

Analiza kvarova uređaja i opreme korištenjem metoda poslovnog izvješćivanja

Milković, Vili

Scientific master's theses / Magistarski rad

2008

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture / Sveučilište u Zagrebu, Fakultet strojarstva i brodogradnje**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:235:179575>

Rights / Prava: [In copyright / Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-04-25**

Repository / Repozitorij:

[Repository of Faculty of Mechanical Engineering and Naval Architecture University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

ANALIZA KVAROVA UREĐAJA I OPREME
KORIŠTENJEM METODA POSLOVNOG
IZVJEŠĆIVANJA

MAGISTARSKI RAD

Vili Milković, dipl. inž. str.

ZAGREB, 2008.

**SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE**

**ANALIZA KVAROVA UREĐAJA I OPREME
KORIŠTENJEM METODA POSLOVNOG
IZVJEŠĆIVANJA**

MAGISTARSKI RAD

Mentor:

prof.dr.sc. Nedeljko štefanić

Vili Milković, dipl. inž. str.

ZAGREB, 2008.

PODACI ZA BIBLIOGRAFSKU KARTICU

UDK :

Ključne riječi: Poslovno izvješćivanje, skladištenje podataka, OLAP sustavi, metode dubinskih analiza, klaster analiza, multidimenzijsko skaliranje, regresijske metode, srednje vrijeme između kvarova.

Znanstveno područje: TEHNIČKE ZNANOSTI

Znanstveno polje: STROJARSTVO

Institucija u kojoj je rad izrađen: Fakultet strojarstva i brodogradnje, Sveučilište u Zagrebu

Mentor rada: Prof. dr. sc. Nedjeljko Štefanić

Broj stranica: 135

Broj slika: 41

Broj tablica: 19

Broj grafikona: 1

Broj korištenih bibliografskih jedinica: 41

Datum obrane: 19.9.2008.

Povjerenstvo: Prof. dr. sc. Nedeljko Štefanić -Fakultet strojarstva i brodogradnje -Zagreb
Prof. dr. sc. Ivo Čala - Fakultet strojarstva i brodogradnje -Zagreb
Prof. dr. sc. Josip Zavada – Fakultet prometnih znanosti- Zagreb

Institucija u kojoj je rad pohranjen: Fakultet strojarstva i brodogradnje, Sveučilište u Zagrebu



Zagreb, 27. 05. 2008.,

Zadatak za magistarski rad

Kandidat: VILI MILKOVIĆ, dipl.inž.strojarstva

Naslov zadatka: ANALIZA KVAROVA UREĐAJA I OPREME KORIŠTENJEM
METODA POSLOVNOG IZVJEŠĆIVANJA

Sadržaj zadatka:

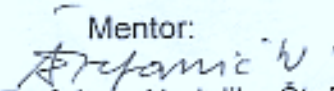
U poslovnim sustavima i u njihovoj okolini kontinuirano se generiraju podaci i informacije koje je potrebno obraditi odgovarajućim alatima i metodama kako bi oni predstavljali podloge prilikom odlučivanja poduzeća. Da bi iz tako generiranih podataka i informacija slijedile upravljačke akcije potrebno je iste prikupiti, pohraniti te ih odgovarajućim alatima obraditi kako bi od njih stvorili nova znanja koja će menadžmentu poduzeća (strateškom, taktičkom i operativnom) omogućiti aktivnosti (donošenje optimalnih odluka, izbor ispravnih pravaca djelovanja) što će pak poduzeću osigurati stabilnu poziciju na tržištu. Koncept poslovnog izvješćivanja razvijen je s ciljem obrade prethodno uskladištenih podataka poduzeća, a obuhvaća širok spektar različitih metoda.

U radu je potrebno:

- detaljno razraditi koncept poslovnog izvješćivanja,
- navesti načine upravljanja podacima koji se generiraju u poduzeću,
- sistematizirati metode koje se koriste pri analizi podataka,
- na primjeru funkcije održavanja primijeniti navedene metode te analizirati dobivene rezultate,
- procijeniti ekonomske i ostale vrste efekata koji se mogu postići primjenom metoda poslovnog izvješćivanja.

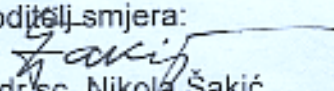
Zadatak zadan: 17.06.2008.

Rad predan:

Mentor:

Prof.dr.sc.Nedeljko Štefanić


Predsjednik Odbora za
poslijediplomske studije:

Prof.dr.sc. Tomislav Filetin

Voditelj smjera:

Prof.dr.sc. Nikola Šakić

ZAHVALA

Najveću zahvalnost dugujem mojoj obitelji - majci, teti, tetku, sestri, koji su me u svakom trenutku podržavali pri izradi ovoga rada.

Posebno bih se zahvalio mentoru dr. Nedeljku Štefaniću bez čije pomoći ovaj rad ne bi bio moguć.

Također se ovom prigodom moram prisjetiti velikog čovjeka koji mi je pomogao u trenucima kad je ovaj rad dobio svoj pravi smisao, a to je dr. Ivo Čala.

Posebna mi je čast i zadovoljstvo poznavati ova dva čovjeka koji su mi u nekim teškim trenucima izrade rada davali sigurnost i samopouzdanje, te imali veliko razumijevanje i strpljenje, što je danas rijetkost u ovome bešćutnom i okrutnom životnom okruženju.

Hvala Vam na svemu. Vili

SADRŽAJ

PREDGOVOR	VII
SAŽETAK RADA	VIII
KLJUČNE RIJEČI / KEY WORDS	IX
POPIS OZNAKA	X
POPIS SLIKA	XI
POPIS TABLICA / GRAFIKONA	XIII
1. UVOD	1
1.1. Metodologija izrade rada	1
1.2. Hipoteza i kompozicija rada	2
2. KONCEPT POSLOVNOG IZVJEŠĆIVANJA	4
2.1. Definicije poslovnog izvješćivanja	6
2.1.1. Neki od osnovnih razloga uvođenja sustava poslovnog izvješćivanja	8
2.1.2. Neki od najčešćih oblika otpora uvođenju sustava poslovnog .. izvješćivanja	9
2.1.3. Vrijednost poslovnih informacija	10
2.1.4. Metapodaci	15
2.2. Skladištenje podataka	16
2.2.1. Zahtjevi koji se postavljaju pred skladište podataka	17
2.2.2. Mjesto i funkcije skladišta podataka u informacijskom sustavu tvrtke	18
2.3. Ahitekture skladišta podataka	20
2.4. Spiralni pristup razvoju sustava poslovnog izvješćivanja	22
2.5. OLAP alati	24

2.5.1. Struktura OLAP sustava	24
3. DUBINSKE ANALIZE (data mining).....	27
3.1. Metode dubinskih analiza.....	36
3.1.1. Klasteriranje (metoda klastera).....	37
3.1.1.1. Hijerarhijska metoda	41
3.1.1.2. K-means klasteriranje.....	44
3.1.2. Stablo odlučivanja.....	46
3.1.3. Multidimenzijsko skaliranje	53
3.1.4. Neuronske mreže	58
3.1.5. Neizrazita logika.....	61
3.1.5.1. Neizraziti sustavi.....	64
3.1.6. Genetički algoritmi	66
3.2. Softverska rješenja poslovnog izvješćivanja na inozemnom i	73
domaćem tržištu	
3.2.1. SAP - Vodeći svjetski proizvođač poslovnih softvera.....	75
3.2.2. Oracle warehouse builder (programski paket).....	79
3.2.3. Analiza tržišta OLAP alata.....	82
3.2.4. Inozemne tvrtke i njihovi partneri na hrvatskom tržištu.....	84
4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA...	87
4.1. Opis ključnih agregata klimatiziranih putničkih vagona HŽ-a	90
4.2. Održavanje putničkih vagona Hrvatskih željeznica	93
4.3. Analiza rezultata uporabom metoda poslovnog izvješćivanja.....	99
4.3.1. Primjena klaster postupka u analizi kvarova	99
4.3.1.1. Metoda najbližeg susjeda (Kvadratni Euklid)	100
4.3.1.2. Metoda najbližeg susjeda, prosječna udaljenost	107
4.3.2. Primjena postupka multidimenzijskog skaliranja u analizi	110

kvarova	
4.3.3. Primjena PARETO postupka u analizi kvarova.....	113
4.3.4. Primjena regresijske analize	115
4.4. Prijedlog poboljšanja trajnosti pojedinih agregata prema proizvođaču	119
4.5. Odabir optimalne metode u rješavanju problema	124
5. PRIMJENA STATISTIČKIH ALATA U ANALIZI KVAROVA.....	127
6. ZAKLJUČAK.....	130
7. LITERATURA.....	133
ŽIVOTOPIS.....	135

PREDGOVOR

U svojoj relativno kratkoj profesionalnoj karijeri, tijekom pet godina rada na poslovima održavanja putničkih vagona na Hrvatskim željeznicama, pored svakodnevnih radnih aktivnosti bio sam u mogućnosti pratiti velike količine generiranih podataka vezanih uz tehnološki proces održavanja i popravaka putničkih vagona, a koji su bili pospremani u arhive te posebne vagonске kartone. Osim konkretnih informacija i znanja, izuzetno korisno je bilo to što sam znao tko ima određene informacije, vodi evidencije, te gdje se one mogu naći. Pokušavajući pronaći model kako „iskoristiti“ veliku količinu generiranih podataka, u radu će se primijeniti metode poslovnog izvješćivanja kako bi se mogao bar djelomično riješiti taj problem. Na kraju, na osnovu dobivenih rezultata upotrijebit će se statistički alati vezani uz proces održavanja kojim će se pokušati riješiti problemi i predložiti rješenja za zamjenom i ugradnjom novih agregata na putničkim vagonima što bi dovelo do smanjenja troškova održavanja i vremena zadržavanja vagona u radionici zbog popravaka što je i primarni cilj ovoga rada..

SAŽETAK RADA

U analizi podataka i njihovoj optimalnoj iskoristivosti u tvrtci Održavanje vagona d.o.o. (u vlasništvu Hrvatskih željeznica), potrebno je primjenjivati suvremene metode dubinskih analiza koje su sastavni dio poslovnog izvješćivanja. U radu je dat pregled primjena metoda poslovnog izvješćivanja u iskoristivosti informacija i podataka bitnih za daljnje analize. Prikupljeni su podaci o kvarovima na svim serijama klimatiziranih modernih putničkih vagona u razdoblju od 1.1.2005. godine do 1.1.2007. godine, te je napravljena njihova obrada i analiza raznim metodama. Na temelju dobivenih rezultata analize, definirani su najznačajniji uzročnici kvarova na klimatiziranim vagonima. Potom su uspoređivani rezultati pojedinih metoda i odabrana je optimalna metoda.

PAPER SUMMARY

In analysing data for Maintenance of Railway Carriages LLC which is owned by the Croatian Railways and with regard to their optimal utilization, it is necessary to apply modern methods of depth analyses which are component parts of the business intelligence. The paper provides an overview of the applied business intelligence methods with regard to the utilization of the information and data necessary for further analyses. Covering the period from 1 January 2005 to 1 January 2007, data on defects in all model ranges of the air-conditioned modern passenger carriages were collected, processed and analyzed by applying different methods. Based on the results of the analysis, the most important causes of defects in the air-conditioned carriages were identified. After that, the results of the individual methods were compared and the optimal method was chosen.

KLJUČNE RIJEČI / KEY WORDS

Poslovno izvješćivanje

Skladište podataka

OLAP alati

Dubinske analize

Klaster analiza

Multidimenzijsko skaliranje

Stablo odlučivanja

Neuronske mreže

Regresijska analiza

Business Intelligence

Data Warehouse

OLAP tools

Data mining

Cluster analysis

Multidimensional scaling

Decision tree

Neural networks

Regression analysis

POPIS OZNAKA

Oznaka	Naziv (engl.)	Naziv
OLAP	On-line Analytical Processing	On-line analitička obrada
ID3	Interactive Dichtomizer 3	Interaktivna dihotomija (3. verzija)
MDS	Multidimensional scaling	Multidimenzijско skaliranje
AID	Automatic Interaction Detection	Automatizirana interakcijska detekcija
CRT	Classification and Regression Trees	Klasifikacija i regresija stabla
CHAID	Chi - square Automatic Interaction Detection	Hi - kvadrat automatizirana interakcijska detekcija
NN	Neural Networks	Neuronske mreže
SVIK	Mean time between failures	Srednje vrijeme između kvarova

POPIS SLIKA

Slika 1.	Metodologija izrade rada	3
Slika 2.	Vrijednost informacije u vremenu [3]	12
Slika 3.	Kontinuum vrijednosti informacije [3]	13
Slika 4.	Mjesto skladišta podataka u informacijskom sustavu tvrtke [3]	18
Slika 5.	Dvoslojna arhitektura s jednim zajedničkim skladištem podataka [3]	21
Slika 6.	Razvoj sustava poslovnog izvješćivanja [3].....	22
Slika 7.	Trodimenzionalni prikaz podataka.....	25
Slika 8.	Pregled tehnologija dubinskih analiza [6]	29
Slika 9.	Pregled klaster algoritama [8]	37
Slika 10.	Euklid udaljenost [8]	39
Slika 11.	City Block udaljenost [8].....	40
Slika 12.	Primjer dendograma.....	44
Slika 13.	K-means algoritam [3]	45
Slika 14.	Podprocesi u stvaranju stabla odlučivanja [7]	46
Slika 15.	Prikaz primjera podjele stabla odlučivanja fizikalnih mjerenja [7]	47
Slika 16.	Grafički prikaz multidimenzionalne analiza (MDS analysis).....	57
Slika 17.	Tipovi neuronskih mreža [8]	58
Slika 18.	Matematički model neurona [29].....	59
Slika 19.	Funkcije pripadnosti izrazitog i neizrazitog skupa VRUĆE [31].....	63
Slika 20.	Shematski prikaz neizrazitog sustava [31].....	65
Slika 21.	Evolucijski program – širi pojam od genetičkog algoritma.....	67
Slika 22.	Pristup problemu s pomoću genetičkog algoritma i evolucijskog programa	68

Slika 23.	Struktura evolucijskog programa	68
Slika 24.	Ponuđači alata poslovnog izvješćivanja 2002./2003. godine [39].....	74
Slika 25.	Ponuđači alata poslovnog izvješćivanja u 2007. godini [41].....	75
Slika 26.	Tvrtka SAP [37].....	78
Slika 27.	Glavni prozor Oracle Warehouse Buildera [40].....	79
Slika 28.	Uvodni prozor New Module Wizarada u ORACLE-u [40].....	82
Slika 29.	Vagonski kartoni sa evidencijom datuma kvara i vrstom kvara.....	89
Slika 30.	Blok shema uređaja za opskrbu vagona električnom energijom	90
Slika 31.	Matrica udaljenosti prema kvadratnom Euklidu za grupe kvarova.....	100
Slika 32.	Dendrogram kvarova (metoda najbližeg susjeda, kvadratni Euklid).....	101
Slika 33.	Matrica udaljenosti prema metodi prosječne udaljenosti (average	107
	distance) za grupe kvarova.	
Slika 34.	Dendrogram kvarova (metoda najbližeg susjeda, average distance).....	108
Slika 35.	Multidimenzionalni prikaz grupa kvarova (u logaritamskom mjerilu).....	110
Slika 36.	Pareto analiza frekvencije kvarova na klimatiziranim putničkim vagonima...	113
Slika 37.	Prikaz rezultata frekvencije kvarova na klimatiziranim putničkim vagonima..	114
Slika 38.	Prikaz linearne regresije ovisnosti broja kvarova statičkih pretvarača.....	118
	i broja prijeđenih kilometara za pojedinu grupu serija vagona	
Slika 39.	Koraci za rješavanje problema klaster analizom [8]	125
Slika 40.	Polje podataka klaster analize u software-u Statgraphics verzija 5.1.....	127
Slika 41.	Polje odabira metoda klaster analize u Software-u Statgraphics 5.1.	128

POPIS TABLICA

Tablica 2.1.	Neke od uporaba OLAP-a [34]	26
Tablica 3.1.	Prikaz nedovoljnog broja podataka za dubinske analize..... postupcima regresije ili neuronske mreže [6]	32
Tablica 3.2.	Prikaz povijesnih koraka u razvoju dubinskih analiza [32].....	35
Tablica 3.3.	Formule za izračun udaljenosti klastera (koeficijent fuzije) pri..... odabranom aglomerativnom postupku [8]	43
Tablica 3.4.	Usporedbe algoritama stabla odlučivanja AID-a, CHAID-a, XAID-a.	51
Tablica 3.5.	Razlike između ekspertnih sustava i neuronskih mreža [31]... ..	66
Tablica 3.6.	Prikaz vodećih tvrtki na svjetskom OLAP tržištu i njihov udjel u tom tržištu u 2001., 2002. i 2003. godinu [39]	84
Tablica 4.1.	Podjela modernih putničkih vagona Hrvatskih željeznica obzirom na seriju, tip statičkog pretvarača i brzinu vagona	88
Tablica 4.2.	Održavanje putničkih vagona prema uputi 241-1 Hrvatskih..... željeznica	94
Tablica 4.3.	Troškovi održavanja za pojedinu seriju vagona (u kunama) bez..... cijene utrošenog materijala i dodatnih radova	98
Tablica 4.4.	Broj kvarova na klimatiziranim putničkim vagonima u razdoblju..... 1.1.2005.- 1.1.2007.godine	99
Tablica 4.5.	Aglomeracijski raspored klastera udaljenosti pojedinih grupa kvarova metodama klastera (najbliži susjed)	102
Tablica 4.6.	Aglomeracijski raspored klastera udaljenosti pojedinih grupa..... kvarova metodom prosječne udaljenosti (average distance)	109
Tablica 4.7.	Prikaz rješenja metodom multidimenzijskog skaliranja.....	111
Tablica 4.8.	Prikaz izračuna funkcije STRESS1 multidimenzijskog skaliranja...	112

Tablica 4.9.	Prikaz prijeđenih kilometara i kvarova na pojedinim komponentama klimatiziranih vagona u razdoblju od 1.1.2005.- 1.1.2007. godine	115
Tablica 4.10.	Rezultati regresijske analize (ovisnost broja kvarova statičkih..... pretvarača o broju prijeđenih kilometara grupa pojedinih serija vagona)	116
Tablica 4.11.	Kvarovi komponenti statičkog pretvarača Siemens EVA/B.....	122
Tablica 4.12.	Kvarovi komponenti statičkog pretvarača Končar VIS 50-1.....	123

POPIS GRAFIKONA

Grafikon 1.	Prikaz veličine svjetskog tržišta OLAP alatima [39].....	83
-------------	--	----

1. UVOD

Prema [33] je procjena američke tvrtke Gartner da je količina skupljenih informacija u cijelom svijetu od 2004.- 2007. jednaka svim podacima koje je čovječanstvo prikupilo i pohranilo počevši od kamenog doba do godine 2000. Prosječna tvrtka danas koristi manje od deset posto prikupljenih podataka, procjena je znanstvenika i istraživača. Ukoliko se tome pridodaju i vanjski podaci (konkurencija, makroekonomski pokazatelji...), situacija je daleko od savršene. Činjenica je da su danas tvrtke pretrpane velikim količinama podataka, ali se s druge strane traži dosta znanja. U poslovnim krugovima postoji potreba za analizom takvih količina podataka. Kada je riječ o poslovnim izvješćivanju, potrebno je reći da ona danas objedinjuje nekoliko vrlo bitnih metodologija i pripadajućih tehnologija pomoću kojih je moguće poboljšati proces odlučivanja, a da pri tome koriste sustave za potporu odlučivanja temeljene na poslovnim činjenicama i podacima. Pri tome se prije svega podrazumijevaju metodologije skladištenja i dubinskih analiza podataka. Poslovno izvješćivanje, skladištenje podataka, dubinske analize i ostalo su tehnologije o kojima se u Hrvatskoj još ne razmišlja dovoljno. Prije nekoliko godina još je imalo smisla razgovarati o tome koja od hrvatskih tvrtki ima skladište podataka kao podlogu sustave poslovnog izvješćivanja.

