

Heartfaid project: A web based knowledge discovery service

KDSA 2008

Rajko Horvat

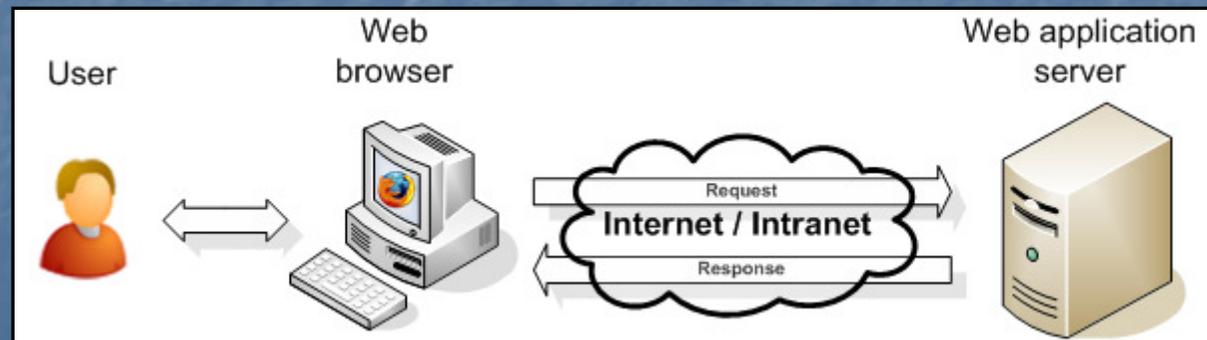
Laboratorij za informacijske sustave

Zavod za elektroniku

Institut Ruđer Bošković

Što je web aplikacija?

- Internet aplikacija koja se izvodi na poslužitelju koji je povezan na Internet ili intranet
- Bazirana je na klijent - server arhitekturi (klijent daje zahtjev koji server obradi i vrati rezultat obrade klijentu)
- Web preglednik služi kao klijent
- Komunikacija se odvija preko HTTP protokola
- Google i internet pošta su primjeri web aplikacije
- Bilo koja interaktivna web stranica se može nazvati web aplikacijom

