejico
1;Rule 1;VSCme < 1;Višina srednje cone < 1 => Raziskovalni tip;VSCme
2;Rule 2;(VSCme >= 1) && (VSCme<=2.5);Višina srednje cone med 1 in 2.5 => Raziskovalni tip ali Realistični tip;VSCme
3;Rule 3;(VSCme > 2.5) && (VSCme<=7);Višina srednje cone med 2.5 in 7 => Podjetniški tip ali Socialni tip;VSCme
4;Rule 4;Math.abs(SSCme - VSCme)< 0.2;Širina srednje cone enaka kot višina=> Podjetniški tip ali Socialni tip;VSCme,SSCme
5;Rule 5; SSCme <= (VSCme/2) ;Širina srednje cone za pol ali več manjša od višine=> Realistični (praktični tip) ali Socialni tip;VSCme,SSCme
6;Rule 6;Math.abs(VZCme - VSCme) < 0.2;Višina zgornje cone enaka višini srednje cone=> Socialni tip;VSCme,VZCme
7;Rule 7;(VZCme - 2*VSCme) > 0.0;Višina zgornje cone je 2 krat višja od višine srednje cone=> Umetniški tip;VZCme,VSCme
8;Rule 8;((SZCme - SSCme) < 0) && (Math.abs(SZCme-SSCme)>=0.2);Širina zgornje cone ožja kot širina srednje cone=> Socialni tip ali Realistični
```

Slika 4.20: Prikaz dela vsebine datoteke *rules.conf*

poljubne logične operatorje iz algebre *Bool*. Za notacijo operatorjev smo uporabili notacijo, ki je del jezika *Java*.

4.7 Rezultati opravljenih meritev

Točnost sistema pri opravljanju meritev smo preverili na 10-tih rokopisih. Primerjali smo rezultate meritev z rezultati, ki smo jih dobili od grafologa. Za ocenjevanje napak smo uporabili spremenjene enačbe za relativno in absolutno napako. Absolutno napako za vsako skupino meritev predstavlja enačba (4.13). Relativna napaka je predstavljena z enačbo (4.14)

$$\Delta = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (4.13)$$

$$\delta = \frac{\Delta}{\frac{1}{n} \sum_{i=1}^n y_i} \quad (4.14)$$

Omeniti je potrebno, da pri primerjavi nismo upoštevali diskretnih atributov. Slednji predstavljajo meritve, ki so rezultat uporabnikove odločitve in tako ne izhajajo neposredno iz sistema. Če bi vključil diskretne attribute, bi primerjali uporabnikove odločitve z odločitvami strokovnjaka grafologa, kar pa ni cilj pri ocenjevanju točnosti sistema. Da bi zmanjšali vpliv, ki ga imajo ustrezne izkušnje in strokovnost pri izvajanju meritev, smo za vsako meritev upoštevali le najbolj pogosto in povprečno vrednost (v kolikor sta prisotni). Tabela 3.1 prikazuje rezultate primerjave. Pri tem simbol Δ_{me} in δ_{me} pomenita absolutno in relativno napako za najbolj pogosto vrednost, simbol

Meritev	Δ_{me}	δ_{me}	Δ_{avg}	δ_{avg}
Višina srednje cone	0.78	0.36	0.51	0.21
Širina srednje cone	1.07	0.58	0.15	0.07
Višina ovalov	0.50	0.23	0.24	0.11
Širina ovalov	0.77	0.30	0.64	0.32
Višina zgornje cone	1.13	0.33	0.89	0.28
Širina zgornje cone	0.92	0.61	0.96	0.65
Dolžina spodnje cone	1.32	0.27	0.62	0.13
Širina spodnje cone	1.80	1.32	1.38	0.78
Širina velikih začetnic	5.53	1.00	1.59	0.30
Višina velikih začetnic	2.10	0.28	1.05	0.16
Višina diakrtičnih znakov	0.40	0.29	0.37	0.28
Presledki med besedami	1.65	0.25	1.44	0.23
Presledki med črkami	0.24	0.24	0.22	0.17
Presledki med vrsticami	2.84	0.47	1.44	0.22
Naklon	16.39	0.23		
Duktus	0.10	0.35		

Tabela 4.1: Tabela napak pri opravljanju meritev

Δ_{avg} in δ_{avg} pa absolutno in relativno napako za povprečno vrednost trenutno obravnavane meritve i . Vse absolutne napake so v mm , razen meritve naklon.

