

Oznaka poročila: ARRS-RPROJ-ZP-2013/150



ZAKLJUČNO POROČILO RAZISKOVALNEGA PROJEKTA

A. PODATKI O RAZISKOVALNEM PROJEKTU

1. Osnovni podatki o raziskovalnem projektu

Šifra projekta	J2-2194
Naslov projekta	Kvalitativno modeliranje na osnovi podatkov
Vodja projekta	16324 Janez Demšar
Tip projekta	J Temeljni projekt
Obseg raziskovalnih ur	5310
Cenovni razred	A
Trajanje projekta	05.2009 - 04.2012
Nosilna raziskovalna organizacija	1539 Univerza v Ljubljani, Fakulteta za računalništvo in informatiko
Raziskovalne organizacije - soizvajalke	
Raziskovalno področje po šifrantu ARRS	2 TEHNIKA 2.07 Računalništvo in informatika 2.07.07 Inteligentni sistemi - programska oprema
Družbeno-ekonomski cilj	13.02 Tehnološke vede - RiR financiran iz drugih virov (ne iz SUF)

2. Raziskovalno področje po šifrantu FOS¹

Šifra	1.02
- Veda	1 Naravoslovne vede
- Področje	1.02 Računalništvo in informatika

B. REZULTATI IN DOSEŽKI RAZISKOVALNEGA PROJEKTA

3. Povzetek raziskovalnega projekta²

SLO

Kvalitativni modeli so modeli, ki, za razliko od klasifikacijskih in regresijskih, opisujejo kvalitativne povezave, kot npr. $y = Q(+x, -z)$ if $t > u$, "y narašča z x in pada z z, če je t večji od u". Po tem, ko so jih začeli pred več kot 60 leti uporabljati v ekonomiji, so postali popularni in uporabljeni v mnogih področjih znanosti in tehnologije ter industrijskih panogah. Čeprav v načelu ne morejo doseči točnosti, ki jo dosegajo numerični modeli, so v praksi navadno

robustnejši ter enostavnejši za razumevanje in uporabo.

Presenetljivo je, da obstaja le nekaj metod za avtomatsko sestavljanje tovrstnih modelov. Namen projekta je razviti praktično uporabne metode za ta namen.

Delo smo začeli z analizo obstoječih definicij kvalitativnih povezav. Čeprav je to delo videti preprosto, je predstavljalo precejšen problem v dosedanjih raziskavah s tega področja. Med različnimi potencialnimi definicijami smo določili takšno, ki temelji na parcialnem diferencialnem odvodu. To se je izkazalo za dobro odločitev, saj je olajšalo razvoj ostalih metod na temeljih matematike in statistike.

Srž projekta je bil razvoj metod za računanje kvalitativnih povezav - kvalitativnih parcialnih odvodov. Razvili smo precej metod in njihove lastnosti preučili tako teoretično kot empirično, s poskusi na sintetičnih podatkih kot podatkih iz praktičnih problemov. Posebej smo opazovali tudi robustnost glede na nastavitve parametrov učenja ter vpliv koreliranih in diskretnih atributov.

Družina razvitih algoritmov je na voljo pod skupnim imenom Pade in je dokumentirana v dveh člankih, ki smo ju objavili v zadnjem letu izvajanja projekta. Z izjemo mejnih primerov, ki so bili konstruirani posebej za preskušanje določene metode, se je v splošnem izkazala za najhitrejšo in najrobustnejšo metoda, ki temelji na primerno uteženi linearni regresiji izračunani na ustreznem podprostoru v okolici opazovane točke.

Poseben problem je predstavljalo ocenjevanje dobljenih modelov. Področje kvalitativnega modeliranja trenutno nima splošno uveljavljene mere kvalitete, zato smo jo definirali v okviru projekta. Izdelali smo tudi metode, s katerimi model poda oceno zanesljivosti svojih napovedi.

Pretvorba kvalitativnih modelov v kvantitativne se je izkazala za manjši problem, kot smo prvotno pričakovali, saj nam je uspelo z računanjem parcialnih odvodov v posameznih točkah pretvoriti začetni problem v klasičen problem strojnega učenja, ki ga lahko rešimo s poljubno standardno metodo brez posebnih prilagoditev.

Projekt je v celoti dosegel zastavljene cilje. Kot kaže delni seznam objav, ki kaže tudi uporabo metode v kontekstih, povsem drugačnih od izvirnega, so rezultati projekta raznoliki in odmevni.

ANG

Qualitative models are models which, for a contrast from classification and regression models describe qualitative relations like $y=Q(+x, -z)$, if $t > u$, "y increases with x and decreases with z if t is greater than u". First considered in economy more than 60 years ago, they became popular and are used in many areas of science, technology and industry. Although they in principle cannot be as exact as numerical models, they are in practice more robust and easier to understand and use.

It is thus rather surprising that there exist only a few methods for automatic construction of models of this kind. The purpose of the project is to provide practically useful methods for this task.

We began by cataloguing the existing definitions of qualitative relation. Although seemingly trivial, this task has proven quite difficult for some of the previous research in the field. Among different options, we have defined qualitative relations in terms of partial derivatives. This later proved to be a good decision as it allows us to develop other methods by utilizing the existing approaches from mathematics and, in particular, statistics.

The core of the project was development of methods for computation of qualitative partial derivatives. We developed a bunch of different methods and explored their properties both in theory and through empirical experiments on synthetic and real-world data. We also observed the method's robustness with regard to its parameters and the effect of correlated and discrete attributes.

The whole family of the developed algorithms is available under the common name Pade and documented in two papers published in the last year of the project. With exception of border cases that were constructed for testing particular properties of specific algorithms, the fastest and most robust method was the one based on weighted linear regression over appropriately chosen neighborhood of the observed reference point.

Assessing the quality of induced models represented a problem since the field currently has no

generally useful and accepted measure for this purpose. We thus defined some suitable methods within this project. We also developed algorithms that are capable of estimating the reliability of their predictions.

Turning qualitative models into quantitative ones was a much easier task than anticipated: calculating partial derivatives at individual points allowed us to transform the original problem into a standard machine learning problem, suitable for any popular machine learning method.

The project successfully reached its goals. As also evident from the partial list of publications, the method is also applicable in contexts that are quite different from the original one.