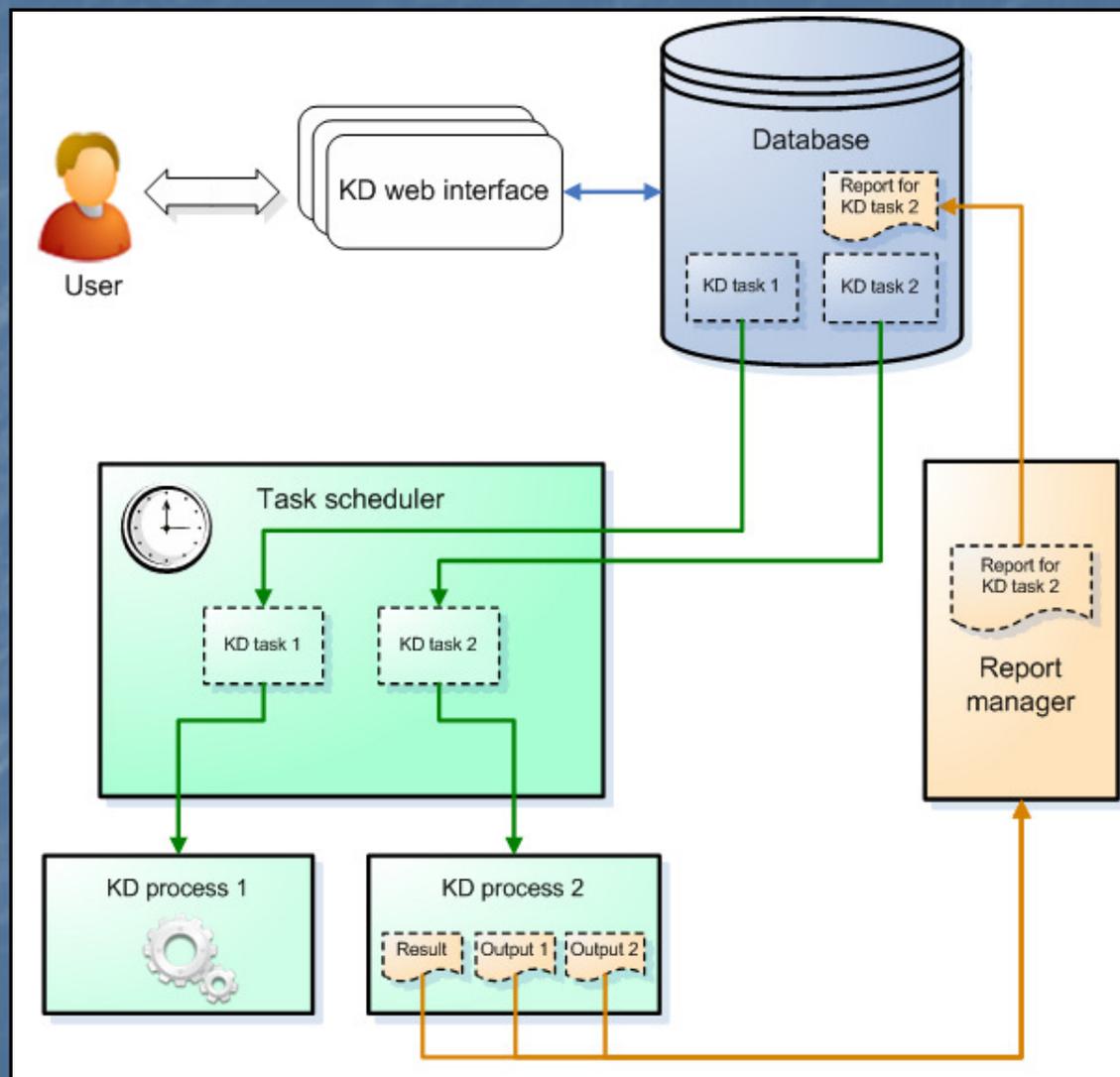


Prednosti web aplikacije

- Nepotrebna instalacija (komunikacija se odvija preko web preglednika)
- Aplikacija je neovisna o sustavu na kojem se nalazi web preglednik (Windows, Linux, MAC...)
- Aplikacija je dostupna preko Interneta/intraneta
- Obrada podataka se odvija na strani mrežnog poslužitelja, umjesto na lokalnoj radnoj stanici
- Svaka izmjena na aplikaciji je odmah dostupna svim korisnicima

Funkcionalni dijagram web aplikacije za otkrivanja znanja

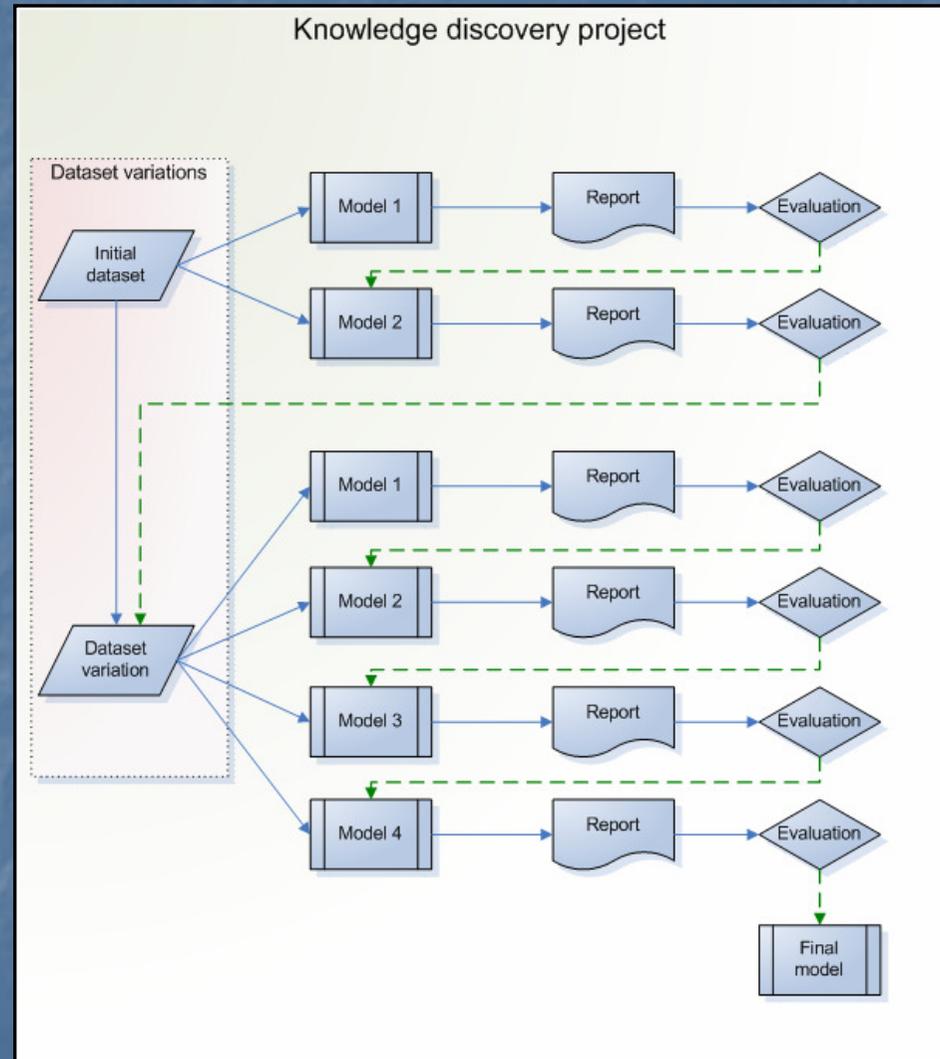
- Interaktivne web stranice
- Baza podataka
- Upravitelj zadatka
- Generator izvještaja
- Alati za otkrivanje znanja