1.1. Metodologija izrade rada

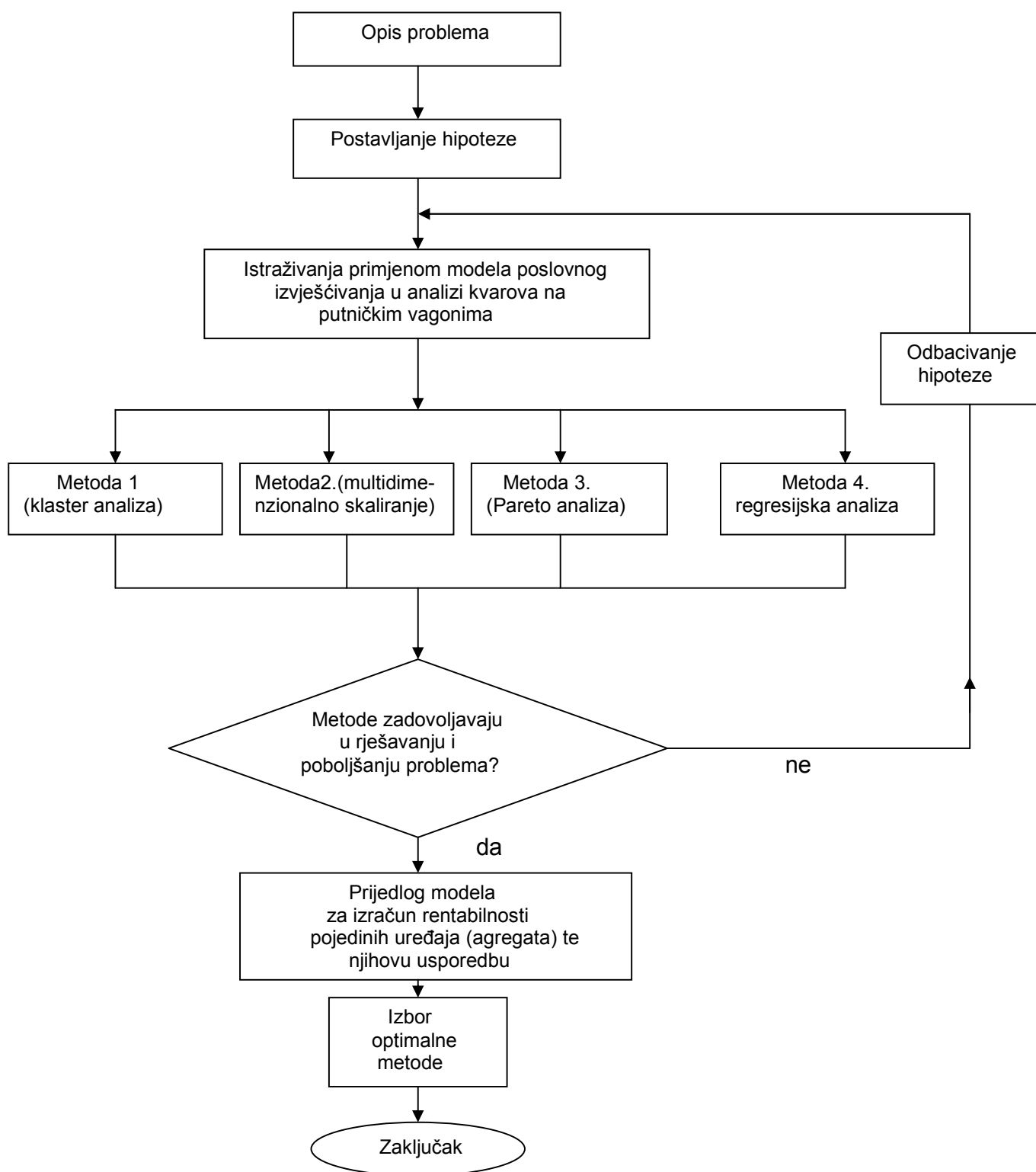
Metoda općenito znači planski postupak istraživanja i ispitivanja neke pojave, odnosno način rada za ostvarivanje nekog cilja na filozofskom, znanstvenom ili političkom području. Metoda znači gotov model, proceduru, redoslijed po kojoj se odvija neka praktična djelatnost. Metoda znači i misaoni postupak da se lakše i što točnije otkriju i sustavno obrade znanstvene činjenice, podaci i informacije. Znanstvena metoda je skup različitih postupaka kojima se znanost koristi u znanstveno-istraživačkom radu da bi istražila i izložila rezultate znanstvenog istraživanja u određenom znanstvenom području ili znanstvenoj

disciplini. Istodobno, znanstvena metoda je i put istraživanja kojim se oblikuje i izgrađuje znanost [1].

1.2. Hipoteza i kompozicija rada

Primjenom sustava poslovnog izvješćivanja tj. uporabom različitih metoda, moguće je unaprijed doprinijeti donošenju optimalnih odluka i izbora pravaca djelovanja vrhovnog (top) menadžmenta koji će tvrtci omogućiti snažniju poziciju na današnjem tržištu. Polazište rezultata dobivenih istraživanjem je odluka o zamjeni pojedine opreme (agregata) povećane raspoloživosti, tj. većim srednjim vremenom izmađu kvarova (SVIK) onima manje raspoloživosti koji su se pokazali nerentabilnim. Istraživanje se sastoji u dubinskim analizama snimaka kvarova na pojedinim serijama putničkih vagona u razdoblju od 1.1. 2005. do 1.1. 2007. godine uporabom različitih metoda poslovnog izvješćivanja, njihovom usporedbom, te odabirom optimalne metode. U istraživanju će se koristiti različite metode poslovnog izvješćivanja: 1. klasteriranje, 2. multidimenzionalno skaliranje, 3. regresijska metoda i kao pomoćna metoda Pareto analiza. Metodologija izrade rada prikazana je slikom 1.

U **drugom poglavlju** opisani su suvremeni koncepti poslovnog izvješćivanja, osnovne definicije, opis današnjih OLAP sustava, kao i vrijednosti poslovnih informacija. U **trećem poglavlju** objašnjen je teorijski aspekt suvremenih metoda poslovnog izvješćivanja, te softverska rješenja na inozemnom kao i na domaćem tržištu. U **četvrtom poglavlju** date su metodologije provedbe istraživanja kvarova, kao i dobiveni rezultati pojedinim metodama, te odabir optimalne metode. U sklopu poglavlja dat je opis pojedinih agregata putničkih vagona, kao i način njihova održavanja. Na kraju poglavlja objašnjen je model rješavanja problema kvarova uporabom analitičkih formula za izračun srednjeg vremena između kvarova usporedbom pojedinih agregata, te prijedlog odabira opreme sa većom rentabilnošću. U **završnom poglavlju** opisana su pojedina softverska rješenja neophodna u izradi ovog rada, te zaključak rada.



Slika 1. Metodologija izrade rada

2. KONCEPT POSLOVNOG IZVJEŠĆIVANJA

Mnogi slučajevi dobrih i loše donesenih odluka ostali su zabilježeni u povijesti čovječanstva. Pitanje je kako je Paris pobijedio dotad nepobjedivog ratnika Ahila ?. Pobijedio ga je, jer je imao informaciju o „Ahilovoj ranjivoj peti“ i donio odluku kako iskoristiti tu slabost. Ovdje je riječ naravno o mitu, a svrha mita je izvlačenje pouke. Prema [10], Ben Glad poznat kao osnivač „Competitive Intelligencea“ prilikom analize konkurenata propisuje pronalaženje „Crvenog gumba“¹, koji predstavlja „Ahilovu petu“.

Danas se na tržištu odvijaju mnoge borbe, te svako želi pronaći protivnikovu slabu stranu, tj. Ahilovu petu, a isto tako i svoju, a to je sve moguće primjenom poslovnog izvješćivanja. U ovom radu će se koristiti mnogi pojmovi koji su u svakodnevnoj uporabi u engleskoj verziji. Kako se želi neke pojmove koji su na hrvatskom jeziku koristiti u ovome radu, daje se paralelni pregled korištenih pojmova i njihove definicije .

Hrvatski	Engleski	Kratice	Definicija
Poslovno Izvješćivanje	Business Intellegence	BI	Uporaba kolektivnog znanja organizacije sa ciljem postizanja konkurentske prednosti.
Skladište podataka	Data Warehouse	DW	Kopija transakcijskih podataka posebno strukturiranih za upite i izvještavanje.
Dubinske analize	Data Mining	DM	Analiza neprimjetnih veza među podacima pomoću sofisticiranih procedura.
OLAP	OLAP	OLAP	Analitičko procesiranje. Višedimenzionalna analiza podataka za podršku odlučivanju. Obuhvaća analizu posl. informacija od ad hoc upita do predviđanja, modeliranja i simulacija

¹ Red Button je definiran kao osjetljivo područje. „Aktiviranje“ crvenog gumba od strane protivnika izaziva unaprijed određenu poznatu reakciju.

U stručnoj i znanstvenoj literaturi, kao i u novinarstvu tj. dnevnom tisku i financijskim časopisima, poslovno izvješćivanje je sve prisutnije. U tumačenju pravog značenja ovog pojma uočljive su nesuglasice, te u velikoj mjeri pogrešni pristupi. Stvarnost je da se pojavi poslovnog izvješćivanja može pristupiti sa najmanje dva gledišta [3]:

- s gledišta makroaspekta, i
- s gledišta mikroaspekta.

Prema [3], s makroaspekta, poslovno je izvješćivanje složena kategorija koja se oblikuje sustavnim prikupljanjem podataka o makroekonomskim kretanjima u određenoj sredini, njihovim evidentiranjem odnosno pohranjivanjem, te računskom obradom zbog otkrivanja trendova.

Razmatranje sa mikroaspekta, višestruke interpretacije otežavaju istraživanje zbog nemogućnosti da se jednoznačno odredi predmet istraživanja. Jedna od definicija poslovnog izvješćivanja promatrane s mikroaspekta je: „Poslovno izvješćivanje nije niti proizvod niti sustav. To je arhitektura i kolekcija integriranih operativnih aplikacija i aplikacija za potporu odlučivanju, te baza podataka koje poslovnim korisnicima omogućuju lak pristup poslovnim podacima“ [3].

Ranih 60. tih godina 20. stoljeća, s pojavom vrlo jakih računala treće tehnološke generacije počelo je skupljanje velikih količina podataka [3]. Stručnjaci su uvidjeli da ih je neophodno transformirati u uporabljive podatke koje će biti važni pri donošenju velikih poslovnih odluka. Ogroman napredak u tom smislu je predstavljao razvoj relacijskih baza podataka 80. tih godina 20. stoljeća. Ta su rješenja potisnula u prvi plan ideju o sustavima za podršku odlučivanju (engl. Decision Support System, DSS), koja su se vremenom razvijala u koncept poslovnog izvješćivanja. Svojim evoluiranjem je poslovno izvješćivanje privlačilo sve veći broj stručnjaka, znanstvenika i menadžera za analitičkim primjenama. Proizvođači alata za poslovno izvješćivanje razvili su i proizveli softverske alate koji se mogu upotrijebiti u tvrtkama i to u svim odjelima ili sektorima, s ciljem analiziranja te razmjene informacija iz različitih izvora.

Ovakove primjene tj. aplikacije pružaju uvelike korisnicima gotova rješenja, te znače veliki korak u razvitku sustava za potporu odlučivanju na svim razinama tvrtke.

2.1. Definicije poslovnog izvješćivanja

Ono što je u većini slučajeva zajedničko kod autora kojima je stalo da ozbiljno definiraju pojam *poslovnog izvješćivanja* je objedinjavanje pojmom metoda, koncepata i odgovarajuće informatičke tehnologije s ciljem poboljšanja procesa odlučivanja te upotrebi rezultata sustava u svrhu boljeg, informiranijeg poimanja poslovnih procesa (unutarnjih uvjeta), te boljeg razumijevanja okoline u kojoj poduzeće djeluje (vanjski uvjeti), a s ciljem izgradnje i održavanja konkurentske prednosti. Sustav poslovnog izvješćivanja daje cjelovit pogled na poduzeće i time omogućava proaktivan pristup vođenju, prognozirajući budućnost, nudeći više scenarija i alternativa kako bi poduzeće bilo pripremljeno za većinu situacija u kojima se vremenom može naći. Pojam „Business Intelligence“ (BI) počeo se u Hrvatskoj prevoditi kao „Poslovno izvješćivanje“. U engleskom jeziku riječ „Intelligence“ ima dva značenja [5]:

- 1) sposobnost učenja, razumijevanja, logičkog razmišljanja, sposobnost da se te stvari rade dobro.
- 2) tajna informacija sakupljena o stranoj zemlji, osobito o neprijateljskoj, osobe koje sakupljaju te informacije.

Poslovno izvješćivanje (*Business Intelligence, BI*) je skup metodologija i softverskih alata koji omogućuju korištenje podataka iz skladišta podataka (*Data Warehouse*) i njihovo pretvaranje u informaciju potrebnu za donošenje poslovnih odluka.

Sustav poslovnog izvješćivanja je takav sustav koji čuva informacije i znanje o konkurenciji, kupcima, dobavljačima, procesima. On omogućava poslovno pregovaranje i brojčano argumentirani nastup prema kupcima i dobavljačima, kvalitetno operativno planiranje, praćenje ponašanja konkurencije, promatranje pojedinih tržišnih segmenata, te predviđanje budućih pojava. Također, sustav poslovnog izvješćivanja nudi bolje razumijevanje vlastitih kupaca i spoznaju što ih potiče na određeno ponašanje.

Poslovno izvješćivanje se intenzivno počelo razvijati kada su poduzeća automatizirala svoje poslovne procese, odnosno implementirala različite transakcijske sustave, koji su se vrlo brzo pokazali kao generatori velikih količina podataka. S tehničke strane, poslovno izvješćivanje je proces kojim se sirovi podaci pretvaraju u informacije. Te informacije se zatim analiziraju i koriste u procesu odlučivanja.

U nastavku rada navedeno je nekoliko reprezentativnih definicija i popratnih tumačenja poslovnog izvješćivanja (Business intelligence) pojedinih autora:

Bernard Liataud [3] :

„Inteligencija podiže informacije u organizaciji na višu razinu. Podaci i informacije su stvari. Inteligencija je organska materija; izvedena iz informacija, ona pridonosi stvaranju takvog stanja organizacije koja se može nazvati kolektivnom inteligencijom. Inteligencija proizlazi iz punog razumijevanja informacija, ranije poduzetih akcija i postojećih mogućnosti, odnosno opcija. Jednom posijana, inteligencija će se sama širiti širom organizacije. Krična masa pojedinaca koji imaju dobar uvid i dijele slične stavove o poslovnim procesima postaje u takvim uvjetima moćna snaga “.

Stevan Dedijer [13] :

„Poslovno izvješćivanje je korištenje kolektivnog znanja organizacije sa ciljem postizanja konkurentске prednosti“.

Ravi Kalakota, Marcia Robinson [3] :

„Pretvaranje podataka u znanje zadatak je aplikacija poznatih pod nazivom poslovno izvješćivanje. Poslovno izvješćivanje je skupina novih aplikacija oblikovanih tako da mogu organizirati i strukturirati podatke o poslovnim transakcijama na način koji omogućuje analizu korisnu u potpori odlučivanju i operativnim aktivnostima kompanije“

Ostale definicije poslovnog izvješćivanja:

Poslovno izvješćivanje je sistematičan način prikupljanja i analiziranja javno dostupnih podataka o aktivnostima konkurencije na temelju kojih se mogu predvidjeti budući poslovni trendovi. Sustavi poslovnog izvješćivanja pomažu u rješavanju upravljačkih problema i čuvaju informacije i znanje o konkurenciji, kupcima, dobavljačima, procesima i vezama među procesima te omogućuju poslovno pregovaranje, argumentirani nastup prema kupcima i dobavljačima.

2.1.1. Neki od osnovnih razloga uvođenja sustava poslovnog izvješćivanja:

Poslovno izvješćivanje je povezano sa činjenicom da si niti jedna ozbiljna tvrtka u uvjetima kakvi danas vladaju na tržištu zbog velike konkurenciju, ne može priuštiti grešku da se poslovne odluke donose na temelju subjektivnih procjena menadžera. Zbog toga, navedeni razlozi koji slijede za uvođenje poslovnog izvješćivanja u tvrtci su čvrsti argumenti za postizanje konkurentne tvrtke [33].

1. *Okolina tvrtke nije više statična.* Danas je ponuda proizvoda i usluga velika, te je stoga i konkurencija velika. Novi uvjeti rada traže nova rješenja, a novi izazovi i velike napore.

2. *Globalizacija.* Globalizacijom tržišta i utjecajem interneta u svaku poru društva, okolnosti su se promijenile. Sada su kupac i prodavač povezani mail-om ili sličnom vrstom konverzacije. Dok su tvrtke pretrpane raznim podacima, s druge strane, postoji nedostatak korisnih i upotrebljivih informacija. Da bi se smanjio raskorak između količine raspoloživih podataka i informacija, potrebno je definirati procese prikupljanja podataka i njihovo oblikovanje u korisne informacije.

3. *Pronalazak novih klijenata* je nekoliko puta skuplje nego zadržavanje postojećih. Ukoliko tvrtka uspije smanjiti odlazak kupaca konkurenciji za 5%, tada može udvostručiti svoju zaradu.

4. *Nezadovoljstva kupaca* je velika opasnost koja tvrtku stavlja u težak položaj. Samo 4% nezadovoljnih kupaca izravno se žali na lošu kvalitetu proizvoda ili usluga. 90% kupaca koji nisu zadovoljni kvalitetom proizvoda, izbjegavat će taj proizvod. Svaki od nezadovoljnih kupaca obavijestit će veliki broj drugih osoba.

5. *Održavanje likvidnosti* je operativni i upravljački problem. Rješenja ovog problema izravno utječu na upravljanje poslovnim rezultatom. Da bi se ovladalo ovim operativnim problemima, treba poznavati svoje klijente, dobavljače, procese i veze među njima.

6. *Ciklus operativnog kontrolinga* (prikupljanje podataka, planiranje, analiza i kontrola te upravljanje) u tvrtci funkcionira, ukoliko postoji informacijska infrastruktura.

2.1.2. Neki od najčešćih oblika otpora uvođenju sustava poslovnog izvješćivanja:

Prema [33], novinarka časopisa «DM Review» Susan Osterfelt je u svibanjskom broju 2000. godine navela nekoliko učestalih razloga zbog kojih određene interesne skupine pružaju otpor uvođenju sustava poslovnog izvješćivanja. Vjerojatnost je da će se na svakom od projekata poslovnog izvješćivanja pojaviti bar nekoliko uzroka otpora uvođenju. Voditelj projekta ima zadatak pronaći način kako probleme uzrokovane time riješiti ili zaobići, a o tome uvelike ovisi konačan uspjeh projekta. Razlozi su:

1. *Manjak strateškog razmišljanja* – Zbog velike preopterećenosti operativnim problemima, nema se vremena strateški pozicionirati. Tu bi se moglo dodati i druga dva vida nedostatka strategije - situacije kada se ne zna i kada se ne želi strateški razmišljati.

2. *Manjak razumijevanja vrhovnog menadžmenta* – u kojem menadžment ne shvaća koje rezultate može dati projekt poslovnog izvješćivanja, te ne daje potpunu podršku i to je ključni razlog za propast projekta.

3. *Neadekvatni i nekvalitetni podaci kao izgovor* - podaci nisu kvalitetni, te se ne isplati ulagati u skupocjene alate koji će raditi dubinske analize podataka prepunih grešaka i netočnosti.

4. *Borba za položaj* - puno ljudi se u organizaciji bori za vlastite pozicije i interese, te ne žele vidjeti u projektu poslovnog izvješćivanja interes za sebe i neće ga podržati.

5. *Otpor informatičara* – ovaj pogled oslikava se činjenicom da se informatičari ne žele baviti projektima poslovnog izvješćivanja, jer ih mnogi smatraju „šminkerskim“ i beskorisnim izvješćima.