4. Poročilo o realizaciji predloženega programa dela na raziskovalnem projektu³

Delo smo začeli s preučevanjem preteklega dela na področju. Napor drugih raziskovalcev doslej so bili usmerjeni skoraj izključno v preučevanje uporabe kvalitativnih modelov, ki so bili narejeni ročno. Poleg tega je dokaj živo tudi področje gradnje regresijskih modelov, ki so zvesti kvalitativnim omejitvam, ki jih poda ekspert. Poleg metod, ki smo jih v preteklosti razvili v svoji skupini (QCGrid, QCSplines), smo pregledali tudi dela drugih avtorjev, med katerimi prevladujejo takšni, ki gradijo monotona regresijska drevesa. Z nekaterimi od njih smo navezali stike, ki so nam pomagali pri nadaljnjem delu na projektu.

Med različnimi možnimi definicijami kvalitativne povezave smo se odločili za definicijo, ki temelji na parcialnem odvodu. Razlog za izbor je, da je v njegovem ozadju konsistentna matematična teorija, poleg tega pa je v osnovi podoben kvalitativnim proporcionalnostim, kot jih definira Quiera. Alternativna definicija, ki jo uporablja program QUIN, je matematično pomanjkljiva in neenoznačna.

Načelni argumenti proti ad hoc definicijam in v prid uveljavljenim matematičnim pristopom so se izkazali za utemeljene tudi v praksi. Rezultat preteklih pristopov ni bil kvalitativni opis vedenja funkcije v posamičnih točkah temveč le v parih točk. Tako podane relacije so uporabne predvsem (ali izključno?) kot omejitve za postopke modeliranja, zato so morali raziskovalci te tehnike razviti tudi adaptacije klasičnih metod strojnega učenja, ki so dovoljevale vključevanje takšnih omejitev. V praksi se je to obneslo zgolj pri klasifikacijskih drevesih. Naš pristop pa vodi do opisa obnašanja funkcije v posamičnih točkah; s tem lahko izvirni problem prevedemo na klasični problem strojnega učenja, na katerem lahko uporabimo poljubno metodo učenja brez kakršnihkoli sprememb.

V prijavi projekta smo opisali tri prototipne metode, implementirane v okviru algoritma Pade. Za metodi, ki temeljita na triangulaciji se je, kot smo slutili, pokazalo, da v praksi nista posebej uporabni, ker sta omejeni na nizkodimenzionalne podatke. Tretjo metodo, tube regression, pa smo uporabili kot osnovo za razvoj treh izboljšanih metod, LWR, tauregression in parallel pairs. Prva temelji na običajni uteženi lokalni linearni regresiji, dve pa na precej modificirani obliki lokalne regresije. Razvite metode smo opisali v članku v ugledni reviji Artificial Intelligence Journal, starejše metode (obe triangulacijski in tube regression), ki so kljub pomanjkljivostim teoretično zanimive, pa smo predstavili na delavnici Learning Monotone Models from Data in članke o njih objavili v reviji Informatica.

Trenutni trend v strojnem učenju je, da se problemi praviloma prevajajo na optimizacijske probleme (po možnosti konveksne), kar sicer nudi učinkovite rešitve, vendar pogosto brez posebnega uvida v delovanje metode. Metode, ki smo jih predstavljali tu, pa temeljijo na kombinaciji postopkov iz topologije in linearne algebre, s čimer predstavljajo – po naši oceni in po oceni udeležencev srečanj, na katerih smo predstavljali razvite metode – zanimivo osvežitev na področju strojnega učenja.

Razvoj metod za kvalitativno modeliranje funkcije, podane z njeno vrednostjo na oblaku točk, predstavlja srž raziskovalnega projekta. Vzporedno s tem pa so kot "stranski produkti" nastajali številni drugi algoritmi.

Tako smo že v prvem letu izvajanja projekta razvili splošen algoritem za določanje optimalnega praga za razbijanje množice primerov v dve podmnožici, če kriterij za oceno praga ne temelji na meri, izračunani iz posameznih primerov (kot npr. entropija), temveč na meri, ki se izračuna iz parov primerov, kot to navadno delamo v kvalitativnem modeliranju. Članek o tem je bil objavljen v reviji Machine Learning.

Drug takšen algoritem je bil namenjen odvajanju diskretnih spremenljivk. Pri tem se je

pokazalo, da je ključni problem določanje metrike, ki definira "sosednost" primerov. Za obravnavo diskretnih atributov smo razvili novo metodo, s katero je mogoče računati (posplošene) odvode diskretnih spremenljivk. Osnovna ideja metode je, da namesto vrednosti diskretne spremenljivke opazujemo njeno verjetnost. Na ta način lahko odvod izračunamo tako, da pogledamo, kako sprememba vrednosti diskretne neodvisne spremenljivke vpliva na verjetnostno porazdelitev vrednosti odvisne spremenljivke. Rezultat takšnega odvajanja ni zvezen, temveč kvalitativen, "diskretni odvod", ki v bistvu predstavlja ureditev vrednosti diskretne spremenljivke. Ker definicija odvoda temelji na pogojnih verjetnostih, je za njegov izračun najprimernejši bayesovski pristop. V praktičnih problemih število primerov ne zadošča za izračun idealnega bayesovskega modela, zato smo se zatekli k naivnemu bayesovemu klasifikatorju, za katerega pa se pokaže, da ga je potrebno računati lokalno, pri čemer se lokalnost (v smislu tega, katere vrednosti ostalih spremenljivk morajo biti v okolici točke odvajanja konstantne) določa dinamično in sproti. Rezultate teh raziskav smo predstavili na srečanju Qualitative Reasoning 2010, kjer je metoda zaradi svoje izvirnosti požela precej zanimanja.

Iz tega dela smo izpeljali še eno zanimivo aplikacijo, to je, indukcijo modelov za napovedovanje preferenc na podlagi podatkov, ki preference vsebujejo le implicitno. Primeri takšnih podatkov najpogosteje nastopajo na področju obvladovanja stikov s strankami, kjer so znani podatki o uporabnikih in njihove preference glede izbire storitev. Preference lahko v duhu gornjih raziskav obravnavamo kot diskretno spremenljivko, katere vrednost želimo odvajati po lastnostih uporabnikov (ali obratno). Rezultat je model, ki pojasnjuje, kakšne vrste stranke so bolj nagnjene h katerim storitvam oz. kaj so atributi stranke, ki pojasnjujejo njeno izbiro. Rezultat smo objavili v prispevku in predstavili na delavnici Preference Learning v okviru konference ECML/PKDD.