Proces otkrivanja znanja u web aplikaciji

- Proces je osmišljen kao projekt koji sadrži:

- Setove podataka
- Modele (pokretanje alata za otkrivanje znanja)
- Izvještaje
- Bilješke



Kako to izgleda?

The screenshot shows a Mozilla Firefox browser window with the following elements:

- Browser Title Bar:** Heartfaid - Login - Mozilla Firefox
- Menu Bar:** File, Edit, View, History, Bookmarks, Tools, Help
- Address Bar:** http://heartfaid.irb.hr/Login.aspx
- Search Bar:** Google
- Bookmarks Bar:** Most Visited, Getting Started, Latest Headlines, lis:lis_index [ImagoWiki], imagoGallery - Home, TorrentFlux, Icecast Streaming Me..., Main Page - Mono
- Page Header:**
 - HEARTFAID** logo (heart icon)
 - KNOWLEDGE DISCOVERY SERVICES**
 - European Union flag logo
 - Information Society and Media logo
 - IST-2005-027107
 - Description: A knowledge based platform of services for supporting medical-clinical management of heart failure within elderly population
- Navigation Menu (Left):** About, Contact, Login
- Login Form (Center):**
 - Log in** header
 - User name:
 - Password:
 - Log in button
- Page Footer:**
 - Rudjer Boskovic Institute logo
 - LIS Laboratory for Information Systems logo
- Status Bar:** Done

Pregled projekata

KDS - Projects - Mozilla Firefox

http://heartfaid.irb.hr/KDD/Projects.aspx

HEARTFAID

KNOWLEDGE DISCOVERY SERVICES

A knowledge based platform of services for supporting medical-clinical management of heart failure within elderly population

Information Society and Media
IST-2005-027107

About
Contact
Resources
Projects
Settings
Admin
Logout

Description	Created				
iris-second	04/07/2008 16:28:04	Datasets	Tasks	Edit	Delete
iris-test	04/07/2008 16:25:21	Datasets	Tasks	Edit	Delete
Test	02/15/2008 15:36:03	Datasets	Tasks	Edit	Delete

Refresh New project

Rudjer Boskovic Institute LIS Laboratory for Information Systems

Done

Pregled zadatka

Heartfaid - Tasks - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://heartfaid.irb.hr/KDD/Tasks.aspx

Most Visited Getting Started Latest Headlines lis:lis_index [ImagoWiki] imagoGallery - Home TorrentFlux Icecast Streaming Me... Main Page - Mono Bugs - Mono

Heartfaid - Tasks

 **KNOWLEDGE DISCOVERY SERVICES**

A knowledge based platform of services for supporting medical-clinical management of heart failure within elderly population

  IST-2005-027107

Description	Dataset	Priority	Status	Algorithm	Created	Started on	Finished at	
test1	iris	Normal	Finished	Random forest	04/07/2008 16:25:47	04/07/2008 16:43:06	04/07/2008 16:43:08	Stop Restart View Report

Refresh New task

 Rudjer Boskovic Institute  LIS Laboratory for Information Systems

Done

Unos novog zadatka

Heartfaid - Tasks - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://heartfaid.irb.hr/KDD/Tasks.aspx