6. *Strah od promjena* - standardni strah koji u organizaciji postoji od promjene bilo koje vrste i kojim se inače bavi teorija upravljanja promjenama.

2.1.3. Vrijednost poslovnih informacija

Pri utvrđivanju vrijednosti informacija potrebno je uzeti u obzir i određenu dozu subjektiviteta svih pojedinaca ili skupina koje informacije stvaraju, kao i onih koje ih primaju da bi ih upotrijebili u neke njima zanimljive potrebe. Iz mnogo dosad razvijenih pristupa problemu utvrđivanja vrijednosti informacija najpoznatije su tri za koje postoji mišljenje da mogu najviše pridonijeti razumijevanju načina i mjere u kojoj informacije svojom vrijednošću pridonose boljem ili lošijem upravljanju poslovanja poduzeća. Utemeljitelji su prva dva pristupa C.E. Shannon i R.Hamming, dok je treći Bernard Liataud u suradnji sa Markom Hammondom. Prema [3], radi se o slijedećim pristupima:

- Shannonov kvantitativni pristup,
- Hammingov vremenski pristup,
- Liataudov poslovno-pragmatički pristup

a.) Shannonov kvantitativni pristup vrednovanju informacija

Količina informacija je jednaka entropiji polja slučajnih događaja koje generiraju informaciju. Do ove spoznaje je došao C.E.Shannon koji se u počecima svog znanstveno-istraživačkog rada bavio problemom utvrđivanja količine informacija [3].

Prema [3], osnovna formula za izračun entropije pritom glasi:

$$H(X) = -\sum_{i=1}^n p_i \cdot \lg(p_i) \quad (1)$$

Pri čemu je: $i=1, n$

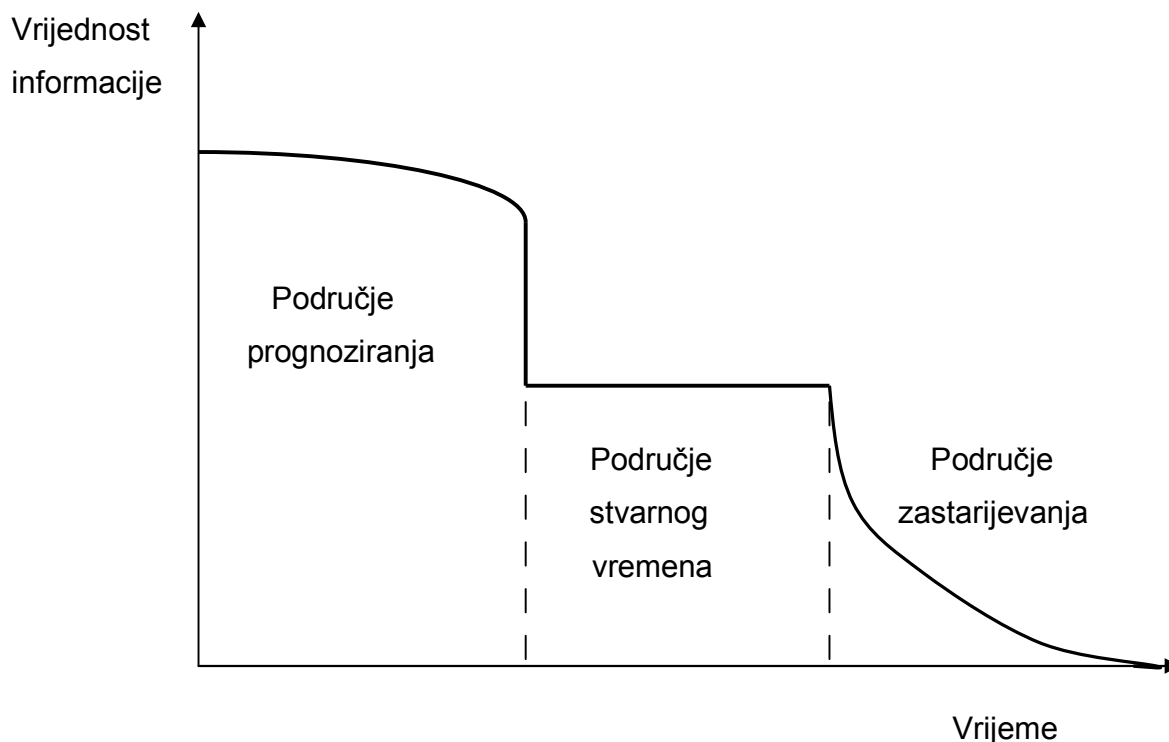
$H(X)$ - entropija polja slučajnih događaja X ,

p_i - vjerojatnost i -tog događaja iz polja X ,

$\lg(p_i)$ – dualni logaritam vrijednosti vjerojatnosti i -tog događaja iz polja X

b.) Hammingov vremenski pristup vrednovanju informacija

Svaka informacija u određenom trenutku nastaje, neko vrijeme traje, da bi nakon isteka tog vremena nestala. Od ove pretpostavke polazi Richard Hamming kao koncepta životnog ciklusa informacije. Vrijednost informacije smatra kao funkciju vremena. Shematski prikaz vrijednosti informacije kao vremenski promjenjive veličine prikazan je slikom 2. na kojoj su vidljiva tri karakteristična područja- 1. *područje prognoziranja*- u kojem se informacija dobiva prije nego što je donešena odluka i tu joj je vrijednost najveća., 2. *područje stvarnog vremena* – u kojem se informacija dobiva upravo u trenutku kada je potrebno donijeti odluku, te 3. *područje zastarijevanja* – koje nastupa kad se informacija dobiva prekasno, te kada nam donošenje odluke više nije potrebno.



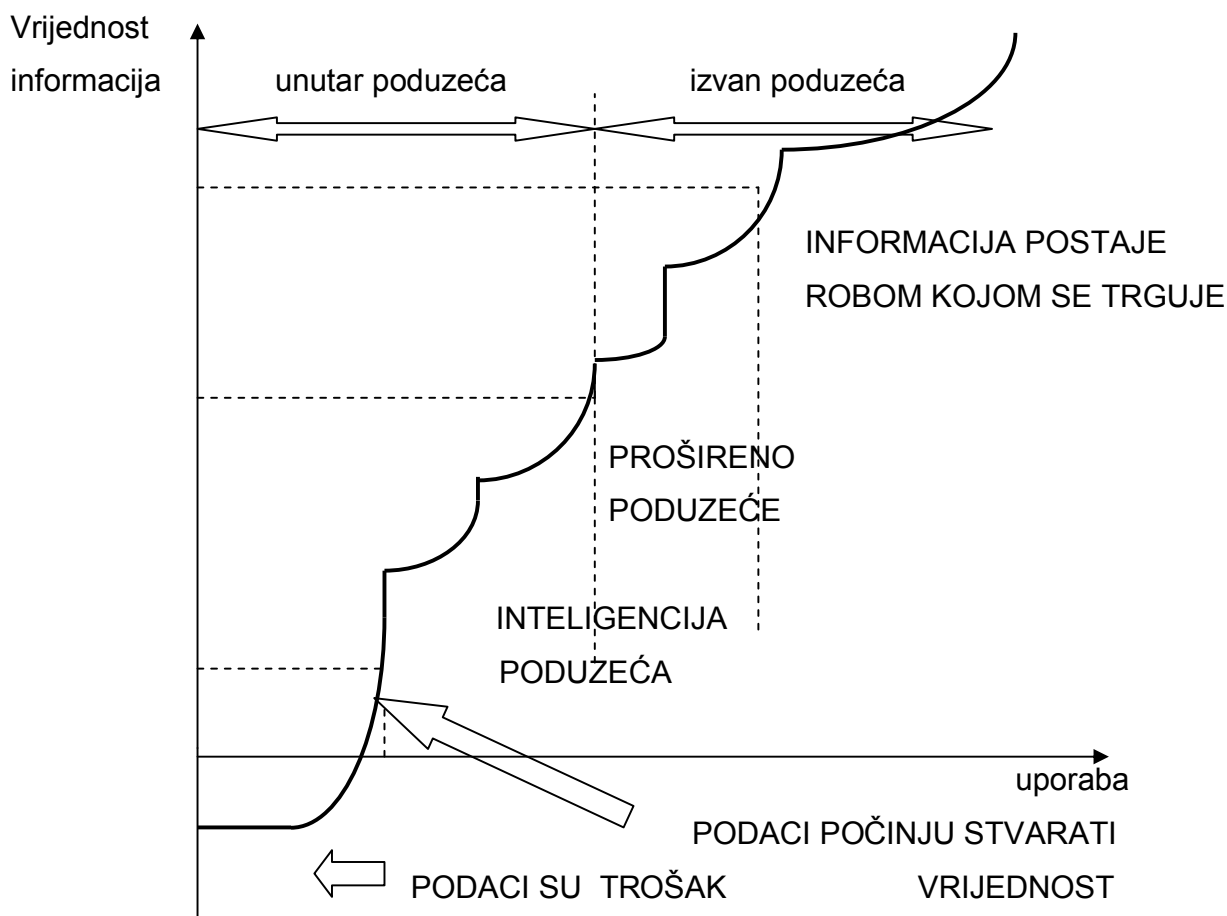
Slika 2. Vrijednost informacije u vremenu [3]

c.) Liautaudov poslovno - pragmatični pristup vrednovanju informacija

Prema [3], Bernard Liautaud polazište za tumačenje i vrednovanje informacija nalazi u činjenici da se vrijednost određene informacije proteže u kontinuumu. Vrijednost informacija se definira kao funkcija broja korisnika na kvadrat tih informacija koji ih mogu dobiti i analizirati ih, te broja poslovnih područja kojima korisnici pripadaju, te se iskazuju slijedećim izrazom:

$$\text{Vrijednost (informacija)} \sim \text{korisnici}^2 \times \text{poslovna područja} \quad (2)$$

U kontinuumu vrijednosti određenih podataka (informacija), može se uočiti pet karakterističnih područja, a koja su prikazana slikom 3.



Slika 3. Kontinuum vrijednosti informacije [3]

a.) Područje u kojem su podaci trošak

Podaci spremjeni u memoriji računala, a koje nitko ne koristi, predstavljaju za tvrtku isključivo trošak, a vrijednost troškova neiskorištenosti ovih informacija je veća od njihove same vrijednosti. Jedini korisnici ovih podataka su informatičari. Vidljivo je da ovaj slučaj ne smije biti cilj uspostavljanja informacijskog sustava u tvrtci, već je neophodno učiniti apsolutno sve da bi postojeće neiskorištene informacije bile zanimljive i dostupne mnogim zaposlenicima tvrtke koji bi iz njih mogli stvarati novu vrijednost i pomogle menadžmentu za donošenje poslovnih odluka.

b.) Područje u kojem podaci počinju stvarati novu vrijednost

Osnovni preduvjet koji treba učiniti da bi se aktivno mogli koristiti podaci unutar jedne tvrtke je izgradnja lokalne mreže danas poznatije kao intranet, gdje će djelatnici moći izravno pristupiti podacima. Što će biti veći broj korisnika ovih informacija, to će i vrijednost samih informacija eksponencijalno rasti, te će biti više koristi za donošenje boljih i kvalitetnijih odluka na razini tvrtke. Ovime se stvara poslovno izvješćivanje na razini jedne službe ili odjela.

c.) Područje inteligencije poduzeća

U ovom području, poslovno izvješćivanje stvoreno na razini odjela ili službe, tvrtka čini dostupnom ostalim službama ili odjelima, te u ovom slučaju vrijednost informacija izrazito naglo raste ukoliko je poprimi nova grupa korisnika iz ostalih službi ili odjela. Na iste informacije različiti korisnici gledaju iz drugih kuteva, što u konačnici može dovesti do nastajanja pozitivnih sinergijskih efekata.

d.) Područje proširenog poduzeća

Svaka tvrtka je u bilo kakvom kontaktu sa svojom okolinom, dakle poslovnim partnerima, kupcima, dobavljačima, te sa njima razmjenjuje kontrolirano informacije, jačajući svoje poslovno izvješćivanje. U ovom se slučaju stvara Liautaudov pristup vrednovanju informacija gdje vrijednost informacija raste umnošku kvadrata broja korisnika i broja poslovnih područja. Tvrtka se na taj način proširuje, prerastajući u tzv. ekstrapoduzeće. Raspoloživo tehnološko organizacijsko rješenje za uspostavljanje ekstrapoduzeća zove se ekstranet.

Ekstranet ili prošireni Internet je privatna poslovna mreža nastala povezivanjem intraneta nekoliko partnerskih organizacija. Ekstraneti služe prije svega unapređenju odnosa s klijentima tvrtke, njenim dobavljačima i poslovnim partnerima i to putem komunikacijskih kanala, pa se u ovoj fazi još uvijek pretjerano ne inzistira na

tretiranju informacije kao robe koja ima određenu tržišnu vrijednost i koja se može prodati po određenoj cijeni. Uloga poslovnog izvješćivanja u ekstranetu se može usporediti s ulogom bankomata u bankovnom poslovanju. Bankomat, dakle nudi uslugu samoposluživanja u bankarstvu. Ukoliko nam nedostaje novca, a poslovnica banke u kojoj jesmo ne radi, moguće je otići do obližnjeg bankomata i podići novac.

e.) Područje u kojoj informacija postaje roba kojom se trguje

Koncept elektroničkog poslovanja podrazumijeva intenzivnu primjenu informacijske i internetske tehnologije pri izvršavanju ključnih poslovnih funkcija tvrtke. Primjena tog današnjeg oblika poslovanja pogoduje stvaranju poslovnog izvješćivanja. Poslovno izvješćivanje se može stvarati i u uvjetima tradicionalnog poslovanja, ali uz puno više problema, troškova i potrebnog vremena te uz skromnije rezultate negoli je to slučaj onda kada se tvrtka usmjeri konceptu elektroničkog poslovanja. Uz uporabu poslovnog izvješćivanja kumuliranog e-poslovanjem, tvrtke se mogu posvetiti traženju novih poslovnih modela napuštajući posebno one tradicionalne. U jednom takvom novom modelu, tvrtke se usmjeravaju pronalaženju informacija do kojih su same došle i njihovoj prodaji drugim zainteresiranim poslovnim subjektima. Postupajući tako, one ulaze u onu vrijednosnu zonu u kojoj informacije postaju roba kojom se trguje.

2.1.4. Metapodaci

Metapodaci su podaci o podacima, te opisuju tvrtku u smislu njenih poslovnih aktivnosti i objekata nad kojima ili pomoću kojih se poslovne aktivnosti obavljaju. Metapodaci su od velike važnosti za poslovno izvješćivanje, jer olakšavaju pretvorbu poslovnih podataka u informacije. Postoji više vrsta metapodataka kojima je potrebno upravljati. Vrste metapodataka su :

- administrativni metapodaci,
- poslovni metapodaci,
- operativni metapodaci.

Administrativni metapodaci uključuju sve informacije neophodne za uspostavu i korištenje skladišta, opisi izvornih baza podataka, krajnji korisnički alati, derivirani podaci, fizička organizacija podataka, ekstrakcija podataka, čišćenje, pravila transformacije, profili korisnika itd. *Poslovni metapodaci* uključuju poslovne termine i definicije te vlasništvo nad podacima. *Operativni metapodaci* uključuju informacije koje se sakupljaju tijekom rada skladišta: odnos transformiranih i migriranih podataka, aktualnost podataka i nadzorne informacije kao što su statistika korištenja, izvješća o greškama i inspekcijama.

2.2. Skladištenje podataka

William H. Inmon je 1992. godine predstavio koncept i definiciju tzv. skladišta podataka prema kojem je to oblik baze podataka kojem su svojstvena četiri obilježja [3] :

1. usmjerenost predmetima,
2. sadržajna nepromjenljivost,
3. integriranost i
4. vezanost uz vrijeme (vremenska određenost).

Svaki sustav poslovnog izvješćivanja temelji se na podacima. Za potrebe analize, podaci se prikupljaju u skladišta podataka. Skladište podataka je skup podataka orijentiran analitičkom korištenju, što znači da je optimiziran za upite bazi koji dohvaćaju veću količinu podataka. Za razliku od skladišta podataka, transakcijski sustav je orijentiran na veliki broj istovremenih korisnika, te brzo zapisivanje, čitanje i izmjenu malih količina podataka, najčešće samo jednog sloga. Glavni građevni elementi skladišta podataka su činjenične tablice (engl. Fact tables) koje sadrže kvantitativne, numeričke podatke koji se analiziraju, te dimenzijske tablice (engl. dimension tables), koje sadrže kvalitativne, strukturne podatke koji služe za kreiranje hijerarhija za zbrajanje i analizu. Još jedna bitna značajka skladišta podataka je veličina, možda je čak ispravnije reći: *ogromnost*. Naime, zbog velikog povijesnog perioda koji podaci u skladištu podataka pokrivaju, zbog broja aplikacija iz kojih dolaze te zbog činjenice da u procesu punjenja podataka dolazi do denormalizacije i kreiranja dodatnih kolona u tablicama, skladišta podataka su najčešće vrlo velika. U tom

kontekstu skladište od 100 GB smatra se relativno malim skladištem, dok se u velika skladišta mogu ubrojiti ona veličine od preko 1 TB. Najveća svjetska skladišta podataka od gotovo 30 TB mogu se naći u telekomunikacijskim tvrtkama. Naravno, porast količine podataka ne donosi povećanje dimenzija samo diskovnog prostora, nego i procesorske i memorijske snage, te komunikacijske propusnosti sustava kako bi svi podaci pravovremeno bili učitani.

2.2.1. Zahtjevi koji se postavljaju pred skladište podataka

Zahtjevi koji se postavljaju pred skladišta podataka su [3]:

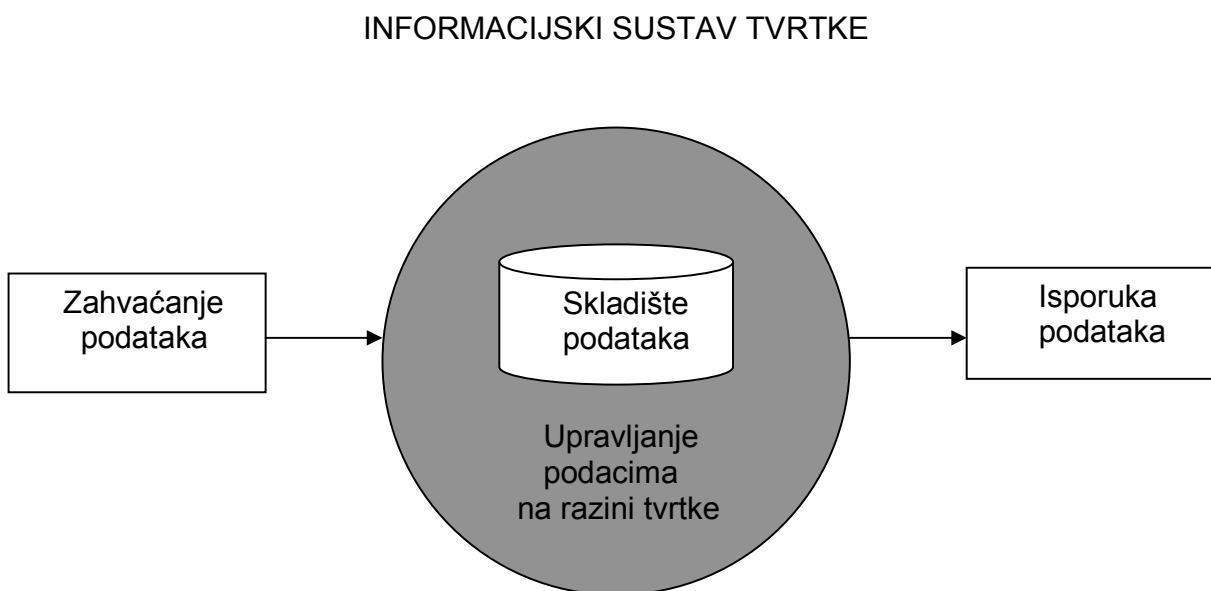
1. *Sadržajnost velike količine podataka.* Tvrtka mora imati pregled nad svime što se zbivalo tijekom svih marketinških kampanja koje je vodila i mjeriti njihovu učinkovitost, a to može samo ako ima trajno zabilježene sve pojedinosti u vezi sa tim kampanjama.
2. *Skladište podataka treba kontinuirano ažurirati podacima o izvršenim marketinškim, prodajnim i drugim poslovnim transakcijama.*
3. *Skladište podataka mora služiti velikom broju ljudi - menadžerima i zaposlenicima tvrtke.*
4. *Neki od zaposlenika će željeti postavljati upite i pretraživati čitavo skladište podataka kako bi došli do nekih zaključaka.*
5. *Skladište podataka mora uvijek biti raspoloživo.* Tvrtka si ne smije dopustiti luksuz da ono gdje kad ne bude dostupno onome kome je potrebno, jer ta osoba tada neće moći pružiti odgovarajuću uslugu klijentu i može se dogoditi da ih tvrtka zbog toga zauvijek izgubi.