Kot najboljši pristop kot ocenjevanju zanesljivosti napovedi so se izkazale ločene mere za umetne koncepte, v katerih poznamo pravilne napovedi in za resnične probleme, kjer pravilne napovedi navadno niso znane. V prvem primeru lahko uporabljamo klasične mere pravilnosti, kot jih poznamo v strojnem učenju in statistiki (klasifikacijska točnost oz. izguba 01, krivulje ROC ipd). V drugem primeru kot indikator točnosti uporabljamo stabilnost klasifikatorja. Leta je v osnovi podobna varianci klasifikatorja glede na spremembe v učni množici, varianca klasifikatorja pa, skupaj s kvadratom pristranskosti, tvori celotno napako. Ta, druga komponenta, pristranskost, je sicer nemerljiva, vendar je v sorazmerno visoko dimenzionalnih podatkih, s kakršnimi znajo delati predlagane metode (do nekaj deset dimenzij) navadno glavna težava redka pokritost prostora primerov in s tem povezana varianca, zato pristranskosti v takšnih primerih ne le v kvalitativnih, temveč tudi v klasifikacijskih kot v regresijskih problemih navadno posvečamo manj pozornosti.

V povezavi z metodami za ocenjevanje zanesljivosti posameznih napovedi smo se odzvali na povabilo farmacevtskega giganta Astra Zeneca, ki potrebuje takšne metode za ocenjevanje zanesljivosti napovedane toksičnosti načrtovanih učinkovin v zdravilih. S podjetjem smo se dogovorili za več projektov, med katerimi je tudi ocenjevanje zanesljivosti. V ta namen smo testirali obstoječe metode, za kar smo vključili tudi doc. Zorana Bosnića, ki sicer ni član projektne skupine, vendar smo ga financirali iz sredstev industrijskega projekta, in več do in podiplomskih študentov, predvsem mladih raziskovalcev, ki delajo v naši skupini.

Rezultate projekta smo uporabili tudi na več drugih praktičnih problemih, med katerimi velja posebej omeniti analizo podatkov o infekcijskih obolenjih pri starostnikih, ki so bili zbrani na ljubljanski infekcijski kliniki. Modeli, ki jih dobimo z uporabo algoritmov, razvitih v okviru projekta, so po oceni ekspertov smiselni in uporabni.

Rezultati projekta so bili predstavljeni v več publikacijah v revijah na indeksu SCI in INSPEC ter na več konferencah in delavnicah.

5. Ocena stopnje realizacije programa dela na raziskovalnem projektu in zastavljenih raziskovalnih ciljev⁴

V tem delu poročila sledimo opisu ciljev, kot so bili navedeni v prijavi projekta (točka 16, Upravljanje projekta: podroben načrt uresničevanja). Projekt je v vseh točkah dosegel zastavljene cilje.

1. Formalna definicija kvalitativne povezave

Kot navajamo v prejšnji točki poročila, smo preučili različne možne definicije kvalitativne

povezave. S teoretičnimi argumenti smo določili primerno povezavo, ki se je v kasnejšem izvajanju projekta pokazala kot zelo primerna.

2. Razvoj algoritmov za učenje kvalitativnih modelov

Ta del je predstavljal bistvo projekta. Razvili smo več algoritmov, primernih za različne kontekste.

2.1 Vpliv parametrov

Vpliv parametrov metod smo analizirali na sintetičnih in praktičnih podatkih. Različne metode so bile različno robustne; za najuspešnejšo se je, v splošnem, izkazala metoda, ki smo jo poimenovali tau-regresija.

2.2 Reševanje problema koreliranih atributov

Vpliv koreliranih atributov smo analizirali predvsem v kontekstu metod, ki temeljijo na linearni regresiji, saj so praktično uporabnejše. Pokazalo se je, da je metoda nanje le zmerno občutljiva, tako da zadošča že uporaba psevdo inverza oz. ridge regression.

2.3 Kvalitativno modeliranje časovnih zaporedij

Modeliranje časovnih zaporedij ni zahtevalo razvoja ločenih metod, saj jih lahko obravnavamo kar kot modeliranje enodimenzionalnih funkcij. Pri tem pa ne moremo uporabiti tau-regresije, ki po svoji naravi deluje le v več dimenzijah, temveč metode, ki izvirajo iz topologije in smo jih razvijali predvsem v začetku izvajanja projekta. Kot najzanimivejši pristop se je pokazala uporaba diskretne Morsejeve teorije.

2.4 Obravnava diskretnih atributov

Diskretne attribute smo obravnavali precej podrobneje, kot je bilo prvotno predvideno, saj smo med izvajanjem projekta dobili idejo za verjetnostno obravnavo diskretnih vrednosti. Detajli so opisani v prejšnji točki in v objavah.

3. Mere za ocenjevanje kvalitativnih modelov

Ugotovili smo, da so za empirično vrednotenje metod na sintetičnih podatkih primerne klasične mere iz strojnega učenja, za podatke, pri katerih ni znan "pravi" model, pa opazujemo stabilnost modelov, ki je, ob določenih predpostavkah, tudi indikator njihove točnosti.

4. Mere za ocenjevanje zanesljivosti posameznih napovedi

Razvili smo nekaj metod za ocenjevanje zanesljivosti.

5. Konstrukcija kvantitativnih modelov s pomočjo kvalitativnih

Ta del projekta se je izkazal za bistveno preprostejšega, kot je bilo predvideno. Razlogi so opisani v prejšnji točki poročila.

6. Razvoj pomožnih metod

Med izvajanjem projekta je bila razvita kopica "stranskih produktov", ki so prav tako opisani v prejšnji točki.

7. Ocenjevanje razvitih algoritmov in pomožnih metod

8. Objave razvitih metod

Projekt je bil zelo uspešen tudi v tem pogledu, saj je v njegovem okviru nastalo večje število odmevnih objav.

6. Utemeljitev morebitnih sprememb programa raziskovalnega projekta oziroma sprememb, povečanja ali zmanjšanja sestave projektne skupine⁵

Med potekom projekta je prihajalo zgolj do manjših sprememb projektne skupine v skladu z vsebinskimi potrebami projekta.