Most Visited Getting Started Latest Headlines lis:lis_index [ImagoWiki] imagoGallery - Home TorrentFlux Iccast Streaming Me... Main Page - Mono Bugs - Mono

Heartfaid - Tasks

HEARTFAID

KNOWLEDGE DISCOVERY SERVICES

A knowledge based platform of services for supporting medical-clinical management of heart failure within elderly population

Information Society and Media
IST-2005-027107

New task

Description

Priority

Dataset

Algorithm

Class attribute

Number of Trees

Attribute importances

Ok Cancel

Rudjer Boskovic Institute

LIS Laboratory for Information Systems

Done

Primjer izvještaja alata za otkrivanje znanja

Sample KD report

Dataset summary

Dataset name: sample.arff
 Number of trees in RF: 100
 Number of cases: 6851
 Number of attributes: 14
 Classes in the problem: 2 (Alive, Dead)
 Number of used attributes: 13
 Attributes to split on: 4

Problem runtime options:

Generate confusion matrix: on
 Generate variable importances: on
 Generate variable interactions: on
 Proximate cases radius: 200
 Number of prototypes per class: 2
 Class weights reassignment: Alive=1, Dead=4
 Runtime command:
 ./parf -t MIS_ANMCOM65_resm65_short_TS.arff -n 100 -s -t -i -p 200 -yn 2
 -yp 100 -ya -w 1,4 -v verbose

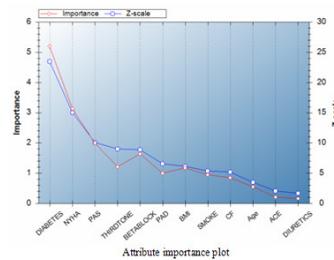
Model summary:

OOB transect accuracy: 33.73%
 Kappa: 0.2635

		Confusion matrix		
		Target class		
Classified as	Not classified	No Class	Alive	Dead
	Alive	0	3477	1545
	Dead	0	755	1074

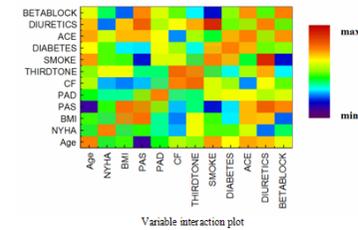
Attribute importances

Tag	Imp	Z-Sc	Significance
DIABETES	5.19	23.45	0.000000
NYHA	3.12	15.03	0.000000
PAS	1.99	10.06	0.000000
THIRDTONE	1.21	9.01	0.000000
BETABLOCK	1.64	8.87	0.000000
PAD	0.99	6.54	0.000000
BMI	1.17	6.11	0.000000
SMOKE	0.96	5.36	0.000000
CF	0.86	5.18	0.000000
Age	0.54	3.45	0.00282
ACE	0.21	2.02	0.021435
DIURETICS	0.16	1.64	0.050576



Variable interactions

	Age	NYHA	BMI	PAS	PAD	CF	THIRD-TONE	SMOKE	DIURETICS	ACE	DIURETICS	BETABLOCK
Age	0	-5	-23	-128	-84	-39	-3	-24	-4	-12	-20	-20
NYHA	-5	0	-24	-29	-17	-24	-24	-20	-24	-24	-200	-74
BMI	-23	-24	0	-1	-20	-100	-21	-37	-39	-11	-14	-109
PAS	-128	-29	-1	0	-20	-84	-24	-114	-39	-11	-11	-72
PAD	-84	-17	-20	-20	0	-71	-19	-33	-21	-24	-40	-21
CF	-39	-24	-100	-84	-71	0	-12	-14	-14	-39	-20	-21
THIRDTONE	-3	-24	-21	-24	-21	-12	0	-11	-35	-45	-9	-38
SMOKE	-24	-20	-37	-114	-33	-14	-6	-14	-37	-24	-112	-72
DIURETICS	-4	-11	-19	-39	-21	-14	-15	-11	0	-10	-24	-44
ACE	-12	-24	-11	-21	-20	-39	-20	-27	-10	0	-38	-4
DIURETICS	-20	-20	-14	-11	-30	-21	-9	-21	-21	-30	0	-24
BETABLOCK	-20	-74	-109	-14	-37	-33	-38	-113	-44	-4	-25	0



Hvala na pažnji!