6. *Treba poduzeti rigorozne mjere zaštite integriteta i tajnosti osjetljivih podataka pohranjenih u skladištu.* Poznato je da veliki broj ljudi odbija kupovati ili obavljati neke druge poslove online zbog straha da će njihovi osobni podaci biti zloupotrebljeni, a privatnost ugrožena.

Ukratko, skladište podataka je zlatni rudnik informacija, pa predstavlja moćno sredstvo upravljanja općenito, a posebice odnosima s klijentima. Da bi svi njegovi potencijali mogli biti u potpunosti iskorišteni, prilikom njegova uspostavljanja, implementacije i korištenja mora se čvrsto pridržavati ovih navedenih šest zahtjeva.

2.2.2. Mjesto i funkcije skladišta podataka u informacijskom sustavu tvrtke

U informacijskom sustavu tvrtke skladište podataka je mjesto u kojem se skupljaju i pohranjuju podaci, ali isto tako i izvor informacija koje će se koristiti u odlučivanju i pri stvaranju poslovnog izvješćivanja. Mjesto skladišta podataka u okviru informacijskog sustava tvrtke prikazan je shematski na slici 4.



Slika 4. Mjesto skladišta podataka u informacijskom sustavu tvrtke [3]

Primarne funkcije skladišta podataka su slijedeće :

- skladište podataka je posljedica pravila poslovanja koje se primjenjuju u tvrtci - ne samo na razini pojedine poslovne funkcije ili organizacijske jedinice – prilikom donošenja strateških poslovnih odluka. Ono mora biti takvo da se može brzo i jednostavno prilagođavati promjenama u poslovnim pravilima, što uključuje sposobnost prihvatanja novih podataka, kao i promjena u hijerarhijskim i logičkim odnosima među podacima.
- skladište podataka je „muzej“ strateških informacija koja iziskuje primjenu tehnike modeliranja koja omogućuje jednostavnu ugradnju povijesne perspektive u pohranjeni sadržaj.
- skladište podataka je izvor informacija koje se izvode iz spremišta podataka (Data Mart). Ta se spremišta mogu iskoristiti za dubinske analize, istraživanje, postavljanje upita ili analitičku online obradu.

Osnove koje moraju zadovoljavati spremišta podataka (Data Mart) su :

- samostalnost (samosvojstvo),
- organizacijski aspekti,
- aspekt zaštite podataka kroz djelomični osvrt na podatke,
- smanjenje količine (volumena) podataka,
- raspodjela tereta,
- neovisnost od aktualnih ciklusa skladišta podataka (Data Warehouse)

Podaci za skladište su, prema [15]:

Unutarnji podaci – koji opisuju aktivnosti unutar poduzeća (uzimaju se iz podsustava za obradu transakcija). To su:

- financijski podsustavi,
- logistički podsustavi (planiranje materijala, nabava, transport, distribucija.....),
- podsustavi održavanja industrijskih postrojenja
- podsustavi proizvodnje (upravljanje i kontrola proizvodnje, praćenje radnih naloga, upravljanje kvalitetom...),
- podsustavi ljudskih resursa (plaće, radna mjesta, obrazovanje),
- informacijski podsustavi (izvještajni podsustavi).

Vanjski podaci – koji opisuju aktivnosti izvan poduzeća su:

- podaci o konkurentnosti (proizvodi, usluge...),
- ekonomski podaci (valutne promjene i promjene kamata),
- strukovni podaci (tehnološki i marketinški trendovi),
- financijski podaci, robni podaci (cijene sirovina),
- ekonometrijski podaci (ponašanje kupaca),
- psihometrijski podaci (profiliranje kupaca),
- meteorološki podaci,
- demografski podaci,
- marketinški podaci.

2.3. Arhitekture skladišta podataka

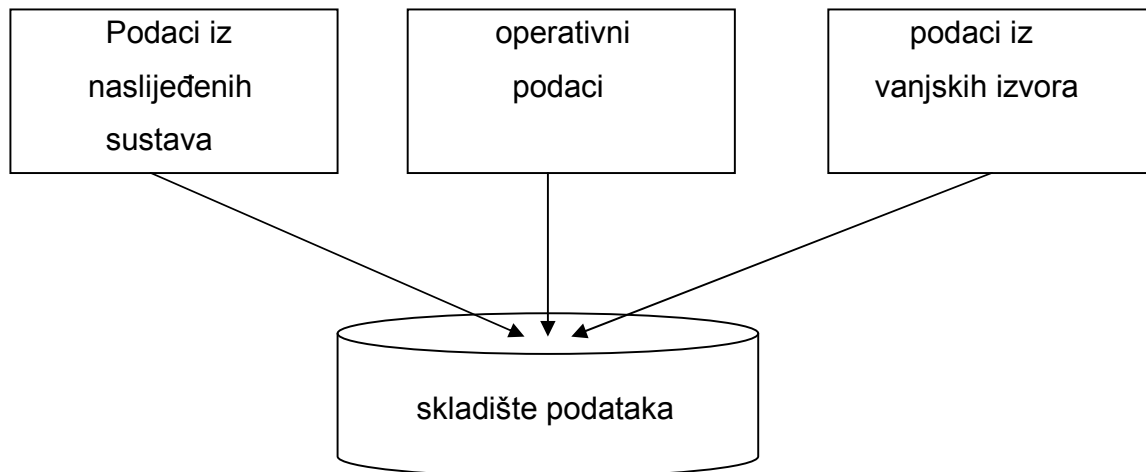
Do danas su razvijene i koriste se tri osnovne arhitekture skladišta podataka [3]:

- dvoslojna arhitektura s jednim zajedničkim skladištem podataka,
- dvoslojna arhitektura s većim brojem nezavisnih lokalnih spremišta podataka,
- troslojna arhitektura s jednim zajedničkim skladištem podataka i većim brojem povezanih lokalnih spremišta podataka.

U praksi je moguće pronaći primjere sva tri modela skladištenja podataka.

Dvoslojna arhitektura s jednim zajedničkim skladištem podataka je tradicionalni način uporabe skladišta podataka, pri kojemu središnje skladište podataka ekstrahira podatke iz izvora poput starijih arhitektura, kao što su IMS ili DB2, online sustava za obradu transakcija, interneta te drugih komunikacijskih mreža ili kanala.

Troslojna arhitektura s jednim zajedničkim skladištem podataka ima prednost u odnosu na dvoslojne arhitekture zbog smanjenja opterećenja informatičara i djelatnika tvrtke, unapređuje se točnost informacija bez obzira iz kojeg su izvora uzete, te se uklanjaju poteškoće u komunikaciji i suradnji jedinica organizacije i pojedinih dijelove tvrtke na poslovima koji se trebaju obavljati zajednički. Dvoslojna arhitektura s jednim zajedničkim skladištem podataka prikazan je slikom 5.

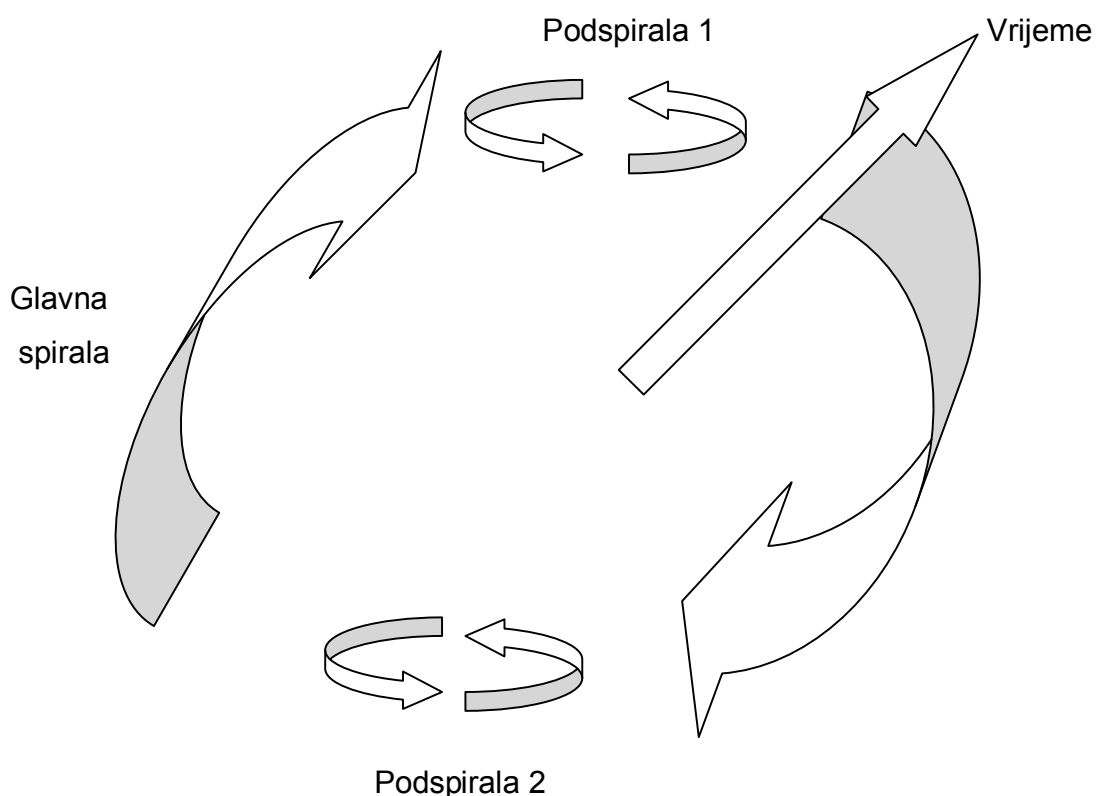


Slika 5. Dvoslojna arhitektura s jednim zajedničkim skladištem podataka [3]

Skladišta podataka u tradicionalnoj dvoslojnoj arhitekturi su velika i složena, što stvara određene poteškoće u njihovoj primjeni.

2.4. Spiralni pristup razvoju sustava poslovnog izvješćivanja

Osnovna karakteristika spiralnog sustava je sposobnost prilagođavanja novim poslovnim uvjetima u okolini gdje su ti sustavi prisutni. Temeljna ideja na kojoj se temelji razvoj sustava poslovnog izvješćivanja proizlazi iz čovjekova životnog učenja. Uz pravilan dotok informacija radi kvalitetne obrade i pružanja relevantnih podataka, također je vrlo važno stečeno iskustvo proizašlo iz povijesnih obrada i poslovnih događaja. Iterativnim procesom se može prikazati svaki pojedinačni slučaj prikupljanja i obrade podataka te sugeriranja rješenja. Iz prethodnih činjenica je vidljivo da se sustavi poslovnog izvješćivanja mogu prikazati kao dva međusobno povezana spiralna procesa, i to jedan osnovni sačinjen od niza manjih spiralnih procesa [3]. Slika 6. prikazuje metodologiju razvoja sustava poslovnog izvješćivanja.



Slika 6. Razvoj sustava poslovnog izvješćivanja [3]

Slika 6. prikazuje glavna spiralu koja objašnjava glavni poslovni ciklus, a sadržava niz manjih spirala. Velika (glavna) spirala predstavlja ciklus u području strateških pravaca tvrtke, dok se manji subspiralni procesi odnose na rješavanje niza problema na razini taktičkog i operativnog menadžmenta. Ulaz u novi glavni ciklus uvjetovan je promjenom strateških planova tvrtke s obzirom na novonastale stanje na tržištu, kao i uspjehom ostvarenih rezultata nakon završetka glavnog ciklusa.

Provedba spiralnog pristupa

Za provedbu spiralnog pristupa neophodna su četiri koraka [3]:

1. prepoznavanje problema,
2. pretprocesiranje podataka, uporaba metoda dubinskih analiza s ciljem otkrivanja zakonitosti među podacima,
3. primjena otkrivenih zakonitosti na testnom uzorku,
4. mjerenje rezultata.

Mjerenje rezultata se provodi na uzorku za testiranje, s obzirom na to da se prilikom primjene metoda dubinskih analiza osnovni uzorak dijeli u uzorak za treniranje i uzorak za testiranje. Postupak se ponavlja sve dok se ne dođe do zadovoljavajućih rezultata mjerenja. Stoga se poslije svakog utvrđenog nedozvoljenog odstupanja ulazi u novi spiralni ciklus, uzimajući u obzir stečena iskustva u prethodnoj spirali. Četiri navedena osnovna koraka mogu se uspješno primijeniti na dijelove unutar glavne spirale, koji se odnose na operativne i taktičke probleme a rješavaju primjenom metoda dubinskih analiza i skladištenja podataka, ali i na glavnu spiralu. Podaci na koje se oslanjaju sustavi poslovnog izvješćivanja najviše su iz unutarnjih (internih) izvora (transakcijski podaci), ali isto tako i na one iz vanjskih (eksternih) izvora podataka. Cjelokupna metodologija razvoja sustava poslovnog izvješćivanja povlači korijene iz metodologije umjetne inteligencije u kojoj se pokušavaju shvatiti i primijeniti ljudske osobine i metodologija učenja. U zaključku se može reći da spiralni sustav omogućava neprestane korekcije podsustava poslovnog izvješćivanja, ali i sustava poslovnog izvješćivanja.

Sustavi poslovnog izvješćivanja su za razliku od tradicionalnih sustava – dinamičniji, a time i zahtijevniji tijekom implementacije, razvoja te samog održavanja.

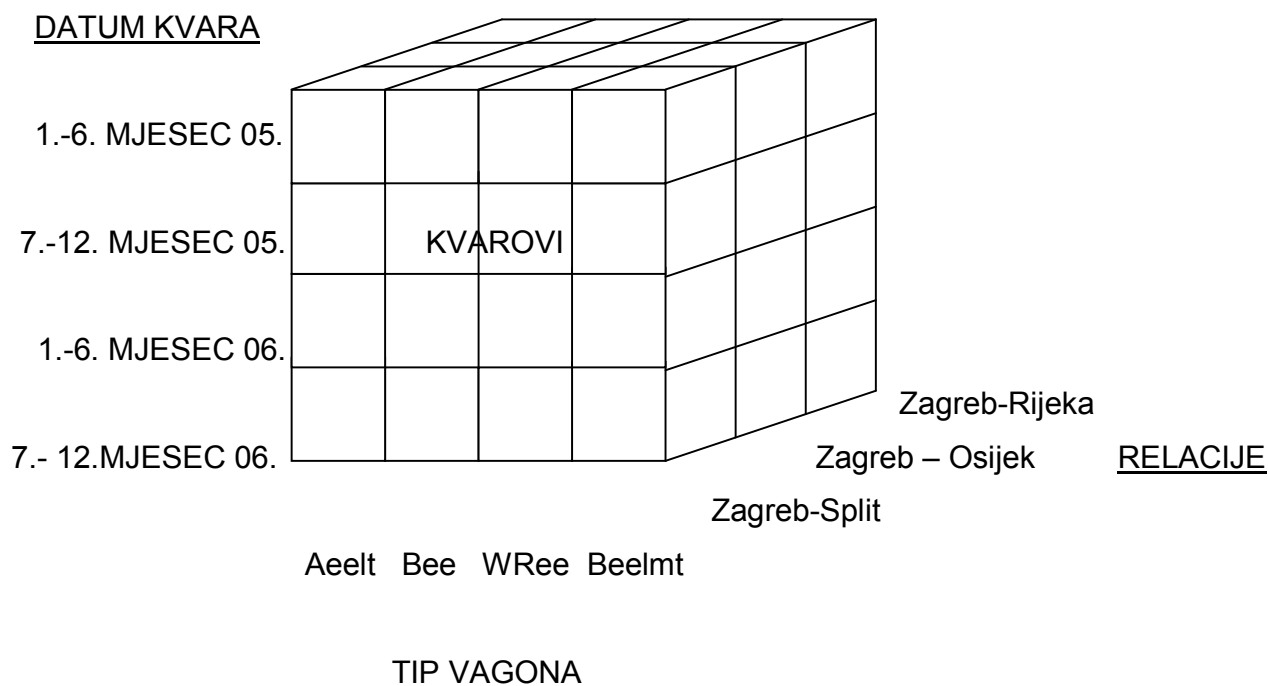
2.5. OLAP alati

OLAP je višedimenzionalni pogled na podatke u skladištu. Izraz je stvorio britanski matematičar E.F.Codd, koji je postavio temelje tzv. relacijskog računa (engl. Relational Calculus) i 1970. godine predstavio relacijski model podataka. Ako se želi analizirati prodaju proizvoda neke tvrtke, tada se može analizirati u vremenu, po gradovima, po regijama, po prodajnim mjestima, po proizvodima, po kupcima. Ideja OLAP-a se zasniva na tehnologiji koja višedimenzionalne kumulativne tablice kreira automatski i daje ih korisniku na raspolaganje, čime se stvara mogućnost kombiniranja različitih uvjeta koji, neovisno o svojem smislu, korisniku daju odgovor na postavljeni upit u realnom vremenu. Premda su razvijeni prije koncepta poslovnog izvješćivanja, OLAP alati se na neki način mogu smatrati začetnikom tog koncepta [3]. Sami OLAP alati obuhvaćaju širok spektar mogućnosti, od navigacija i pretraživanja, proračuna do ozbiljnih analiza kompleksnog modeliranja i vremenskih serija. Klasična OLAP struktura se sastoji od OLAP poslužitelja koji se nalazi između korisnika i skladišta podataka i povezanog s njime online.

2.5.1. Struktura OLAP sustava

Snaga OLAP sustava proizlazi iz mogućnosti strukturiranja podataka na način sukladan načinu na koji ljudi prirodno analiziraju pojave i procese. Najvažnije među tim strukturama su dimenzije i hijerarhija podataka. Skupovi podataka strukturirani po dimenzijama mogu se podvrgavati analizi po tim dimenzijama ili dimenzijskoj analizi. Dimenzija je pripadnost entiteta na kojem se podaci odnose na određenu klasu, skupinu ili kategoriju.

Na slici 7. dat je trodimenzionalni prikaz u obliku kocke. Promatrano je područje kvarova na pojedinim agregatima i uređajima (osjenčana ploha) klimatiziranih putničkih vagona, a tri relevantne dimenzije su tip vagona, datum kvara i relacija gradova na kojima je vagon prometovao.



Slika 7 .Trodimenzijski prikaz podataka

Promatrano je područje kvarova (osjenčana ploha), a tri osi su: tip vagona, datum kvara i relacija vožnje. Tablicom 2.1. dat je prikaz uporabe OLAP sustava danas u svijetu.

Na današnjem se tržištu mogu naći četiri glavne vrste OLAP alata:

- **višedimenzijski OLAP** – predstavlja skup proizvoda koji koriste vlastitu bazu podataka n-dimenzijske matrične strukture. Najveći proizvođači OLAP alata na ovom modelu su tvrtke APPLIX, COGNOS, ORACLE i SAS.
- **relacijski OLAP ili ROLAP** – koristi relacijski model kao osnovu svoje baze podataka. Najveći proizvođači ovih alata danas su IBM, MICROSOFT i ORACLE.
- **OLAP za stolna računala ili DOLAP** – je namijenjen primjeni na stolnim računalima

manjih i skromnijih memorijskih kapaciteta i mogućnosti obrade za razliku od velikih računala. Najistaknutiji proizvođači ove inačice OLAP alata je BUSINESS OBJECTS,

- **hibridni OLAP ili HOLAP** – predstavlja kompromisno rješenje višedimenzionalnog OLAP-a i relacijskog OLAP-a. Poznati proizvođači ovih inačica su APPLIX i ORACLE.