7. Najpomembnejši znanstveni rezultati projektne skupine⁶

	Znanstveni dosežek		
1.	COBISS ID	7550548	Vir: COBISS.SI
	Naslov	Algorithms for subsetting attribute values with Relief	

		ANG	Algorithms for subsetting attribute values with Relief
	Opis	SLO	Članek opisuje postopek, s katerim lahko poiščemo vrednost zveznega atributa, ki optimalno razdeli primere v dve podmnožici, pri čemer je optimalnost določena z mero, ki ne seštevata prek posamičnih primerov (kot npr. entropija), temveč prek parov primerov. Algoritem zmanjša kvadratno časovno kompleksnost v linearno. Za področje kvalitativnega modeliranja je postopek potreben, ker so "vektorji sprememb" (kot so definirani v, npr. literaturi o QUINu) definirani na parih primerov in ne na posamičnih primerih, kot je to običajno pri gradnji običajnih klasifikacijskih in regresijskih dreves.
		ANG	The paper presents an algorithm for finding the value of a discrete attribute which optimal splits a set of examples into two subsets, where the optimality is defined with regard to a criterion which is not computed over individual examples (like entropy) but on pairs of examples. The algorithm reduces the time complexity from quadratic (which we would get with a brute force algorithm) to linear. This is important for qualitative modelling where qualitative change vectors (as used in QUIN) are defined on pairs of examples and not on single examples.
	Objavljeno v		Kluwer Academic Publishers; Machine learning; 2010; Vol. 78, no. 3; str. 421-428; Impact Factor: 1.956; Srednja vrednost revije / Medium Category Impact Factor: 1.534; WoS: EP; Avtorji / Authors: Demšar Janez
	Tipologija		1.01 Izvirni znanstveni članek
2.	COBISS ID		7926100 Vir: COBISS.SI
	Naslov	SLO	Preference learning from qualitative partial derivatives
		ANG	Preference learning from qualitative partial derivatives
	Opis	SLO	V članku raziskujemo uporabo metod kvalitativnega modeliranja, ki smo jih razvili v okviru projekta, za sestavljanje preferenčnih modelov. Algoritem zna iz podatkov, ki preferenčne relacije vsebujejo zgolj implicitno, sestaviti model, ki zna za posamezni konkretni primer (npr. osebo) napovedati, katera izmed možnih alternativ ji bo všečnejša oz. urediti alternative po všečnosti.
		ANG	In the paper, we use one of the methods developed within the project, for construction of preference models. Using the data in which the preferences are given only implicitly, the method can construct a model that can predict the preferences for a concrete example (e.g., a person choosing between a set of alternatives).
	Objavljeno v		s. n.]; ECMLPKDD 2010; 2010; Str. 1-13; Avtorji / Authors: Žabkar Jure, Možina Martin, Janež Tadej, Bratko Ivan, Demšar Janez
	Tipologija		1.08 Objavljeni znanstveni prispevek na konferenci
3.	COBISS ID		8863572 Vir: COBISS.SI
	Naslov	SLO	Learning predictive qualitative models with Padé
		ANG	Learning predictive qualitative models with Padé
	Opis	SLO	Članek opisuje metode za gradnjo kvalitativnih modelov, ki temeljijo na topološki analizi prostora ter iskanju hiperravnin, ki se najbolj prilegajo podatkom, ne pa eksplicitno na linearni regresiji, za razliko od članka, ki smo ga objavili v reviji Artificial Intelligence Journal, ki temelji na različnih načinih uporabe linearne regresije.
		ANG	The paper describes methods for induction of qualitative models based on topological analysis and fitting hyper-planes to the data, instead of (explicitly) using linear regression, a method which is described in a separate paper in Artificial Intelligence Journal.
			Slovene Society Informatika; Informatica; 2011; Vol. 35, no. 4; str. 435-

	Objavljeno v		444; Avtorji / Authors: Žabkar Jure, Možina Martin, Bratko Ivan, Demšar Janez
	Tipologija		1.01 Izvirni znanstveni članek
4.	COBISS ID		7881812 Vir: COBISS.SI
	Naslov	SLO	Learning qualitative relations from categorical data
		ANG	Learning qualitative relations from categorical data
	Opis	SLO	Članek opisuje posplošitev metod odvajanja na kategorične domene. V ta namen smo definirali verjetnostne kvalitativne parcialne diferencialne odvode, ki povedo, kako sprememba vrednosti spremenljivke vpliva na verjetnost kategorične odvisne spremenljivke. Za izračun takšnih odvodov predlagamo metodo, ki temelji na lokalni različici naivnega bayesovskega modela.
		ANG	The paper describes a generalization of derivation to categorical domains. We defined probabilistic qualitative partial derivatives that describe how the change of the variable affects the probability of categorical dependent variable. We propose a local naive bayesian method for computation of such derivatives.
	Objavljeno v		s. n.]; QR 2010; 2010; Str. 105-111; Avtorji / Authors: Žabkar Jure, Možina Martin, Bratko Ivan, Demšar Janez
5.	COBISS ID		8324436 Vir: COBISS.SI
	Naslov	SLO	Learning qualitative models from numerical data
		ANG	Learning qualitative models from numerical data
	Opis	SLO	Članek opisuje novo metodo kvalitativnega učenja, ki ocenjuje parcialne odvode ciljne funkcije iz učnih primerov in jih uporablja za indukcijo kvalitativnega modela ciljne funkcije. Zasnovali smo tri metode izračuna odvodov, ki temeljijo na linearni regresiji, izračunani v lokalni okolici referenčne točke. Metode smo testirali na sintetičnih in resničnih podatkih. Članek opisuje tudi študijo na resničnem primeru, ki kaže, kako se metoda uporabi v praksi.
		ANG	The paper describes Pade, the crucial method developed within the project. Pade is a new method for qualitative learning which estimates partial derivatives of the target function from training data and uses them to induce qualitative models of the target function. We formulated three methods for computation of derivatives, all based on using linear regression on local neighbourhoods. The methods were empirically tested on artificial and realworld data. We also provide a case study which shows how the developed methods can be used in practice.
	Objavljeno v		Elsevier; Artificial intelligence; 2011; Vol. 175, no. 9/10; str. 1604-1619; Impact Factor: 2.252; Srednja vrednost revije / Medium Category Impact Factor: 1.427; A': 1; WoS: EP; Avtorji / Authors: Žabkar Jure, Možina Martin, Bratko Ivan, Demšar Janez
	Tipologija		1.01 Izvirni znanstveni članek