Opazimo velika odstopanja, ponekod tudi do več kot 50% od originalne meritve. Pri opravljanju meritev smo postopali enako kot grafolog. Meritve smo opravili v prvi vrstici, na sredini in v zadnji vrstici. Grafolog nato poišče predele rokopisa, za katere meni, da njihove meritve vplivajo na končno oceno. Izbira pomembnih predelov rokopisa je pogojena z znanjem in izkušnjami, vendar sodi v subjektivni del analize. Pomembni predeli rokopisa predstavljajo odstopanja in neposredno vplivajo na najbolj pogosto in povprečno vrednost, ki sta odvisna od števila in izbranih meritev. Najtež-

tnejši razlog za odstopanje meritev vidimo torej v pomanjkanju strokovnega znanja in izkušenj. Še en razlog izhaja neposredno iz delovanja sistema. Da bi zagotovili ustrezno natančnost, smo sistem zasnovali tako, da večino meritev zaokrožuje na eno decimalno natančno. Originalni podatki niso tako natančni. Na primer, meritev *višina srednje cone* in pripadajoča najpogostejša vrednost v 10 rokopisih nima decimalnega mesta.

4.8 Rezultati napovedovanja

V sklopu razvoja sistema smo vključili tudi sposobnost napovedovanja. Kot navajamo v razdelku 3.3, smo v sistem vključili *odločitvena pravila* in metodo *kNN*. S pomočjo grafologa smo pridobili in definirali skupaj 88 odločitvenih pravil. Razvrstili smo jih v dve skupini. V prvo skupino spada 25 pravil, kjer sklepni del napoveduje poklicni tip osebnosti. V drugi skupini so ostala pravila, kjer je sklepni del podan opisno, torej opisuje osebnostno značilnost. Ker nam ni uspelo pridobiti ustreznih pravil, ki bi sklepale nedvoumno na pripadajočo lastnost, smo v sklepnem delu pravil druge skupine podali tako negativen kot pozitiven pomen izpolnjenega pravila. V Dodatku 2 vidimo interpretacijo rokopisa, kot ga je naredil grafolog, Dodatek 3 pa prikazuje dejstva, ki jih je napovedal sistem. Opazimo, da so napovedi splošne. Tako odločitvena pravila prve skupine kot druge podajo množico opisov, iz katerih ne moremo neposredno sklepati na osebnostne značilnosti niti na poklicni tip osebnosti. Zaradi majhne velikosti učne množice (30 rokopisov) in 138 atributov ter relativno nizke zastopanosti nekaterih poklicnih tipov v učni množici (tudi do 2 krat) metode *kNN* in točnosti uvrščanja nismo podrobno obravnavali.

Dvoumnost napovedi, ki jih vrnejo odločitvena pravila, pripisujemo njihovi splošnosti. Predvidevamo, da bi z ustrezno kombinacijo parametrov in nedvoumno definicijo sklepnega dela napovedi postale natančnejše. Prav tako predvidevamo, da se bo klasifikacijska točnost metode *kNN* višala s povečevanjem velikosti učne množice.

Poglavje 5

Sklepne ugotovitve

Sistem smo načrtovali z namenom popolne avtonomnosti, ki smo jo opredelili v razdelku 4.1. Zaradi težavnosti problemov *segmentacije* in *razpoznavanja*, smo se pri implementaciji omejili le na *merjenje* in *napovedovanje*. Implementacijo *segmentacije* in *razpoznavanja* prepustimo nadaljnjemu delu. Točnost sistema pri opravljanju meritev, ki pokriva področje *merjenje* prikazuje tabela 3.1. Menimo, da bi rezultati opravljenih meritev odstopali v manjši meri oziroma sploh nebi odstopali, če bi jih opravljali strokovno usposobljeni grafologi. Ne glede na rezultate primerjave meritev trdimo, da nam je uspelo razviti sistem, ki nadomesti standardna orodja, ki jih grafolog uporablja pri svojem delu. Napovedovanje sistema ni zadovoljivo. Razloge za neuspeh vidimo v velikosti učne množice in splošnosti odločitvenih pravil. Menimo, da bi se metodi *kNN* z zadostno velikostjo učne množice, točnost uvrščanja poklicnih tipov, bistveno izboljšala. Prav tako trdimo, da bi s kompleksnejšo vendar s strokovnega stališča še vedno smiselno kombinacijo atributov v pogojnem delu odločitvenih pravil in posledično poenostavitvijo sklepnega dela, odločitvena pravila postala učinkovito orodje tako pri napovedovanju poklicnih tipov kot tudi osebnostih značilnosti. Področji *segmentacije* in *razpoznavanja* zaradi tehnološkega napredka v prihodnosti obetata veliko. Priložnost vidimo v razširjenosti mobilnih telefon z zaslonom občutljivim na dotik in odprtosti trga mobilnih aplikacij. V razdelku 4.1 smo

omenili programsko opremo, ki se uči lastnih karakteristik posameznikove pisave in na ta način izboljša razpoznavanje. Nismo zasledili učinkovitosti omenjenega učenja, vendar predvidevamo, da ustrezen izbor značilk, ki sestavljajo lastno karakteristiko pisave posameznika, razpoznavanje bistveno izboljša. Razširjenost mobilnih telefonov in aplikacija, ki se uči karakteristik posameznikove pisave, lahko vodi do ogromnih baz podatkov, s pomočjo katerih bi izboljšali oziroma razvili nove metode za uspešno rešitev problema razpoznavanja. Prednost sistema vidimo v shranjevanju podatkov in posledično širjenju učnih primerov. Prav tako mislimo, da je sistem primeren kot orodje za pomoč pri učenju grafološke analize. Predavatelj lahko interaktivno demonstrira opravljanje meritev, jih shrani in kasneje prikliče. Z ustrezno množico odločitvenih pravil lahko predavatelj definira pomen posameznih meritev, katerega potem študenti spoznavaajo z opravljanjem meritev. Zaključimo lahko, da nam je uspelo razviti sistem, ki ga grafologi lahko uporabljajo pri svojem delu in poučevanju.