Tablica 2.1. Neke od uporaba OLAP-a [34]

Broj	Područje	Dimenzije	Činjenice	Svrha (namjena)
1.	Prodaja	Područja, trgovci, kupci, proizvodi...	Cijena, količina, iznos....	Multidimenzionalna analiza prodaje i profita
2.	Narudžbe (kupci)	Vrijeme, proizvodi, regija, kupci....	Prosječna cijena, kvantitet	Analiza cijene klijenata i sveobuhvatnih troškova
3.	Cijene	Vrijeme, regija, proizvodi.....	Prosječna cijena..	Usporedna analiza prosjeka
4.	Popis inventara	Vrijeme, područje, skladište podataka...	Prosječna cijena, kvantitet	Analiza inventara i zaliha
5.	Financijska izvješća	Vrijeme, indeksna kategorija...	Vrijednost...	Univerzalni oblik financijskog izvješća
6.	WEB statistika	Vrijeme, stranice, posjetitelji...	Vrijeme boravka na stranici, iznos.....	Analiza mrežnog prometa i posjetitelja na WEB-u
7.	Rezultati političkih izbora	Područja, kandidati stranke	Vrijednost i broj glasova	Analiza izborne statistike po regiji, kandidatima.....
8.	Proizvodnja	Vrijeme, mjesto, proizvodi..	Kvantitet, cijena	Analiza proizvodne strukture, usporedba po regijama...
9.	Potrošnja električne energije	Vrijeme, područje, potrošači...	Kvantitet, cijena	Analiza potrebe potrošača po regijama
10.	Plaće zaposlenih	Vrijeme, djelatnost....	Iznos....	Analiza potrošnje plaće, zaposlenici....

3. METODE DUBINSKIH ANALIZA (Data mining)

Pojam dubinske analize se razvijao godinama, a u stvari pod ovim pojmom prije pet do deset godina su mnogi podrazumijevali OLAP analize. Pojavom velike količine izvještaja i mogućnošću skeniranja podataka prema zemljopisnoj regiji, funkcionalnim područjima i proizvodnji, daju krajnjem korisniku bolju sliku o tome što se dešava u njegovoj organizaciji. Logička pitanja u svezi rudarenja podataka uključuju teme kao što su:

- postoji li statistička povezanost među varijablama u podacima i da li su pouzdani ?
- koliko je jaka povezanost među varijablama ?
- ukoliko se promjene određene varijable, koje odgovarajuće promjene možemo očekivati u varijabli ?
- što se može očekivati u budućnosti ?

Pojam *dubinske analize* se danas opisuje kao tehnologija aplikacije statističkih tehnika povezanih sa matematičkim formulama koje pokušavaju identificirati znakovitu povezanost između varijabli u podacima koje možemo koristiti za predviđanje i izvršenje osjetljivih analiza (što se događa sa zavisnim varijablama ukoliko su se promijenile jedna ili više nezavisnih varijabli). Neke od metodologija koje se pojavljuju u svijetu dubinske analize su: klasteriranje, segmentacija i klasifikacija, neuronske mreže, regresija, asocijacijska analiza. Dubinske analize se mogu opisati kao proces identifikacije neospornih, novih, potencijalno korisnih i razumljivih uzoraka (eng. patterns) i odnosa (eng. relationships) među podacima u skladištu podataka. Dubinska analiza je i proces analiziranja podataka iz različitih perspektiva i njihova pretvaranja u korisne i upotrebljive informacije tj. informacije koje će biti iskorištene za povećanje prihoda, smanjenje troškova i ostalo.

„Dubinska analiza je proces koji analitičaru primjenjuje tehnologiju podataka za određivanje statistički pouzdanih veza među varijablama“ [6].

Godinama se pojam dubinske analize povezivao sa raznim tipovima analitičkih pristupa problemu. Neovisno o vrsti metodologije, dubinske analize vuku svoje korijene iz tradicionalnih analitičkih tehnika. Dubinska analiza inkorporira analitičke procedure zasnovane na tradicionalnim statističkim, matematičkim, poslovnim i ekonomskim teorijama. Većina metodologija dubinske analize vuče korijene iz ekonometrije. Ekonometrija je analitička metodologija koja uključuje aplikaciju matematičkih i kvantitativnih metoda na podacima u konjunktiji sa tradicionalnom statistikom s pokusom na testiranju postojećih ekonomskih i business teorija (Gujarati 1988.) [6].

Slijedeća fraza stavlja ova obilježja u prigodan kontekst :

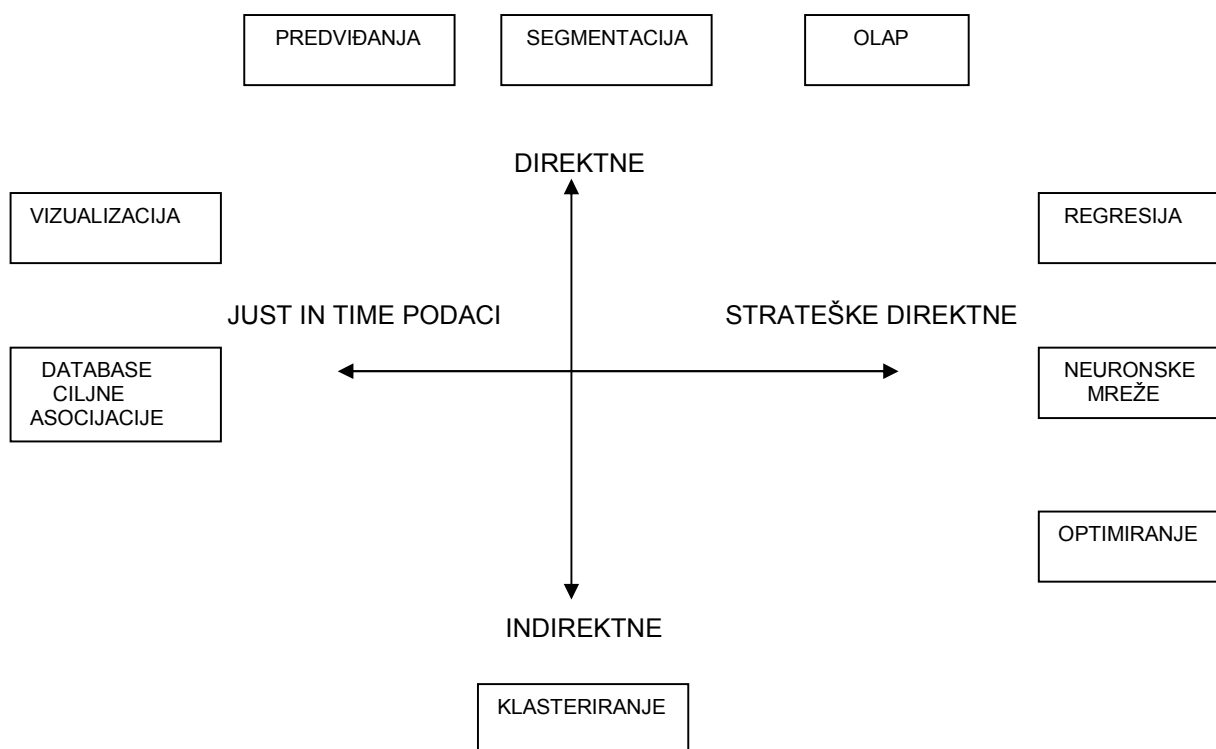
„U pregledu razvoja koncepta statističke analize ekonometrijskih modela, vrlo je lako zaboravljeno da je na početku 21. stoljeća glavno obilježje bilo da li je statistički pristup prikladan za analizu ekonomskih fenomena. Srećom, prepoznavanje znanstvenih vrijednosti sofisticiranih znanstvenih metoda u ekonomiji i business-u je pokopalo tu tvrdnju. Uporaba statistike u sofisticirane svrhe je zahtijevalo puno istraživanja u osnovnim konceptima ekonometrijskog modeliranja koja su danas priznata“ [6].

Tradicionalna metodologija koja se odnosi na metodologiju gore uključuje slijedeću proceduru:

- skupljanje podataka koje uključuje informacije relevantne za teoriju koja će se testirati.
- specifična kvantitativnost (matematički funkcionalni oblik) koja opisuje povezanost između objašnjavajuće i ciljne varijable.
- primjena odgovarajućeg statističkog testa za mjerenje jakosti ili ispravnosti kvantitativnog modela u potpori ekonomske ili poslovne teorije.
- optimiranje kvantitativnog modela prema statističkim rezultatima,
- izvlačenje zaključaka iz postupka.

Podaci mogu biti činjenice, brojevi ili tekstovi obrađeni računalom. Danas, organizacije pohranjuju velike količine podataka u različitim formatima i bazama podataka. To uključuje:

- *operacijske ili transakcijske podatke* - kao što su prodaja, troškovi, inventari, platni spiskovi, računi i sl.
- *neoperacijske podatke* - kao što su industrijska prodaja, prognozirajući podaci i makroekonomski podaci.
- *Metapodaci* - podaci o samim podacima, kao što su dizajnirane logičke baze podataka.



Slika 8. Pregled tehnologija dubinskih analiza [6]

Dubinske analize mogli bismo definirati kao pronalaženje zakonitosti u podacima. Ti podaci mogu biti organizirani u baze podataka, ali isto tako mogu biti i tekstualni podaci, nestrukturirani podaci proizašli iz web-a ili pak podaci organizirani u vremenske serije. Zakonitosti pronalazimo primjenom metoda koje svoje korijene vuku iz različitih područja kao što su: statistika, matematika, baze podataka, teorija informacija, teorija vjerojatnosti i umjetna inteligencija. Ovo je vrlo mlado područje, te postoji niz metodoloških pristupa problematici, kao i preferencije primjene metoda koje nagnju k određenom području ovisno o autorima koji obrađuju tu problematiku. Pojedine metode koje se deklariraju kao metode dubinske analize razvijene su u svojim rudimentarnim oblicima još u sedamdesetim i osamdesetim godinama 20.stoljeća. S obzirom na interdisciplinarnost ovog područja vrlo je teško povući jasnu granicu i proglasiti pojedine metode takvima koje služe isključivo za dubinske analize. Slika 8. prikazuje metodologiju otkrivanja znanja primjenom metoda dubinske analize.

Dubinske analize se sastoje od 5 elemenata:

- ekstrakcije, transformacije i punjenja transakcijskih podataka u sistem skladišta podataka,
- pohranjivanje i upravljanje podacima u multidimenzijском sustavu baze podataka,
- pristup snabdijevanje podacima poslovnim analitičarima i profesionalnim informacijskim tehnologijama,
- analiza podataka uporabom aplikacijskog software-a,
- prezentacija podataka u korisnom formatu, kao što su dijagrami ili tablice.

Koraci koji su potrebni za uspjeh procesa dubinskih analiza [6] :

Korak 1. Treba odlučiti što se želi znati

Dubinske analize pokazuju vrijednosti koje su primijenjene za rješavanje poslovnog problema. Poslovni problem koji treba biti riješen mora biti okvir u sadržaju pitanja koji će biti kvantitabilno odgovoreno. Npr., prodavač ne treba znati koji najbolji oglašivački medij

treba koristiti. Potrebno pitanje bi trebalo biti „koja je prednost oglašavanja putem televizije u odnosu na radio za neki proizvod?“

Korak 2. Određivanje relevantnog postupka mjerenja

Većina algoritama dubinskih analiza zahtijeva da je izvršenje mjerenja u brojkama (numerička vrijednost), dakle sve što je u vrijednostima - dolarima, metrima itd. Ukoliko su mjerenja postavljena sa odgovorima „da“ ili „ne“, potrebno je „da“ zamijeniti sa „1“ i „ne“ zamijeniti sa „0“.

Korak 3. Odlučiti kakav će biti oblik podataka

Pogledom u uzorak podataka, algoritam dubinskih analiza će usporediti svaku drukčiju zabilješku u bazi podataka. Broj stupaca će zahtijevati varijacije između algoritama. Segmentacija i klaster algoritmi zahtijevaju stotinu do tisuću podataka (zabilješki) da bi rezultat bio dobar. Regresija i neuronske mreže zahtijevaju daleko manje podataka. Slijedeći podaci zahtijevaju slične i usporedive formate (oblike) (oglašavanje potrošnje i prihoda mora odgovarati nekim mjesečnim periodima). Nemogu se uspoređivati tjedni i mjeseci, mjeseci i kvartali između varijabli u modelu. Tablica 3.1. daje prikaz oblika podataka.

Tablica 3.1. Prikaz nedovoljnog broja podataka za dubinske analize postupcima regresije ili neuronske mreže [6]

Godina	Mjeseci	Reklamacija troškova (vozači)	Reklamacija prihoda (vozači)
1999	Siječanj	100.000 kn	1.500.000 kn
1999	Veljača	120.000 kn	1.750.000 kn
1999	Ožujak	115.000 kn	1.800.000 kn
1999	Travanj	105.000 kn	1.900.000 kn
1999	Svibanj	128.000 kn	1.800.000 kn
1999	Lipanj	100.000 kn	1.500.000 kn
1999	Srpanj	120.000 kn	1.750.000 kn
1999	Kolovoz	115.000 kn	1.800.000 kn
1999	Rujan	180.000 kn	2.225.000 kn
1999	Listopad	105.000 kn	1.900.000 kn
1999	Studeni	128.000 kn	1.800.000 kn

Ukoliko nema dovoljno podataka, potrebno je:

- Dati više podataka. Ukoliko je to moguće, to je najbolje rješenje.
- Primijeniti više pristupa rudarenja podataka koji mogu raditi s manjom količinom podataka.
- Pokušati primijeniti double (dvostruki) roll-up. Ovo radi samo ako postoji istinska varijacija između dvije dimenzije.

Ukoliko postoji pregršt podataka (milijunski podaci), potrebno je:

- Imati uzorke podataka. To je jednostavan način, ali brz i efektivan.
- Podjela analize na nekoliko manjih analiza. Odvojene analize po npr. regiji ili demografskoj grupi.
- Postići unapređenje podataka. Npr. od nekoliko tjedana do mjeseci.

Korak 4. Identifikacija vodećih faktora

Vodeći faktori su varijable koji imaju utjecaj na izvršenje mjerenja (u slučaju klaster algoritama, to su varijable koje pomažu u podjeli podataka). Granica broja vodećih faktora koju možemo uključiti u proces varira prema algoritmu i sadržaju podataka. Dodatni faktori će rezultirati pogoršanjem rezultata. Mnogi sustavi dubinskih analiza naročito regresijski bazirani modeli imaju problem ukoliko su dva ili više faktora linearno ovisni.

Korak 5. Dobivanje podataka

Ukoliko neki od podataka nisu na raspolaganju, potrebno je:

- Pokušati pronaći zamjensku varijablu,
- Vidjeti kako možemo napredovati bez nje,
- Početi skupljati tako da možemo poboljšati analizu u budućnosti.

Ukoliko u jednom ili dva polja nedostaju neki podaci postoji nekoliko opcija:

- Ostaviti polje prazno,
- Umetnuti prosječnu vrijednost. Prosječna vrijednost je neutralna i neće odlučiti o rezultatu. Izbjegavanjem umetanja nule mnogi algoritmi će doslovce interpretirati broj, što vjerojatno nije to što se želi postići.
- Ukoliko su podaci vremenske serije, može se interpolirati ili koristiti prethodna zabilježena vrijednost.
- Mogu se predvidjeti nedostajuće vrijednosti vodećeg faktora na osnovu drugih izgradnjom regresijskog modela zasnovanog na preostalim podacima.

Korak 6. Vizualni pregled podataka

Na osnovu izrade karte svih podataka potrebno se je zapitati da li izgleda točno i ispravno. Pretraživač (miner) ima opciju korekcije ili ignoriranja podataka. Ukoliko tu postoje greške, odgovarajući model će biti nepouzdan.

Korak 7. Dubinske analize

Korak 8. Pregled rezultata

Ukoliko sustav dubinske analize može pronaći uzorke u podacima, potrebno se je pitati da li su pouzdani i vjerodostojni. Ukoliko nisu ili se nemogu pronaći uzorci, potrebno je:

- Pitati se da li različitost podataka ili različite transformacije postojećih podataka mogu dati više informativnih rezultata.
- Ne pokušati modificirati algoritam pretraživanja.
- Ukoliko uzorak ne može biti pronađen, mora se prihvatiti da ga nema.

Tablicom 3.2 dat je prikaz povijesnih koraka u razvoju dubinskih analiza

Dubinske analize su izrazito multidisciplinarno područje. Ono obuhvaća područja baza podataka ekspertnih sustava, teorije informacija, statistike, matematike, logike, te čitav niz pridruženih područja i potpodručja. Kvaliteta podataka presudan je faktor o kojemu ovisi uspješnost dubinskih analiza. U kvaliteti podataka vrlo značajnu ulogu, osim izvora, imaju i postupci čišćenja i pretprocesiranja podataka.

Tablica 3.2. Prikaz povijesnih koraka u razvoju dubinskih analiza [32]

EVOLUCIJSKI KORACI	PITANJA POSL- OVNIH LJUDI	OMOGUĆENE TEHNOLOGIJE	PROIZVOĐAČI	KARAKTERISTIKE
SKUPLJANJE PODATAKA 1960. te godine	„kolika je moja ukupna dobit u zadnjih 5 godina?“	KOMPJUTERI, VRPCE, DISKOVI	IBM, CDC	RETROSPEKTIVNA STATIČKA ISPORUKA PODATAKA
PRISTUP PODACIMA (1980.tih godina)	„koje su bile prodajne količine u New Englandu prošlog ožujka?“	RELACIJSKE BAZE PODATAKA (RDBMS), STRUKTURIRANI JEZICI (SQL), QDBC	ORACLE, IBM, MICROSOFT	RETROSPEKTIVNA DINAMIČKA ISPORUKA PODATAKA NA REKORDNIM NIVOIMA
SKLADIŠTENJE PODATAKA I POTPORA ODLUČIVANJU (1990.te godine)	„koje su bile prodajne količine u New Englandu prošlog ožujka?“ Drill down u Bostonu.	OLAP, MULTIDIMENZIJSKE BAZE I SKLADIŠTA PODATAKA	PILOT, ARBOR, COGNOS, COMSHARE	RETROSPEKTIVNA DIN. ISPORUKA PODATAKA NA VIŠESTRUKIM NIVOIMA
DUBINSKE ANALIZE (DATA MINING)- danas	„što će se desiti slijedećeg mjeseca sa količinama prodaje u Bostonu?“	POSEBNI ALGORITMI MULTIPROC. I KOMPJUTORI, VELIKE BAZE PODATAKA	PILOT, IBM, SGI, LOCKHEAD	PERSPEKTIVNA PROAKTIVNA ISPORUKA IMFORMACIJA

Analitičar najviše vremena provede čisteći i pretprocesuirajući podatke, i to do 80% vremena utrošenih na analize, da bi tek 20% vremena nad tako pripremljenim podacima primjenjivao metode dubinskih analiza. Jedna od glavnih prednosti sustava poslovnog izvješćivanja očituje se u korištenju skladišta podataka kao izvora podataka za analize, što analitičaru znatno štedi vrijeme. No, analitičar mora ponekad izravno pristupiti izvornoj bazi podataka - ako neki podatak ne postoji u skladištu podataka, ili ako se ne pojavljuje na zadovoljavajućem stupnju granulacije.