8. Najpomembnejši družbeno-ekonomski rezultati projektne skupine²

Družbeno-ekonomski dosežek			
1.	COBISS ID		Vir: vpis v poročilo
	Naslov	SLO	Gostovanje na Univerzi v Kyotu in inštitutu CHORDJ v Nari, Japonska
		ANG	Guest lecturer at University of Kyoto and institute CHORDJ, Nara, Japan

	Opis	SLO	Vodja projekta, dr. Janez Demšar, je v poletju 2010 tri mesece gostoval SLO na ugledni Univerzi v Kyotu, kjer je predaval vsebine s področja strojnega učenja. Poleg tega je v tem času raziskovalno sodeloval z inštitutom CHORDJ v Nari, na Japonskem.	
		ANG	PI of the project, Dr. Janez Demšar, was a guest lecturer for three months ANG in summer 2010 at Kyoto University, where he taught a short course on machine learning and data mining. In this time he also collaborated with research institute CHORDJ in Nara, Japan.	
	Šifra		B.05	Gostujoči profesor na inštitutu/univerzi
	Objavljeno v		(brez objave)	
	Tipologija		3.14	Predavanje na tuji univerzi
2.	COBISS ID			Vir: vpis v poročilo
	Naslov	SLO	Član uredništva revije Informatica Medica Slovenica	
		ANG	Member of the editorial board of Informatica Medica Slovenica	
	Opis	SLO	Vodja projekta je od leta 2009 član uredniškega odbora revije Informatica Medica Slovenica	
		ANG	From 2009, PI of the project is a member of editorial board of Informatica Medica Slovenica.	
	Šifra		C.04	Uredništvo mednarodne revije
	Objavljeno v		(brez objave)	
	Tipologija		4.00	Sekundarno avtorstvo
3.	COBISS ID			Vir: vpis v poročilo
	Naslov	SLO	Vzpostavitev sodelovanja s podjetjem Astra Zeneca	
		ANG	Establishing a collaboration with Astra Zeneca	
	Opis	SLO	Skupina je navezala aktivno sodelovanje s farmacevtskim gigantom Astra Zeneca, ki nas je povabil k skupnemu razvoju algoritmov in programske opreme. Delo delno sovpada tudi z vsebino projekta. Skupna vrednost doslej podpisanih pogodb znaša 140.000 evrov.	
		ANG	The group established active collaboration with pharmaceutical giant Astra Zeneca, which invited us to joint development of algorithms and their software implementation. The work partially overlaps with the proposed project. The collaboration so far resulted in grants in amount of 140.000 EUR.	
	Šifra		F.17	Prenos obstoječih tehnologij, znanj, metod in postopkov v prakso
	Objavljeno v		(brez objave)	
	Tipologija		3.25	Druga izvedena dela
4.	COBISS ID		8024660	Vir: COBISS.SI
	Naslov	SLO	Učenje kvalitativnih odvisnosti	
		ANG	Learning qualitative dependencies	
	Opis	SLO	Vodja projekta je bil somentor pri doktorski disertaciji Jureta Žabkarja. Disertacija je s področja kvalitativnega modeliranja in večina v njej predstavljenih metod je nastala v okviru tega projekta.	
		ANG	PI was a co-advisor of graduate student Jure Žabkar. His dissertation is from the field of qualitative modelling; most of the presented methods were developed within this project.	
	Šifra		D.09	Mentorstvo doktorandom

	Objavljeno v		[J. Žabkar]; 2010; X, 100 str.; Avtorji / Authors: Žabkar Jure	
	Tipologija		2.08 Doktorska disertacija	
5.	COBISS ID		Vir: vpis v poročilo	
	Naslov	SLO	Pridobitev projekta v okviru 7. okvirnega programa EU	
		ANG	Obtaining a project within the 7. EU framework	
	Opis	SLO	Na podlagi rezultatov tega in nekaj podobnih projektov smo bili ob koncu izvajanja projekta povabljeni v konzorcij, s katerim smo uspešno pridobili projekt v okviru 7. okvirnega programa EU (AXLE - Advanced Analytics for Extremely Large European Databases (318633), 2012–2015).	
		ANG	Based on results of this and several other projects, the group has been invited into a consortium that successfully proposed a project within the 7th EU FP (AXLE - Advanced Analytics for Extremely Large European Databases (318633), 2012–2015)	
	Šifra		D.01 Vodenje/koordiniranje (mednarodnih in domačih) projektov	
	Objavljeno v		Survey of activities, Faculty of Computer and Information Science, 2012	
	Tipologija		3.25 Druga izvedena dela	

9. Drugi pomembni rezultati projektne skupine⁸

--

10. Pomen raziskovalnih rezultatov projektne skupine⁹

10.1. Pomen za razvoj znanosti¹⁰

SLO

Razvite metode predstavljajo pionirsko delo na področju kvalitativnega modeliranja. Še posebej pa je delo zanimivo zaradi inovativnega povezovanja različnih področij - topologije, simboličnega računanja, strojnega učenja, numerične analize, verjetnosti in statistike.

V osnovi pa projekt sodi na področje umetne inteligence. Zaradi omejenega časa seveda nismo pričakovali, da bo kvalitativno modeliranje že v času izvajanja projekta doseglo enako stopnjo zrelosti kot učenje klasifikacijskih in regresijskih modelov, ki ju že pol stoletja razvija velika skupnost z vsega sveta. Menimo pa, da smo postavili dobre temelje za njegov nadaljnji razvoj.

Projekt je enako pomemben tudi za razvoj drugih vej znanosti, ki uporabljajo strojno učenje in odkrivanje znanja iz podatkov. To vključuje vse znanosti, ki sestavljajo hipoteze iz empiričnih podatkov, denimo moderno genetiko in medicino. Razviti algoritmi bodo uporabni, denimo, za analizo odvisnosti med geni v genetski mreži, kar lahko uporabimo za modeliranje in zdravljenje bolezni na nivoju genov. Drugi primeri znanosti, ki temeljijo na sklepanju iz podatkov, so socialne vede, psihologija in ekonomija, pa tudi večina drugih področij moderne znanosti.

ANG

Developed methods represent a pioneering work in the field of qualitative modeling. The work is especially interesting because of innovative connecting of techniques from different fields - topology, symbolic computation, machine learning, numerical analysis, probability and statistics.

The core of the project, however, belongs to the field of artificial intelligence. Due to the limited time and resources available we did not expect the project to develop qualitative modeling to the same level of maturity as that of classification and regression learning which have been developing for half a century by a large community. We however believe that we provided a good basis for its future development.