Literatura

- [1] K. Amend, M. S. Ruiz, *Handwriting Analysis*, Franklin Lakes, NJ: The Career Prees, 1980, str. 3-5.
- [2] P. W. Frey, D. J. Slate, "Letter Recognition Using Holland-Style Adaptive Classifiers", *Machine Learning*, št. 6, str. 166-182, 1991.
- [3] I. Kononenko, M. R. Šikonja, *Inteligentni sistemi*, Ljubljana: Založba FE in FRI, 2010.
- [4] S. Lowe, "Michon and The Birth of Scientific Graphology", *Vanguard graphological journal*, 2007
- [5] M. Metelko, *Grafologija II - Ovrednotene lastnosti pisca rokopisa, 4. dopolnjena izd.*, Ljubljana: Evropska šola grafologije, GE-Global Elite d.o.o., 2010.
- [6] S. H. Ow, K. S. Teh, "An Overview of the Handwriting Analysis Systems in the Market", *CMU. Journal*, št. 4(2), str. 231-238, 2005
- [7] N. Potrebuješ, B. Tvrđi, A. Valmarska, "Graphology analysis through machine learning", Seminarška naloga, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, 2010.
- [8] S. Prasad, V. K. Singh, A. Sapre, "Handwriting Analysis based on Segmentation Method for Prediction of Human Personality using Support Vector Machine", *International Journal of Computer Applications (0975 – 8887)*, št. 12, zv. 8, 2010.

-
- [9] G. Sheikholeslami, S. N. Srihari, V. Govindaraju, Computer Aided Graphology. Dostopno na:
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.46.8214&rep=rep1&type=pdf>
- [10] C. Y. Suen, M. Berthod, S. Mori, "Automatic Recognition of Handprinted Characters-The State of the Art", *Proceedings Of The IEEE*, št. 68, zv. 4, 1980.
- [11] A. Trstenjak, *Človek in njegova pisava*, Ljubljana: Zavod za alternativno izobraževanje-GESD, 2006, str. 5-55
- [12] (2012) *Codehause*. Dostopno na:
<http://mvel.codehaus.org/>
- [13] (2012) *Graphology*. Dostopno na:
<http://en.wikipedia.org/wiki/Graphology>
- [14] (2012) *Graphoanalysis*. Dostopno na:
<http://en.wikipedia.org/wiki/Graphoanalysis>
- [15] (2012) *Handwriting recognition*. Dostopno na:
http://en.wikipedia.org/wiki/Handwriting_recognition
- [16] (2012) *Learning classifier system*. Dostopno na:
http://en.wikipedia.org/wiki/Learning_classifier_system

Slike

2.1	Meritev <i>višina srednje cone</i>	6
2.2	Meritev <i>višina zgornje cone</i>	6
2.3	Meritev <i>dolžina spodnje cone</i>	6
2.4	Meritev <i>višina diakritičnih znakov</i>	7
2.5	Meritev <i>presledki med besedami</i>	7
2.6	Meritev <i>presledki med črkami</i>	8
2.7	Meritev <i>presledki med podaljški</i>	8
2.8	Meritev <i>20 horizontalnih presledkov</i>	8
2.9	Meritev <i>dolžina črtice na t</i>	9
2.10	Meritev <i>višina črtice na t</i>	9
3.1	Distribucija razredov	14
4.1	Prikaz primerov generiranih črk	20
4.2	Proces razpoznavanja	20
4.3	Beseda <i>tudi</i> v prvem rokopisu	21
4.4	Beseda <i>tudi</i> v drugem rokopisu	21
4.5	Funkcionalnost sistema	23
4.6	Datotečna struktura programa	28
4.7	Osnovno okno	29
4.8	Prikaz elementov menija <i>Datoteka</i>	30
4.9	Prikaz elementov menija <i>Meritve</i>	30
4.10	Prikaz elementov menija <i>Prikaži</i>	30
4.11	Meritev črke	31

4.12	Dodajanje meritve črke	31
4.13	Merjenje razdalje	32
4.14	Meritev poravnava osnovne linije	33
4.15	Meritev 20 horizontalnih presledkov	33
4.16	Meritev naklon	34
4.17	Meritev širina duktusa	34
4.18	Prikaz rezultatov	37
4.19	Prikaz dejstev	38
4.20	Prikaz dela vsebine datoteke <i>rules.conf</i>	40
5.1	Tabela meritev	52
5.2	Primer rokopisa	58

Tabele

2.1	Prikaz meritev, ki se opravljajo pri posameznih črkah	11
3.1	Klasifikacijska točnost izbranih metod	15
4.1	Tabela napak pri opravljanju meritev	41

Dodatek 1

Referenčna tabela meritev

MERITVE	Najbolj pogosta meritev	Najnižja vrednost	Najvišja vrednost	Povprečna vrednost	Nihanje
Višina srednje cone:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Višina srednje cone je nepravilna:	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %
Širina srednje cone:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Višina ovalov:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Višina ovalov je nepravilna:	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %
Širina ovalov:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Širina ovalov je nepravilna:	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %
Višina zgornje cone:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Višina zgornje cone je nepravilna:	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %
Širina zgornje cone:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Dolžina spodnje cone:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Širina spodnje cone:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Višina osebnega zaimka:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Širina velikih začetnic:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Višina velikih začetnic:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Horizontalni presledki:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Višina diakritičnih znakov:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Presledki med besedami:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Presledki med besedami nepravilni:	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %	<input type="text"/> %
Presledki med črkami:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Presledki med podajški:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Presledki med vrsticami:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Rob - zgornji:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Rob - zgornji (širjenje/ožanje):	<input type="text"/> mm na začetku	<input type="text"/> mm na koncu	<input type="checkbox"/> menjajoč		
Rob - spodnji:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Rob - spodnji (širjenje/ožanje):	<input type="text"/> mm na začetku	<input type="text"/> mm na koncu	<input type="checkbox"/> menjajoč		
Rob - levi:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Rob - levi (širjenje/ožanje):	<input type="text"/> mm na začetku	<input type="text"/> mm na koncu	<input type="checkbox"/> menjajoč	<input type="checkbox"/> odstavki	
Rob - desni:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Rob - desni (širjenje/ožanje):	<input type="text"/> mm na začetku	<input type="text"/> mm na koncu	<input type="checkbox"/> menjajoč		
Robovi - vsi 4:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Poravnava osnovne linije:	<input type="text"/> mm na začetku	<input type="text"/> mm na sredini	<input type="text"/> mm na koncu	<input type="checkbox"/> menjajoča	
Naklon:	<input type="text"/> °	<input type="text"/> °	<input type="text"/> °	<input type="text"/> °	<input type="checkbox"/> menjajoč
Širina duktusa:	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm	<input type="text"/> mm
Razmerje črnilo : belina:	<input type="text"/> % črnila	<input type="text"/> % črnila	<input type="text"/> % črnila	<input type="text"/> % črnila	<input type="text"/> % črnila

Slika 5.1: Tabela meritev

Dodatek 2

Končna ocena grafologa

Primer končne ocene grafološke analize za rokopis prikazan v Dodatku 4, ki jo je opravil strokovnjak grafolog:

- *tehnični tip*
- *fizična aktivnost*
- *stabilnost*
- *samoinicativnost*
- *samostojnost*
- *osamljenost*
- *materialni svet prevladuje nad mentalnim*
- *iskrenost*
- *organizacijske sposobnosti*
- *intuicija*
- *neučakanost*

Dodatek 3

Končna ocena sistema

Primer analize osebnosti za rokopis prikazan v Dodatku 4, ki jo je podal sistem na podlagi opravljenih meritev:

Višina srednje cone med 2.5 in 7 => Podjetniški tip ali Socialni tip

Širina zgornje cone ožja kot širina srednje cone => Socialni tip ali Realistični tip

Horizontalni presledki deljeni z višino srednje cone so večji od 31 => Raziskovalni tip ali Umetniški tip

Presledki med besedami manj kot 2 krat širina srednje cone => Socialni tip ali Podjetniški tip

Desni rob širok več kot 4 cm. => Raziskovalni tip ali Umetniški tip

Presledki med podaljški se NE prekrivajo => Raziskovalni tip ali Realistični tip

Naklon med 40 in 90 stopinj => Raziskovalni tip, Umetniški ali Realistični tip

Višina srednje cone med 2.76 in 3.25 => P: Človek rutine in sedanjega trenutka

Širina zgornje cone 0.5:1 glede na širino srednje cone => P: Človek rutine. Dosleden. N: Brez domišljije.