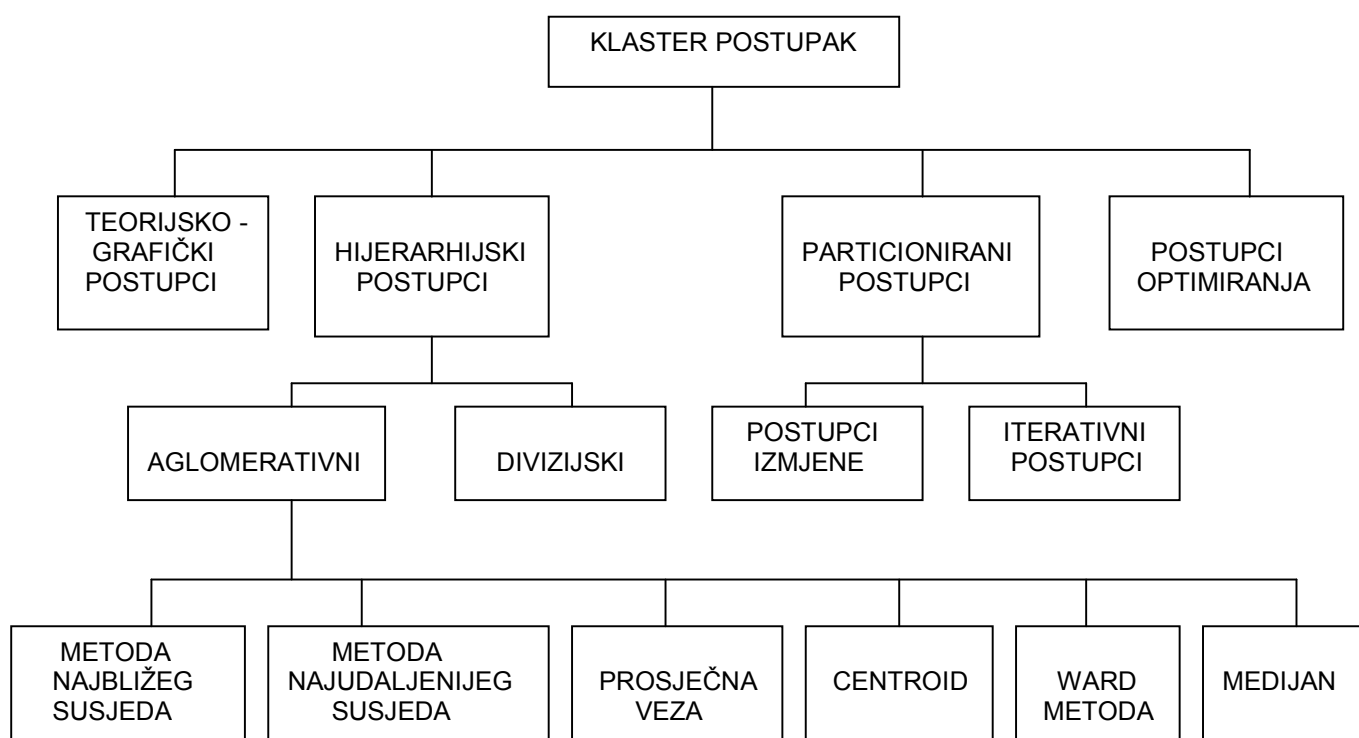
3.1. Metode dubinskih analiza

Postoji čitava lepeza metoda dubinskih analiza. Najznačajnije metode su:

- 1) *Klasifikacijske metode* - svrstavaju grupe u jednu od prethodno definiranih klasa - diskriminantna analiza, logistička regresija, stablo odlučivanja i neuronske mreže.
- 2) *Regresijske metode* - koje utvrđuju povezanost između varijabli (npr. prodaje) i predikcijskih varijabli (npr. prihod kupca, spol, broj članova kućanstva) - linearna, multipla, multivarijatna, nelinearna i polinomna regresija.
- 3) *Metode klasteriranja* - svrstavaju grupe u klastere - klaster analiza, metoda multidimenzijanskog skaliranja, metoda glavnih komponenti.
- 4) *Metode sažimanja* - računanje suma, prosjeka, frekvencija, postotaka.
- 5) *Metode za vizualizaciju i istraživačku analizu podataka* - su interaktivne metode za grafičko prikazivanje veza među varijablama, te uočavanje neobičnih podataka i oblika.
- 6) *Metode za modeliranje zavisnosti među varijablama* - faktorska i komponentna analiza.
- 7) *Metode za analizu vremenskih serija.*

3.1.1. Klasteriranje (metoda klastera)

Klasteriranje se može definirati kao grupiranje ili objedinjavanje objekata sličnih karakteristika. Algoritmi za klasteriranje pokušavaju pronaći sličnosti unutar zadane populacije koristeći zadani skup atributa. Sličnost između članova populacije nad kojom se obavlja analiza računa se primjenom funkcije udaljenosti kao što je Euklid udaljenost. Odabir i podjelu klaster algoritama prikazuje slika 9.



Slika 9. Pregled klaster algoritama [8]

Statistička metodologija

Klaster analiza omogućuje tri različite matematičke metrike za mjerenje udaljenosti između dvije veličine x i y :

a.) Kvadratna Euklidova udaljenost

$$d(x, y) = \sum_{i=1}^p (x_i - y_i)^2 \quad (3)$$

gdje su :

$d(x, y)$ – udaljenost točaka x, y

x_i, y_i – koordinate točaka k, l na p -toj dimenziji ($i = 1, 2, \dots, p$)

b.) Euklidova udaljenost

Kod Euklid metrike udaljenost između dvije točke prema njihovoj najkraćoj udaljenosti se izračunava formulom:

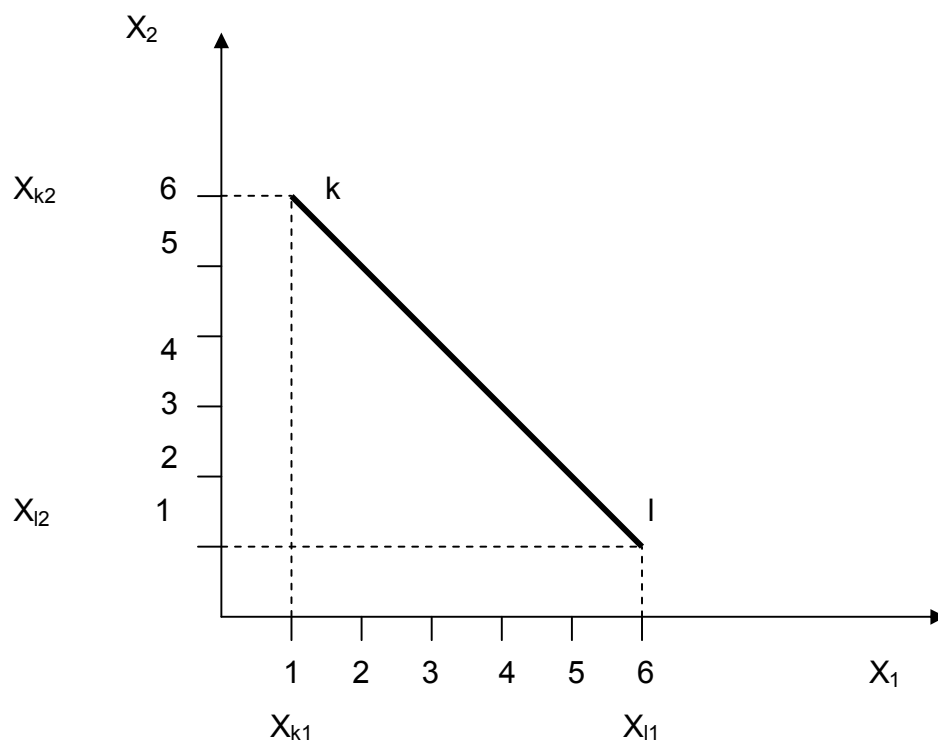
$$d(x, y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (4)$$

gdje su :

$d(x, y)$ – udaljenost točaka x, y

x_i, y_i – koordinate točaka k, l na p -toj dimenziji ($i = 1, 2, \dots, p$)

Slika 10. prikazuje Euklid udaljenost između koordinata (1,6) i (6,2)



Slika 10. Euklid udaljenost [8]

Prema slici 10. udaljenost točke k sa koordinatama $k(1,6)$ i $l(6,2)$ iznosi:

$$d(k,l) = \sqrt{(1-6)^2 + (6-2)^2} = 6,4$$

c.) City - Block udaljenost

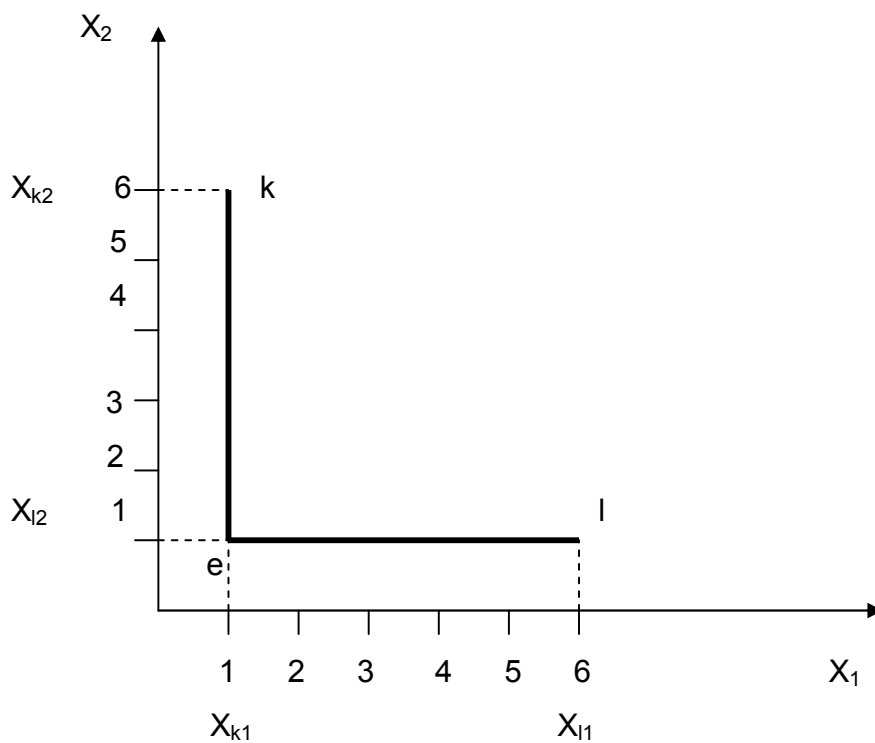
Kod City - Block metrike udaljenost između dvije točke se izračunava kao suma apsolutnih udaljenosti između točaka formulom:

$$d(x,y) = \sum_{i=1}^p |x_i - y_i| \quad (5)$$

gdje su :

$d(x,y)$ – udaljenost točkaka x, y

x_i, y_i – koordinate točkaka k, l na p -toj dimenziji ($i = 1, 2, \dots, p$)



Slika 11. City Block udaljenost [8]

Udaljenost točke k sa koordinatom $(1,6)$ i l sa koordinatom $(6,2)$ iznosi:

$$d_{ke} = |6 - 2| = 4, \quad d_{el} = |1 - 6| = 5, \quad d_{kl} = |4 + 5| = 9$$

d.) Udaljenost prema Minkovskom

Metrika koja obuhvaća u sebi metrike City Blocka i Euklida zove se Minkovski metrika. Za dvije točke, udaljenost će se izračunati kao apsolutna vrijednost diferencija koordinatnih vrijednosti preko svih dimenzija. Te dimenzije će biti potencirane konstantnim faktorom c .

$$d_{kl} = \left[\sum_{r=1}^R |x_{kr} - x_{lr}|^c \right]^{1/c} \quad (6)$$

gdje je:

d_{kl} – je udaljenost točki k, l

X_{kr}, X_{kl} – koordinate točki k, l na r -toj dimenziji ($r = 1, 2, \dots, R$)

c – Minkovski konstanta

$c=1$ je za City Block metriku, $c=2$ je za Euklid metriku

Osnovne dva tipa klaster metoda su:

3.1.1.1. Hijerarhijska metoda

Aglomerativne hijerarhijske klaster metode počinju sa postavljanjem svakih od veličina u posebne (separativne) klastere i kombinacijom klastera zasnovanog na njihovim udaljenostima od drugih klastera. Proces se nastavlja sve dok se ne definira željeni broj klastera.

a.) metoda najbližeg susjeda - definira udaljenost između dva klastera kao minimum udaljenosti između svakog člana jednog klastera i članova drugih klastera.

b.) metoda najudaljenijeg susjeda - definira udaljenost između dva klastera kao maksimum udaljenosti između svakog člana jednog klastera i člana drugog klastera.

- c.) centroid metoda** - definira udaljenost između dva klastera i udaljenost između centroida svakog klastera gdje je centroid određen kao prosječna vrijednost svake varijable podijeljen sa svim članovima klastera.
- d.) Median** - definira udaljenost između dva klastera kao udaljenost između medijana svakog klastera, gdje je median određen kao medijan veličina svake varijable podijeljen sa svim članovima klastera.
- e.) Prosjek grupe** - definira udaljenost između dva klastera kao prosječna udaljenost između svih članova jednog i drugog klastera.
- f.) Wardova metoda** - definira udaljenost između dva klastera u uvjetima povećanja devijacije sume kvadrata oko srednje vrijednosti klastera koja će se dogoditi ukoliko su spojena dva klastera.

Osnovna formula za izračun udaljenosti (koeficijent fuzije) u aglomeracijskom postupku nakon izračuna udaljenosti (Euklidom, City blokom, Euklid kvadratom ili Minkovskim) između grupa objekata i oblikovanja klastera u dendogramu je prema [8]:

$$D(R, P+Q) = A \cdot D(R,P) + B \cdot D(R,Q) + E \cdot D(P,Q) + G \cdot |D(R,P) - D(R,Q)| \quad (7),$$

gdje su: $D(R,P)$ - udaljenost između grupa R i P

$D(R,Q)$ - udaljenost između grupa R i Q

$D(P,Q)$ - udaljenost između grupa P i Q

A,B,E i G – konstante zadane u ovisnosti prema postupku u tablici 3.3.

Tablica 3.3. Formule za izračun udaljenosti klastera (koeficijent fuzije) pri odabranom aglomerativnom postupku [8]:

Postupa	Konstanta				Formule udaljenosti, prema jednadžbi (7)
	A	B	E	G	
Pojedinačna veza (single link.)	0,5	0,5	0	-0,5	$0,5 \cdot \{ D(R,P) + D(R,Q) \} - D(R,P) - D(R,Q) $
Kompletna veza (complete linkage)	0,5	0,5	0	0,5	$0,5 \cdot \{ D(R,P) + D(R,Q) \} + D(R,P) - D(R,Q) $
Prosječna veza (bez tež. koefj)	0,5	0,5	0	0	$0,5 \cdot \{ D(R,P) + D(R,Q) \}$
Prosječna veza (s tež. koef.)	$\frac{N_p}{N_p + N_q}$	$\frac{N_q}{N_p + N_q}$	0	0	$\frac{1}{N_p + N_q} \cdot \{ N_p \cdot D(R,P) + N_q \cdot D(R,Q) \}$
Centroid	$\frac{N_p}{N_p + N_q}$	$\frac{N_q}{N_p + N_q}$	$-\frac{N_q \cdot N_p}{N_p + N_q}$	0	$\frac{1}{N_p + N_q} \cdot \{ N_p \cdot D(R,P) + N_q \cdot D(R,Q) \} - \frac{N_p \cdot N_q}{(N_p + N_q)^2} \cdot D(P,Q)$
Median	0,5	0,5	-0,25	0	$0,5 \cdot (D(R,P) + D(R,Q)) - 0,25 \cdot D(P,Q)$
Ward	$\frac{N_r + N_p}{N_p + N_r + N_q}$	$\frac{N_r + N_q}{N_p + N_r + N_q}$	$\frac{N_r}{N_p + N_r + N_q}$	0	$\frac{1}{N_p + N_r + N_q} \cdot \{ (N_r + N_p) \cdot D(R,P) + (N_r + N_p) \cdot D(R,Q) - N_r \cdot D(R,Q) \}$

U tablici 3.3. su navedene oznake za:

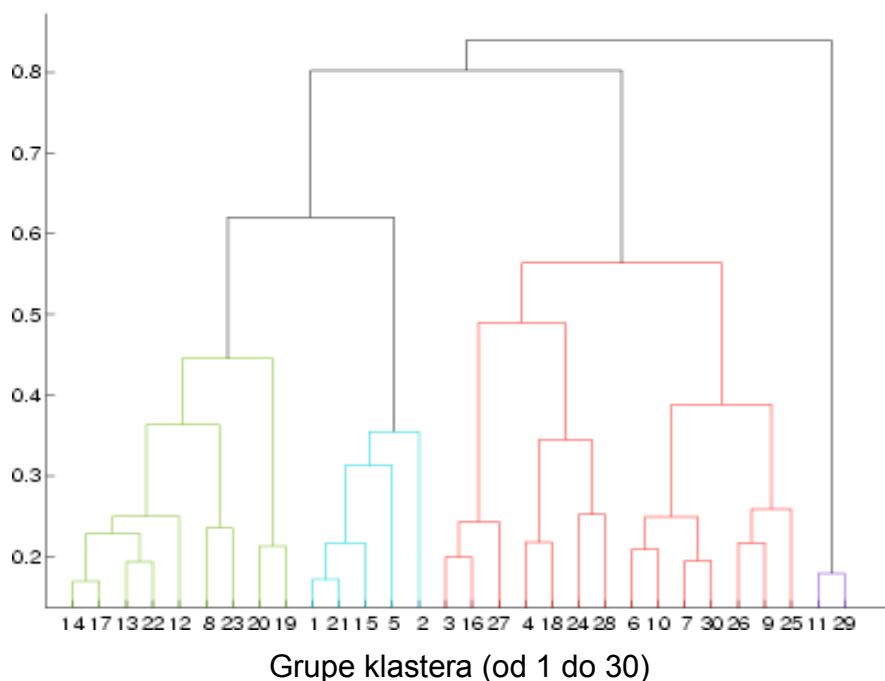
N_r – broj objekata u grupi R

N_p – broj objekata u grupi P i

N_q – broj objekata u grupi Q.

Hijerarhijski se klasteri mogu prikazati dendogramima prema slici 12.

Udaljenost



Slika 12. Primjer dendograma

Primjenom metode klasteriranja osnovni skup podataka može se podijeliti u manje segmente i svaki se može promatrati kao zasebna cjelina koja u sebi krije pravilnosti kakve se mogu otkriti primjenom nekih drugih metoda dubinskih analiza. Klasteri se mogu oblikovati na osnovu čitavog niza varijabla kao što su dob, spol, prosječna plaća i slično.

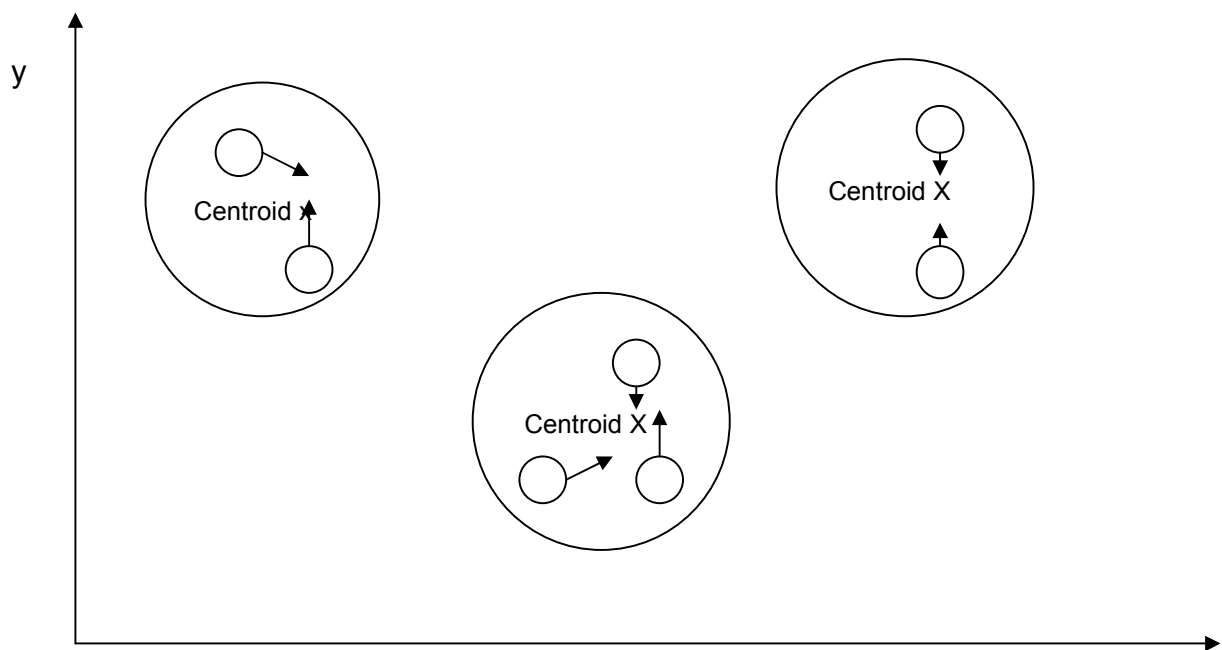
3.1.1.2. K-means klasteriranje

Osnovna karakteristika ove metode se svodi na dijeljenje osnovnog uzorka na k - segmenata, pri čemu svaki od segmenata sadrži n -sličnih elemenata. Na temelju funkcije udaljenosti algoritam procjenjuje sličnost elemenata.