To project's results are useful in many other areas of science that rely on machine learning and data mining. These include all sciences that derive hypotheses from empirical data, most notably modern genetics and medicine. Developed algorithms will be, for instance, useful in analysis of dependencies between genes in genetic network, which can be used in modeling and curing diseases on genetic level. Other examples of scientific fields that rely heavily on drawing conclusions from experimental data are social sciences, psychology and economy and also most other areas of modern science.

10.2.Pomen za razvoj Slovenije¹¹

SLO

Umetna inteligenca je disciplina, v kateri je Slovenija vedno sodila v svetovni vrh glede na uveljavljenost njenih raziskovalcev in njihovih dosežkov, kot tudi z vidika pridobivanja raziskovalnih sredstev iz EU. Z opravljenim delom na doslej slabo raziskanem področju modeliranja bo Slovenija še naprej obdržala stik z vrhom.

ANG

The work of Slovenian researchers in AI has always represented the stateoftheart in the field and was also very successful with regard to obtaining EU-funded research grants. Our work in the unexplored area of qualitative modeling will help it to maintain this position.

11.Samo za aplikativne projekte in podoktorske projekte iz gospodarstva!

Označite, katerega od navedenih ciljev ste si zastavili pri projektu, katere konkretne rezultate ste dosegli in v kakšni meri so doseženi rezultati uporabljeni

Cilj		
F.01	Pridobitev novih praktičnih znanj, informacij in veščin	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.02	Pridobitev novih znanstvenih spoznanj	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.03	Večja usposobljenost raziskovalno-razvojnega osebja	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.04	Dvig tehnološke ravni	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.05	Sposobnost za začetek novega tehnološkega razvoja	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	

F.06	Razvoj novega izdelka	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.07	Izboljšanje obstoječega izdelka	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.08	Razvoj in izdelava prototipa	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.09	Razvoj novega tehnološkega procesa oz. tehnologije	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.10	Izboljšanje obstoječega tehnološkega procesa oz. tehnologije	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.11	Razvoj nove storitve	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.12	Izboljšanje obstoječe storitve	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.13	Razvoj novih proizvodnih metod in instrumentov oz. proizvodnih procesov	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.14	Izboljšanje obstoječih proizvodnih metod in instrumentov oz. proizvodnih procesov	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	

	Uporaba rezultatov	
F.15	Razvoj novega informacijskega sistema/podatkovnih baz	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.16	Izboljšanje obstoječega informacijskega sistema/podatkovnih baz	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.17	Prenos obstoječih tehnologij, znanj, metod in postopkov v prakso	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.18	Posredovanje novih znanj neposrednim uporabnikom (seminarji, forumi, konference)	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.19	Znanje, ki vodi k ustanovitvi novega podjetja ("spin off")	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.20	Ustanovitev novega podjetja ("spin off")	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.21	Razvoj novih zdravstvenih/diagnostičnih metod/postopkov	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.22	Izboljšanje obstoječih zdravstvenih/diagnostičnih metod/postopkov	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.23	Razvoj novih sistemskih, normativnih, programskih in metodoloških rešitev	
	Zastavljen cilj	☐ DA ☐ NE

	Rezultat	
	Uporaba rezultatov	
F.24	Izboljšanje obstoječih sistemskih, normativnih, programskih in metodoloških rešitev	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.25	Razvoj novih organizacijskih in upravljavskih rešitev	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.26	Izboljšanje obstoječih organizacijskih in upravljavskih rešitev	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.27	Prispevek k ohranjanju/varovanje naravne in kulturne dediščine	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.28	Priprava/organizacija razstave	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.29	Prispevek k razvoju nacionalne kulturne identitete	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.30	Strokovna ocena stanja	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.31	Razvoj standardov	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	

F.32	Mednarodni patent	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.33	Patent v Sloveniji	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.34	Svetovalna dejavnost	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	
F.35	Drugo	
	Zastavljen cilj	☐ DA ☐ NE
	Rezultat	
	Uporaba rezultatov	

Komentar

--

12.Samo za aplikativne projekte in podoktorske projekte iz gospodarstva!
Označite potencialne vplive oziroma učinke vaših rezultatov na navedena področja

	Vpliv	Ni vpliva	Majhen vpliv	Srednji vpliv	Velik vpliv	
G.01	Razvoj visokošolskega izobraževanja					
G.01.01.	Razvoj dodiplomskega izobraževanja	☐	☐	☐	☐	
G.01.02.	Razvoj podiplomskega izobraževanja	☐	☐	☐	☐	
G.01.03.	Drugo:	☐	☐	☐	☐	
G.02	Gospodarski razvoj					
G.02.01	Razširitev ponudbe novih izdelkov/storitev na trgu	☐	☐	☐	☐	
G.02.02.	Širitev obstoječih trgov	☐	☐	☐	☐	
G.02.03.	Znižanje stroškov proizvodnje	☐	☐	☐	☐	
G.02.04.	Zmanjšanje porabe materialov in energije	☐	☐	☐	☐	
G.02.05.	Razširitev področja dejavnosti	☐	☐	☐	☐	
G.02.06.	Večja konkurenčna sposobnost	☐	☐	☐	☐	
G.02.07.	Večji delež izvoza	☐	☐	☐	☐	
G.02.08.	Povečanje dobička	☐	☐	☐	☐	
G.02.09.	Nova delovna mesta	☐	☐	☐	☐	

G.02.10.	Dvig izobrazbene strukture zaposlenih	☐	☐	☐	☐	
G.02.11.	Nov investicijski zagon	☐	☐	☐	☐	
G.02.12.	Drugo:	☐	☐	☐	☐	
G.03	Tehnološki razvoj					
G.03.01.	Tehnološka razširitev/posodobitev dejavnosti	☐	☐	☐	☐	
G.03.02.	Tehnološko prestrukturiranje dejavnosti	☐	☐	☐	☐	
G.03.03.	Uvajanje novih tehnologij	☐	☐	☐	☐	
G.03.04.	Drugo:	☐	☐	☐	☐	
G.04	Družbeni razvoj					
G.04.01	Dvig kvalitete življenja	☐	☐	☐	☐	
G.04.02.	Izboljšanje vodenja in upravljanja	☐	☐	☐	☐	
G.04.03.	Izboljšanje delovanja administracije in javne uprave	☐	☐	☐	☐	
G.04.04.	Razvoj socialnih dejavnosti	☐	☐	☐	☐	
G.04.05.	Razvoj civilne družbe	☐	☐	☐	☐	
G.04.06.	Drugo:	☐	☐	☐	☐	
G.05.	Ohranjanje in razvoj nacionalne naravne in kulturne dediščine in identitete	☐	☐	☐	☐	
G.06.	Varovanje okolja in trajnostni razvoj	☐	☐	☐	☐	
G.07	Razvoj družbene infrastrukture					
G.07.01.	Informacijsko-komunikacijska infrastruktura	☐	☐	☐	☐	
G.07.02.	Prometna infrastruktura	☐	☐	☐	☐	
G.07.03.	Energetska infrastruktura	☐	☐	☐	☐	
G.07.04.	Drugo:	☐	☐	☐	☐	
G.08.	Varovanje zdravja in razvoj zdravstvenega varstva	☐	☐	☐	☐	
G.09.	Drugo:	☐	☐	☐	☐	