Višina velikih začetnic za 50% višja od višine zgornje cone => P: Ambiciozen. Sposoben manjših tveganj.

Horizontalni presledki deljeni z višino srednje cone večji ali enaki 33 => P: Neprilagodljiv, potrebuje ogromno svobode.

Presledki med besedami manj kot 2 krat širina srednje cone => P: Želja po fizični bližini. Potrebuje nekoga, ki diha z njim. Primeren za timsko delo. Svoja stališča oblikuje glede na podlagi tujih sugestij. N: Primanjkuje mu kritičnosti do sebe in/ali soljudi. Subjektivni spomin. Strah pred osamljenostjo. Boječ. Skopušтво.

Presledki med vrsticami 4-5 krat višina srednje cone => P: Jasnost mišljenja, sproščenost.

Zgornji rob med 20 in 25 mm => P: Drži se naučenih pravil. Brez strahu pred avtoritetami. Sposoben je uresničiti svoje ambicije.

Presledki med podaljški se NE prekrivajo =>P:Jasnost mišljenja, sproščenost. N:Osamljenost, nedružabnost.

Spodnji rob večji od 26 mm =>P:V medosebnih odnosih ne mara pretirane domačnosti. Talent za estetiko.N: Samoljubnost.

Levi rob se oži proti dnu strani za najmanj 2 x širina srednje cone in več =>P:Previden, premišljen.N:Nezaupljiv.

Pokončen naklon =>P:Skuša obdržati center svojih moči znotraj sebe. Veliko mu pomeni individualnost, neodvisnost, vedno izhaja iz sebe. Ni dozeten za tuje vpliva, saj zaupa v svoje lastne sposobnosti. Notranji in zunanji nadzor. Nagnjen k pretiranemu iskanju in dokazovanju resnice. N:Skuša nadzorovati svoja čustva z razumom. Če bi pokazal čustva, bi postal ranljiv, zato jih skriva.

Desni naklon =>P:Odziven, družaben, rad sodeluje z drugimi. Čustveno odziven.N:Nesamostojen ali odvisen od zunanjega sveta.

Desni rob večji od 26 mm =>N:Strah pred prihodnostjo. Ne čuti se kos ciljem, ki so pred njim.

Najbolj podobni rokopisi:

FRI019, Razdalja: 0.8847307221129761, TIP: ZBIRALCI_PODATKOV

FRI001, Razdalja: 0.8898303002030307, TIP: RAZISKOVALNI_TIP

FRI020, Razdalja: 1.011195052200071, TIP: PRAKTICNI_TIP

FRI016, Razdalja: 1.0667174862791167, TIP: PODJETNISKI_TIP

FRI012, Razdalja: 1.0774546170209487, TIP: USTVARJALNI_TIP

FRI028, Razdalja: 1.0932007150565441, TIP: RAZISKOVALNI_TIP

FRI026, Razdalja: 1.1069256291240666, TIP: SOCIALNI_TIP

Verjetnost razreda ZBIRALCI_PODATKOV: 0.14

Verjetnost razreda RAZISKOVALNI_TIP: 0.29

Verjetnost razreda PRAKTICNI_TIP: 0.14

Verjetnost razreda PODJETNISKI_TIP: 0.14

Verjetnost razreda USTVARJALNI_TIP: 0.14

Verjetnost razreda SOCIALNI_TIP: 0.14

Velika črka *P* v sklepnem delu odločitvenega pravila pomeni pozitivno lastnost, velika črka *N* pa negativno.

Dodatek 4

Primer rokopisa

Moje potovanje na prvi torki
mislim, da bo zelo lepo in se
zelo veselim. Vsaka nedelja
hadim za mojim timu in
svaki mo kosilo. Zelo sem veselo
kjer suhi pride moja družina
in moja družina

Slika 5.2: Primer rokopisa