Algoritamski se k-means klasteriranje može prikazati na slijedeći način [3]:

1. Odabir proizvoljno k-segmenata
2. Određivanje središta (centroida) za svaki od k-segmenata
3. Ponavljanje:
 - svojim najbližim klasterima (izračun se vrši na osnovu centralnih vrijednosti-centroida) pridruživanje pomoću funkcije udaljenosti svih elemenata populacije,
 - proračun nove vrijednosti središta klastera (centroida) za svaki klaster pojedinačno kao prosječne vrijednosti objekata sadržanih unutar svakog klastera.
 - ponavljanje sve dok se mijenjaju vrijednosti klastera (vrijednosti centroida)

Grafički se ova metoda može prikazati kao na slici 13.

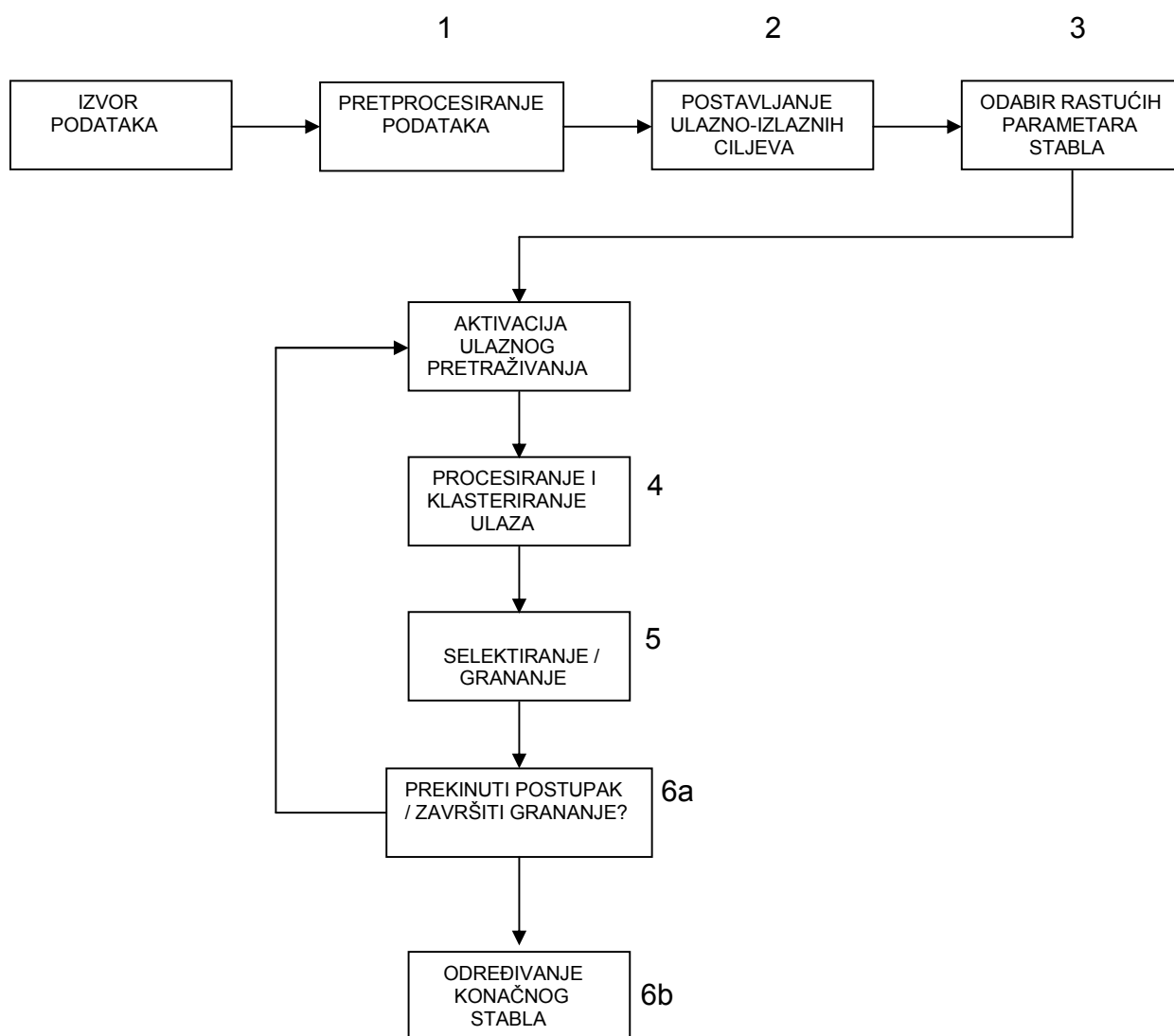


Slika 13. K-means algoritam [3]

Slika 13. prikazuje težnju vezivanja elemenata u iterativnim procesima uz centralne vrijednosti klastera. Klasteriranje svakom slogu dodjeljuje vrijednost pripadnosti klastera, te opcionalno pridružuje vrijednost udaljenosti od centra klastera.

3.1.2. Stablo odlučivanja

Stablo odlučivanja (decision tree) je metoda dubinske analize koja ima korijene u tradicionalnim statističkim disciplinama kao što je linearna regresija. Stabla odlučivanja imaju udio u polju kognitivnih znanosti čiji su „ proizvod “ neuronske mreže. Prva stabla odlučivanja su modelirana nakon bioloških procesa (Belson 1956.) [7]. Stabla odlučivanja su jednostavni, ali snažno djelotvorni oblici analize multivarijabli. Slika 14. prikazuje podprocese u stvaranju jednog stabla odlučivanja.



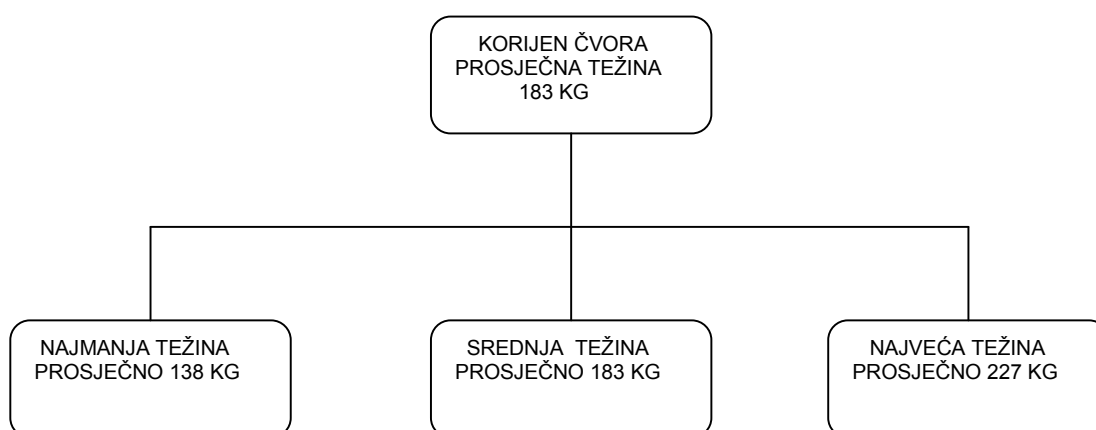
Slika 14. Podprocesi u stvaranju stabla odlučivanja [7]

Šest koraka za oblikovanje stabla odlučivanja su:

1. preprocesiranje podataka za alate grananja stabla odlučivanja,
2. postavljanje ulaznih i ciljnih karakteristika modeliranja,
3. određivanje parametara razvoja stabla odlučivanja,
4. klasteriranje i procesiranje svake ulazno oblikovane grane polja,
5. određivanje ponuđenih grana stabla odlučivanja,
6. kompletiranje oblika i sadržaja krajnje oblikovanog stabla odlučivanja
 - a.) prekid grananja, daljnje grananje ili iteracija stabla odlučivanja
 - b.) određivanje konačnog oblika stabla odlučivanja

Stablo odlučivanja jest klasifikacijski algoritam u formi stablaste strukture u kojoj se razlikuju dva tipa čvorova povezanih granama:

- 1.) a *krajnji čvor* ("*leaf node*") - kojim završava određena grana stabla. Krajnji čvorovi definiraju klasu kojoj pripadaju primjeri koji zadovoljavaju uvjete na toj grani stabla;
- 2.) a *čvor odluke* ("*decision node*") - ovaj čvor definira određeni uvjet u obliku vrijednosti određenog atributa (varijable), iz kojeg izlaze grane koje zadovoljavaju određene vrijednosti tog atributa.



Slika 15. Prikaz primjera podjele stabla odlučivanja fizikalnih mjerenja [7]

Postoji čitav niz algoritama na temelju kojih se može oblikovati stablo odlučivanja. Najpoznatiji su ID3, C 4.5, CHAID, CRT i QUEST.

a.) ID3 algoritmi

ID3 algoritam je oblik strojnog učenja razvijen od strane znanstvenika Rossa Quinlana (University of Sydney, Australia) [7]. ID3 je skraćenica od izraza „Interactive Dichtomizer“ ili Interaktivna dihotomija (podjela). Ovaj algoritam izračunava veličinu uspjeha određivanja strukture stabla odlučivanja, a baziran je na konceptu entropije, razvijen od strane Claude Shannona za opis količine informacija koja je sadržana u signalu. Stoga je ovaj koncept originalno korišten za opis kapaciteta raznovrsnih komunikacijskih kanala koji mogu biti korišteni u stablu odlučivanja za opis komunikacijskih kapaciteta ulaznih veličina i rezultirajućim ograncima u stablu odlučivanja.

ID3 algoritam ima nekoliko nedostataka u odnosu na ostale algoritme poput CHAID-a ili AID-a jer ne može kombinirati slične vrijednosti u ograncima. Ako je na ulazu bilo 3 vrijednosti, proizvest će granu sa 3 čvora, s 5 vrijednosti - 5 čvorova itd. Poboljšanje ID3 algoritma kulminiralo je razvojem C 4.5 metode stabla odlučivanja. C 4.5 algoritam rješava probleme identificirane u originalnim AID i ID3 implementacijama i obrađuje zajedno sa objema kvalitativne i kvantitativne atribute i nedostajuće vrijednosti.

b.) AID algoritmi

AID (Automatic Interaction Detection) pristup problemu je razvijen od strane Morgana i Sonquista [7]. Ovaj je program postao prvi u uporabi za stabla odlučivanja u SAD-u. Vrste problema koje su bile razmatrane programom AID su:

- „neistinite“ relacije povezanosti (npr. prikaz strukture u slučajnim greškama podataka),
- predrasude selekcija ulaza ili predvidljivih veličina,

- AID ne zna kada treba prekinute grananje te oblikuje podjelu (cijepanje) stabla odlučivanja na niže ekstremitete gdje je na raspolaganju nekoliko zabilježenih podataka.

Jedan od prvih pomoćnih sredstava za adresiranje problema sa AID algoritmom je predložen od strane Kass-a (1970.) [7]. Kass je sugerirao uporabu statističkih testova i Bonferronijevih regulacija.

c.) CRT (Classification and Regression Trees)

CRT je tehnika stabla odlučivanja razvijena od Briemanna, Friedmana, Ohlsena i Stonea (1984.) a korištena je za klasifikaciju kompleta (seta) podataka. CRT slijedi originalne AID ciljeve, ali sa poboljšanjima posredstvom primjene opravdanosti i križne („cross“) opravdanosti.

Potpuna metodologija za razvoj i zaustavljanje grananja u CRT-u uključuje slijedeće:

- za kontinuirana polja, mogu se mjeriti i najmanji kvadrat i najmanja apsolutna devijacija.
- za kategorično ovisna polja moguće je koristiti GINI mjerenja. GINI mjerenja su nazvana prema talijanskom statističaru Corradu Giniju [5], a GINI indeks se koristi za mjerenja čistoće „čvorova“, gdje su $p_1, p_2, \dots, p_n$ relativne frekvencije svakog razreda u čvoru. GINI kriterij je predložen od strane Breiman-a:

$$\text{GINI}_{\text{čistoća}} = 1 - \sum_j^r p_j^2 \quad (8)$$

- moguće je izračunavanje troškova netočne raspodjele manjih stabala odlučivanja.
- određivanje stabla odlučivanja sa niskim troškovima kao opcijom.
- troškovi mogu biti regulirani.

Za CRT, ulazi su uglavnom nominalni ili intervalni. Ordinalni ulazi se razmatraju kao intervalni ulazi. Tradicionalni kriterij podjele obuhvaća slijedeće:

- Za intervalne ciljeve, smanjenje varijance i smanjenje najmanje apsolutne devijacije,
- Za nominalne ciljeve uporaba GINI metode.

d.) CHAID metoda (Chi-square Automatic Interaction Detection)

CHAID metoda je pomoć tradicionalnim statističkim testovima signifikacije za formiranje i oblikovanje grupa granica koje određuju vrijednosti ulaza koji oblikuju grane stabla odlučivanja. Tradicionalno, test signifikacije je napravljen oko nul-hipoteze. Kada se uspoređuju distribucije dvije ili više grupa u setu podataka, statističko skupljanje numeričkih evidencija karakterizira dvije ili više grupa i tada se postavlja pitanje „da li su razlike u veličinama unutar grupa toliko velike da nul razlike mogu biti odbijene i neodržive?“. Tablica 3.4. prikazuje usporedbe nekih od algoritama stabla odlučivanja.

Tablica 3.4. Usporedbe algoritama stabla odlučivanja AID-a, CHAID-a, XAID-a i CRT-a [7]

Tehničko obilježje	AID	CHAID	XAID	CRT
Ciljne varijable – kontinuirane	X		X	X
Ciljne varijable – kategorične		X		X
Ciljne varijable – numeričke				X
Kriterij podjele prilagođen broju kategorija u predviđanju		X	X	
Određivanje prestanka grananja u stablu odlučivanja pri nekoliko promatranja	X	X	X	X
Određivanje prestanka grananja u stablu ukoliko grananja nisu znakovita		X	X	
Kriterij za podjelu čvorova- za sve moguće 2-way (2-načina) podjele, izaberi jedan koji objašnjava najveće varijacije	X			X
Kriterij za podjelu čvorova- za sve moguće multi-way (više načina) podjele, izaberi jedan koji objašnjava najveće varijacije		X	X	

Prednosti i slabe strane metode stabla odlučivanja

Prednosti metode stabla odlučivanja su:

- sposobnost generiranja razumljivih modela;
- relativno mali zahtjevi na računalne resurse (vrijeme i memorija);
- sposobnost korištenja svih tipova atributa (kategorički, numerički);

- tabla odlučivanja jasno odražavaju važnost pojedinih atributa za konkretni klasifikacijski odnosno predikcijski problem.

Slabe strane metode stabla odlučivanja

- metode stabla odlučivanja su manje prikladne za probleme kod kojih se traži predikcija kontinuiranih vrijednosti ciljnog atributa;
- sklona su greškama u više-klasnim problemima sa relativno malim brojem primjera za učenje modela;
- u nekim situacijama generiranje stabla odlučivanja može bit računalno zahtjevan problem. Sortiranje kandidata za testiranje na čvorovima stabla može biti zahtjevno, kao i metode 'skraćivanja' stabla, kod kojih je često potrebno generirati velik broj stabala da bi odabrali ono koje je najbolje za klasifikaciju primjera određenog problema;
- stabla odlučivanja nisu dobro rješenje za klasifikacijske probleme kod kojih su regije određenih klasa 'omeđene' nelinearnim krivuljama u više-dimenzionalnom atributnom prostoru. Većina metoda stabla odlučivanja testiraju u svojim čvorovima vrijednosti jednog atributa, i time formiraju pravokutne regije u višedimenzionalnom prostoru.

Stabla odlučivanja spadaju u klasifikacijske metode dubinskih analiza. Često se prilikom analize kombiniraju s metodom klasteriranja. Osnovna im je zadaća određivanje varijabli i njihovih vrijednosti koje determiniraju neki skup pojava. Ova metoda može korisno poslužiti pri segmentaciji tržišta, pri čemu se kao rezultat analize mogu dobiti lako čitljiva pravila segmentacije. Prilikom oblikovanja stabla odlučivanja važno je pratiti procjene pouzdanosti modela, proizašle na temelju uzorka za treniranje i uzorka za testiranje.

3.1.3. Multidimenzijsko skaliranje (engl. MDS analysis)

Multidimenzionalno skaliranje (analiza) ima svoje začetak u psihometriji, gdje je predviđena kao tehnika za pomoć razumijevanju istraživanju ljudi, tj. sličnosti članova u grupi objekata. Znanstvenik Torgeson je predložio prvu MDS metodu i postavio izraze.

MDS prikazuje strukturu skupa objekata (grupa) iz podataka koji aproksimiraju udaljenost između parova objekata (grupa). Svaka grupa (objekt) je prikazana točkom u multidimenzionalnom prostoru. Točke su u prostoru tako raspoređene da udaljenosti između parova točaka imaju najjaču povezanost (bliskost) prema sličnosti između parova objekata. Dva slična objekta su prikazana sa dvije točke koje su približno „ zatvorene“ tj. jako bliske, dok su različiti objekti prikazani dvjema točkama koje su udaljene.

Koraci u provedbi MDS analize [8]:

- Korak 1. Mjerenje sličnosti između objekata
- Korak 2. Izbor modela udaljenosti između objekata
- Korak 3. Pronalaženje međusobne povezanosti (odnosa)
- Korak 4. Broj i interpretacija dimenzija
- Korak 5. Agregacija (skupljanje) sličnih objekata

1. Mjerenje sličnosti između grupa objekata

Postoje mnogobrojne metode u MDS analizi za istraživanje sličnosti između grupa (objekata). Neke od njih su:

a.) metoda rangiranja

Metoda rangiranja je klasična metoda u istraživanju sličnosti između grupa ili objekata. Npr. ukoliko postoji $K=10$ objekata koji sadrže 45 parova, tada je potrebno rangirati od 1 do N . Tada sličnost i (ne)sličnost sadrži:

Podatke za sličnost: 1= najnesličniji par
N= najslučniji par

Podatke za najnesličnost: 1= najsličniji par
N= najnesličniji par

Pr K –broju objekata dobije se $K * (K-1) / 2$ broja parova.

b.) metoda „prikopčanih“ točkaka

2. Izbor modela udaljenosti

Metrika za mjerenje udaljenosti među grupama (objektima) za sličnost je poznata metrika klaster analize (Euklidova, City Block i Minkowski)

3. Pronalaženje međusobnih odnosa između grupa

Postupak MDS-a : sličnost podataka tj. njihova nesličnost u_{kl} (za objekte k i l) u jednom prostoru s najmanjom mogućom dimenzionalnošću mora ispuniti (za udaljenost d_{kl}) *uvjet monotonije*:

Ako je $u_{kl} > u_{ij}$, tada je $d_{kl} > d_{ij}$

Za pronalazak konfiguracije, počinje se sa iteracijom, pri čemu se svaki slijedeći korak poboljšava. Mjera za „dobrotu“ jedne konfiguracije i kao ciljni kriterij za njeno optimiranje naziva se STRESS (**ST**andardized **RE**sidual **S**um of **S**quares- Standardizirani ostatak sume kvadrata) funkcija:

$$STRESS = \sqrt{\frac{\sum_k \sum_l (d_{kl} - \hat{d}_{kl})^2}{Faktor}} \quad (9)$$

gdje je:

d_{kl} – udaljenost između objekata k i l ,

$\hat{d}_{kl}$ - disparitet za objekte k i l .