Komentar

--

13.Pomen raziskovanja za sofinancerje¹²

	Sofinancer		
1.	Naziv		
	Naslov		
	Vrednost sofinanciranja za celotno obdobje trajanja projekta je znašala:		EUR
	Odstotek od utemeljenih stroškov projekta:		%

Najpomembnejši rezultati raziskovanja za sofinancerja			Šifra
	1.		
	2.		
	3.		
	4.		
	5.		
Komentar			
Ocena			

14. Izjemni dosežek v letu 2012¹³

14.1. Izjemni znanstveni dosežek

Ker se je projekt zaključil v začetku leta 2012, najodmevnejši dosežki pa so bili objavljeni v letu 2011, sta na prosojnici opisana dva od njih.

Prosojnica je namerno pripravljena v nekoliko poljudnejšem slogu, ki je, predpostavljamo, primernejši za objavo.

14.2. Izjemni družbeno-ekonomski dosežek

C. IZJAVE

Podpisani izjavljam/o, da:

- so vsi podatki, ki jih navajamo v poročilu, resnični in točni
- se strinjamo z obdelavo podatkov v skladu z zakonodajo o varstvu osebnih podatkov za potrebe ocenjevanja ter obdelavo teh podatkov za evidence ARRS
- so vsi podatki v obrazcu v elektronski obliki identični podatkom v obrazcu v pisni obliki
- so z vsebino zaključnega poročila seznanjeni in se strinjajo vsi soizvajalci projekta

Podpisi:

*zastopnik oz. pooblaščen oseba
raziskovalne organizacije:*

in

vodja raziskovalnega projekta:

Univerza v Ljubljani, Fakulteta za
računalništvo in informatiko

Janez Demšar

ŽIG

Kraj in datum: Ljubljana 28.2.2013

Oznaka prijave: ARRS-RPROJ-ZP-2013/150

¹ Opredelite raziskovalno področje po klasifikaciji FOS 2007 (Fields of Science). Prevajalna tabela med raziskovalnimi področji po klasifikaciji ARRS ter po klasifikaciji FOS 2007 (Fields of Science) s kategorijami WOS (Web of Science) kot podpodročji je dostopna na spletni strani agencije (<http://www.arrs.gov.si/sl/gradivo/sifranti/preslik-vpp-fos-wos.asp>). [Nazaj](#)

² Napišite povzetek raziskovalnega projekta (največ 3.000 znakov v slovenskem in angleškem jeziku) [Nazaj](#)

³ Napišite kratko vsebinsko poročilo, kjer boste predstavili raziskovalno hipotezo in opis raziskovanja. Navedite ključne

ugotovitve, znanstvena spoznanja, rezultate in učinke raziskovalnega projekta in njihovo uporabo ter sodelovanje s tujimi partnerji. Največ 12.000 znakov vključno s presledki (približno dve strani, velikost pisave 11). [Nazaj](#)

⁴ Realizacija raziskovalne hipoteze. Največ 3.000 znakov vključno s presledki (približno pol strani, velikost pisave 11) [Nazaj](#)

⁵ V primeru bistvenih odstopanj in sprememb od predvidenega programa raziskovalnega projekta, kot je bil zapisan v predlogu raziskovalnega projekta oziroma v primeru sprememb, povečanja ali zmanjšanja sestave projektne skupine v zadnjem letu izvajanja projekta, napišite obrazložitev. V primeru, da sprememb ni bilo, to navedite. Največ 6.000 znakov vključno s presledki (približno ena stran, velikost pisave 11). [Nazaj](#)

⁶ Navedite znanstvene dosežke, ki so nastali v okviru tega projekta. Raziskovalni dosežek iz obdobja izvajanja projekta (do oddaje zaključnega poročila) vpišete tako, da izpolnite COBISS kodo dosežka – sistem nato sam izpolni naslov objave, naziv, IF in srednjo vrednost revije, naziv FOS področja ter podatek, ali je dosežek uvrščen v A" ali A'. [Nazaj](#)

⁷ Navedite družbeno-ekonomske dosežke, ki so nastali v okviru tega projekta. Družbeno-ekonomski rezultat iz obdobja izvajanja projekta (do oddaje zaključnega poročila) vpišete tako, da izpolnite COBISS kodo dosežka – sistem nato sam izpolni naslov objave, naziv, IF in srednjo vrednost revije, naziv FOS področja ter podatek, ali je dosežek uvrščen v A" ali A'.

Družbeno-ekonomski dosežek je po svoji strukturi drugačen kot znanstveni dosežek. Povzetek znanstvenega dosežka je praviloma povzetek bibliografske enote (članka, knjige), v kateri je dosežek objavljen.

Povzetek družbeno-ekonomskega dosežka praviloma ni povzetek bibliografske enote, ki ta dosežek dokumentira, ker je dosežek sklop več rezultatov raziskovanja, ki je lahko dokumentiran v različnih bibliografskih enotah. COBISS ID zato ni enoznačen, izjemoma pa ga lahko tudi ni (npr. prehod mlajših sodelavcev v gospodarstvo na pomembnih raziskovalnih nalogah, ali ustanovitev podjetja kot rezultat projekta ... - v obeh primerih ni COBISS ID). [Nazaj](#)

⁸ Navedite rezultate raziskovalnega projekta iz obdobja izvajanja projekta (do oddaje zaključnega poročila) v primeru, da katerega od rezultatov ni mogoče navesti v točkah 7 in 8 (npr. ker se ga v sistemu COBISS ne vodi). Največ 2.000 znakov, vključno s presledki. [Nazaj](#)

⁹ Pomen raziskovalnih rezultatov za razvoj znanosti in za razvoj Slovenije bo objavljen na spletni strani: <http://sicris.izum.si/> za posamezen projekt, ki je predmet poročanja [Nazaj](#)

¹⁰ Največ 4.000 znakov, vključno s presledki [Nazaj](#)

¹¹ Največ 4.000 znakov, vključno s presledki [Nazaj](#)

¹² Rubrike izpolnite / prepisite skladno z obrazcem "izjava sofinancerja" <http://www.arrs.gov.si/sl/progproj/rproj/gradivo/>, ki ga mora izpolniti sofinancer. Podpisan obrazec "Izjava sofinancerja" pridobi in hrani nosilna raziskovalna organizacija – izvajalka projekta. [Nazaj](#)

¹³ Navedite en izjemni znanstveni dosežek in/ali en izjemni družbeno-ekonomski dosežek raziskovalnega projekta v letu 2012 (največ 1000 znakov, vključno s presledki). Za dosežek pripravite diapozitiv, ki vsebuje sliko ali drugo slikovno gradivo v zvezi z izjemnim dosežkom (velikost pisave najmanj 16, približno pol strani) in opis izjemnega dosežka (velikost pisave 12, približno pol strani). Diapozitiv/-a priložite kot priponko/-i k temu poročilu. Vzorec diapozitiva je objavljen na spletni strani ARRS <http://www.arrs.gov.si/sl/gradivo/>, predstavitev dosežkov za pretekla leta pa so objavljena na spletni strani <http://www.arrs.gov.si/sl/analize/dosez/>. [Nazaj](#)

Obrazec: ARRS-RPROJ-ZP/2013 v1.00

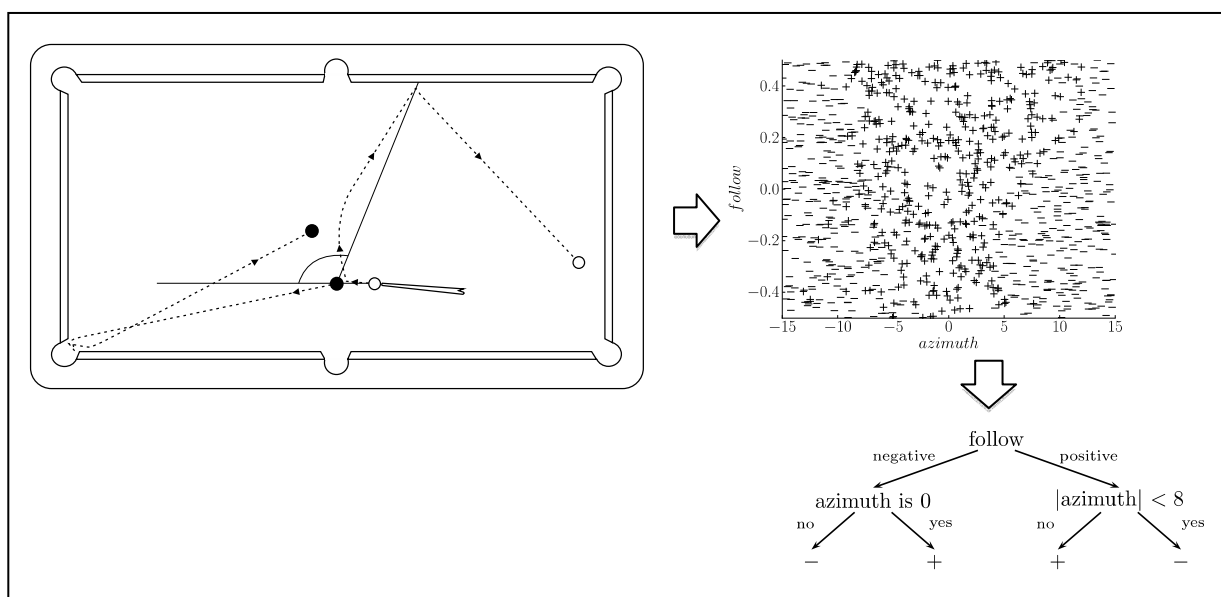
94-E3-C4-98-8A-24-29-D8-A8-D6-29-FC-4E-4B-F7-2B-17-2B-2E-B8

TEHNIKA

Področje: 2.07 Računalništvo in informatika

J. Žabkar, M. Možina, I. Bratko, J. Demšar. Learning qualitative models from numerical data. *Artificial Intelligence*, 2011, vol. 175, no. 9/10, str. 1604-1619.

J. Žabkar, M. Možina, I. Bratko, J. Demšar. Learning predictive qualitative models with Padé. *Informatica* (Ljubljana), dec. 2011, vol. 35, no. 4, str. 435-444.



Kvalitativni modeli so modeli, ki, za razliko od klasifikacijskih in regresijskih, opisujejo kvalitativne povezave, kot npr. $y = Q(+x, -z)$ if $t > u$, "y narašča z x in pada z z, če je t večji od u". Čeprav v načelu ne morejo doseči točnosti, ki jo dosegajo numerični modeli, so v praksi navadno robustnejši ter enostavnejši za razumevanje in uporabo. Enega najzgodnejših primerov uporabe takšnih modelov zasledimo v Samuelsonovi knjigi *Foundations of Economic Analysis* (1947), kasneje pa so jih uporabljali tudi na najrazličnejših drugih področjih.

V okviru projekta Kvalitativno modeliranje na osnovi podatkov smo raziskovali metode za avtomatsko gradnjo takšnih modelov iz podatkov. Prvi od gornjih člankov opisuje razvite metode, ki temeljijo na uteženi linearni regresiji v ustreznih okolicah, drugi pa metode, ki izvirajo iz topologije. Prve so se izkazale za hitrejše in robustnejše, poleg tega pa so zmožne ocenjevati tudi zanesljivost napovedi.

Razvite metode so uporabne na najrazličnejših področjih znanosti, ekonomije in drugod. Slika kaže primer iz članka: na podlagi poskusov s simulatorjem je računalnik avtomatsko odkril, kako se giba žoga za biljard, pri čemer upošteva tudi, na primer, vrtenje žoge zaradi udarca pod ali nad središčem. V kasnejših člankih smo metodo preskusili tudi na drugih, praktičnih področjih, kot je, na primer, napovedovanje bakterijskih okužb pri starostnikih.