Disparitet se računa za parove objekata (k,l) i (m,n) kao:

$$d_{kl} = d_{mn} = \frac{d_{kl} + d_{mn}}{2} \quad (10)$$

Za *Kruskal metodu* upotrebljavaju se STRESS formule 1 i 2:

$$STRESS1 = \sqrt{\frac{\sum_k \sum_l (d_{kl} - \hat{d}_{kl})^2}{\sum_k \sum_l d_{kl}^2}} \quad (11)$$

$$STRESS2 = \sqrt{\frac{\sum_k \sum_l (d_{kl} - \hat{d}_{kl})^2}{\sum_k \sum_l (d_{kl} - \bar{d})^2}} \quad (12)$$

Gdje je: $\bar{d}$ = srednja vrijednost udaljenosti među objektima

Za obje STRESS formule, vrijednosti su jako različite (formula STRESS 2 ima dvostruko veću vrijednost nego STRESS 1).

Kod Kruskovog algoritma za pronalazak optimalne konfiguracije, rukuje se metodički iterativnim postupkom optimiranja, koji je baziran na principu strmog porasta (postupak gradijenta). Ciljna konfiguracija će se iterativno tako dugo poboljšavati, sve dok se ne postigne minimalna vrijednost STRESS funkcije. Posredstvom formule (13), dat će se za vrijednost koordinate x_{kr} objekta (grupe) k na dimenziji r izračun nove vrijednosti koordinate, čija će se pozicija od objekta k poboljšati relativno ka objektu l .

$$x_{kr}^+(l) = x_{kr} + \alpha \left[1 - \frac{\hat{d}_{kl}}{d_{kl}} \right] (x_{lr} - x_{kr}) \quad (13)$$

gdje je: $r=1, \dots, R$

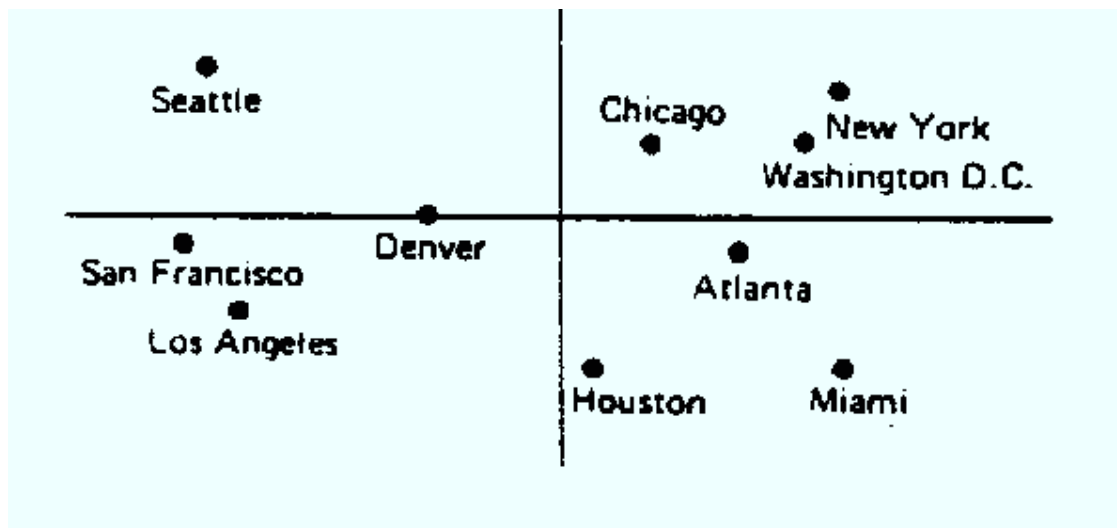
pri čemu je α poznat kao veličina koraka iteracije. Promjena vrijednosti koordinata se dobije samo u slučaju ukoliko postoji razlika između dispariteta $\hat{d}_{kl}$ i udaljenosti d_{kl} .

4. Broj i interpretacija dimenzija

Prostor opažanja će pored metrike biti određen i brojem dimenzija. Broj dimenzija mora odgovarati istinitom dimenzionalitetu prostora opažanja. Formalni kriterij za određivanje broja dimenzija može biti određen mjernom veličinom funkcije STRESS, čije rješenje mora biti numeričko. Stoga je potrebno naznačiti, da će se generalno smanjenjem vrijednosti STRESS povećavati broj dimenzija.

5. Grafički prikaz rješenja

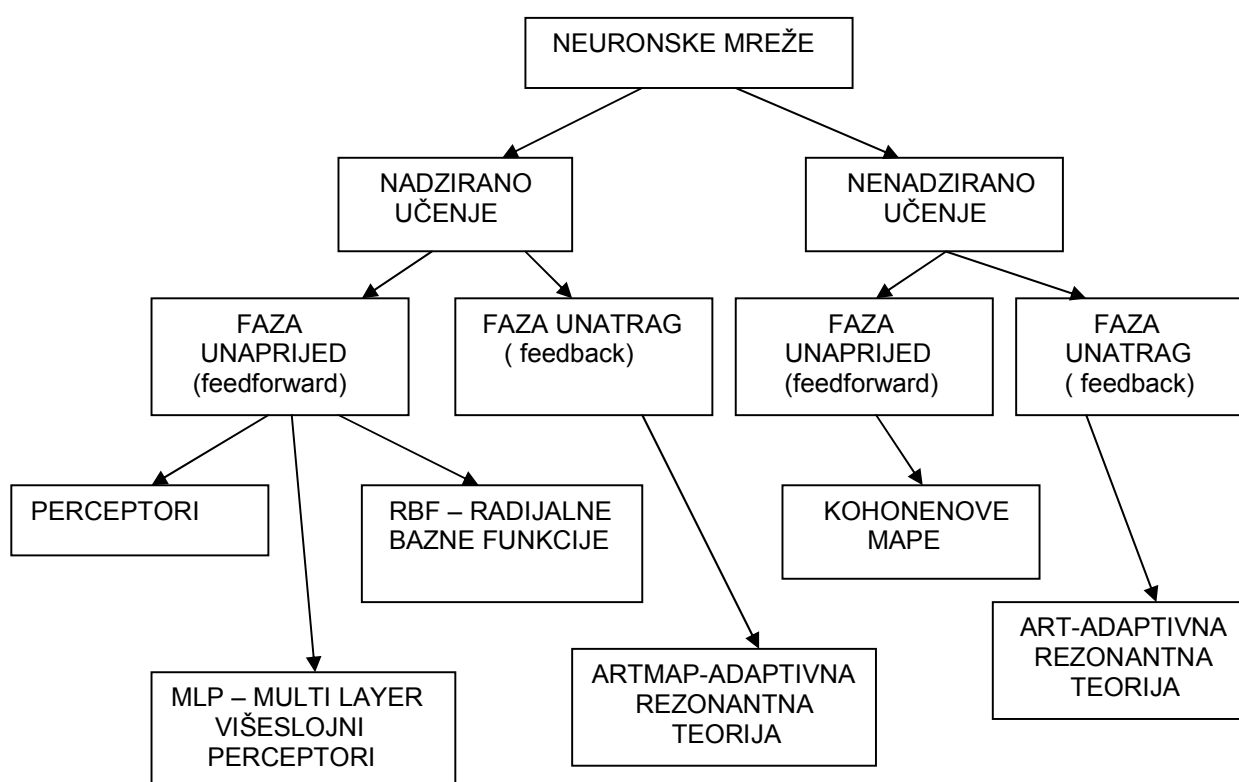
Na slici 16. dat je grafički prikaz rješenja jedne MDS analize na primjeru sličnosti gradova u Sjedinjenim američkim državama. Graf prikazuje da su klasteri gradova (Chicago, New York i Washington) u prvom kvadrantu koordinatnog sustava i gradova (San Francisco, Los Angeles i Denver) u trećem kvadrantu- razdvojeni, a međusobno su unutar sebe slični po nekim karakteristikama.



Slika 16. Grafički prikaz Multidimenzionalne analiza (MDS analysis)

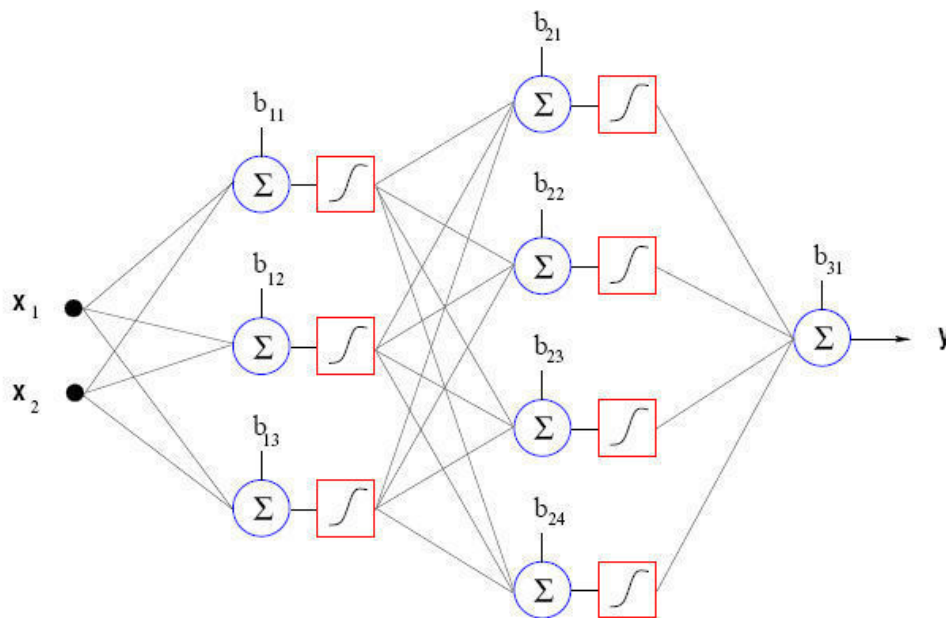
3.1.4. Neuronske mreže

Neuronske mreže su sustavi modelirani od procesirajućih elemenata tzv. neurona, po uzoru na mrežu ljudskih moždanih ćelija. Neuronske mreže su područje vještačke inteligencije, AI – Artificial Intelligence, strukturirane da rade na način kao što radi ljudski mozak. Ključni element ove vrste vještačke inteligencije je specifična struktura obrade informacija sastavljena od velikog broja međusobno povezanih elemenata obrade (neurona) koji zajednički rade na rješavanju problema. Te jedinice su povezane komunikacijskim kanalima (vezama). Neuronska mreža u informatičkim znanostima predstavlja jako povezanu mrežu elemenata koji obrađuju podatke. Podaci koji se ovim kanalima razmjenjuju su obično numerički. Na slici 17. dat je prikaz podjele neuronskih mreža.



Slika 17. Tipovi neuronskih mreža [8]

Osnovna saznanja o biološkom modelu upotrebljena su za oblikovanje matematičkog modela koji se koristi u praksi. U tom matematičkom modelu je osnovna jedinica dizajnirana po uzoru na biološki neuron. Jedinice kombiniraju ulaze u jedinstveni rezultat (najčešće funkcija sumiranja), koji zatim biva preusmjeren u funkciju transformacije koja kalkuliра izlaznu vrijednost i najčešće poprima vrijednost između 0 i 1. Kombinacijska i transferna funkcija zajedno čine aktivacijsku funkciju neurona [3]. Slika 18. prikazuje spomenute odnose.



Slika 18. Matematički model neurona [29]

Rezultat obrade aktivacijske funkcije leži između 0 i 1, a naziva se *threshold vrijednost* o kojoj ovisi aktivacija jedinice. Najučestalije funkcije transformacije su *linearna*, *hiperbolična* i *sigmoidna funkcija*. Linearna funkcija transformacije ima ograničenu primjenu i rijetko se koristi, jer se njezinim korištenjem u jedinicama neuronske mreže model mreže pretvara u alat za linearnu regresiju, dok su sigmoidna i hiperbolična nelinearne funkcije, pa su i rezultati njihove obrade također nelinearni. Razlika među ovim

dvjema funkcijama je u tome što sigmoidna funkcija producira izlaze u intervalu između -1 i 1, a hiperbolična funkcija u intervalu između 0 i 1 [3] .

Najčešće korištena funkcija je sigmoidna funkcija, a ona glasi:

$$S(x) = \frac{1}{1 + e^{-x}} \quad (14)$$

Ovdje je x rezultat kombinacijske funkcije koja je sumarna funkcija produkta težinskih koeficijenata (svaki neuron ima svoj težinski koeficijent) i ulaznih parametara.

Primjena neuronskih mreža

Na koje se vrste problema neuronske mreže mogu primijeniti?

- gdje postoji mnogo prošlih reprezentativnih primjera (barem oko 100, a poželjno je što više),
- gdje nisu poznata pravila dolaženja do odluke (black box), tj. eksperata nema ili nisu dostupni,
- gdje se varijable mogu kvantitativno izraziti (što je gotovo uvijek moguće),
- gdje standardne statističke metode nisu pokazale uspjeh, tj. pojava se ne može predstaviti nekim linearnim modelom,
- gdje je ponašanje pojave često neizvjesno, podaci nepotpuni, te traže robustan alat.

3.1.5. Neizrazita logika

1965. godine uveden je i koncept neizrazite logike u sklopu teorije neizrazitih skupova. Prva publikacija u kojoj se obrađuju pojmovi neizrazite logike (engl. Fuzzy Logic) je izdana od strane Lofti Zadeha [31]. Područje potpune afirmacije neizrazite logike je postala tehnika koja se koristi u širokom rasponu industrijskih i znanstvenih primjena, iako je početni razvoj ukazivao na potencijalnu primjenu u društvenim znanostima. Najznačajnija područja aplikacije neizrazite logike su: automatska regulacija, i sustavi za nadzor i dijagnostiku. Veliki porast zanimanja za neizrazitu logiku i porast istraživanja počinje prije desetak godina. Njemački matematičar G. Cantor u tradicionalnoj teoriji skupova koju je utemeljio, skup predstavlja kolekciju objekata koji dijele neko zajedničko obilježje. Definicija skupova je intuitivna, jer izrazi skup ili klasa predstavljaju sinonime, jednako kao i izrazi objekt ili element. Pripadnost objekta „klasičnom“ skupu egzaktno je određena: ili je objekt unutar skupa ili je izvan njega. Ako se pripadnost objekta skupu želi izraziti funkcijski, tada funkcija pripadnosti objekta x skupu A , $\mu_A(x)$ ima samo dvije vrijednosti i definirana je sa:

$$\mu_A(x) = \begin{cases} 1 & \text{za } x \in A \\ 0 & \text{za } x \notin A \end{cases}$$

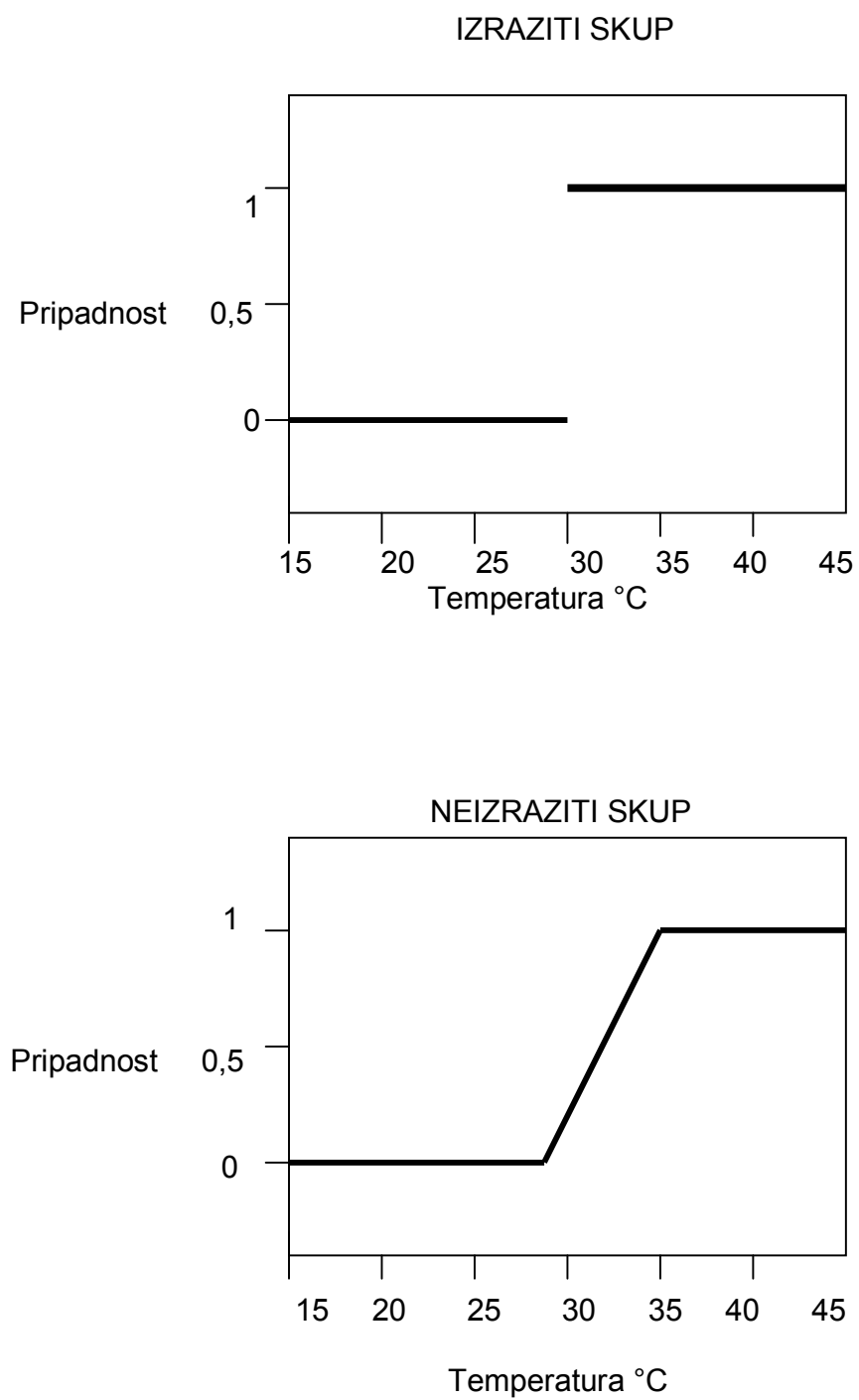
U podjelama objekata na skupove definicija pripadnosti nije prikladna, jer ponekad nije moguće odrediti zadovoljavajuću izrazitu granicu između objekata. Problem matematičkog opisa ljudske percepcije temperature okoline se navodi često kao tipičan primjer, gdje se po iznosu vrlo bliske numeričke vrijednosti mogu naći u različitim skupovima. Tako npr. osoba upitana za klasifikaciju temperature zraka, u skup VRUĆE može smjestiti sve temperature veće od 40°C, dok će intuitivno vrlo bliska vrijednost od 39,9°C pripadati drugom skupu (npr. toplo). Simbolički zapis izrazitog skupa VRUĆE preko funkcije pripadnosti bio bi slijedeći:

$$\mu_A(x) = \left| \begin{array}{ll} 1 & \text{za } x \geq 40^{\circ}\text{C} \\ 0 & \text{za } x < 40^{\circ}\text{C} \end{array} \right.$$

Za razliku od tradicionalnog pristupa, teorija neizrazitih skupova dopušta djelomičnu ili stupnjevitu pripadnost elementa skupu, a funkcijske vrijednosti funkcije pripadnosti se nalaze unutar intervala [0,1]. Pri tome vrijednost 0 označava da je objekt potpuno izvan skupa, a 1 da je potpuno unutar skupa, a bilo koja vrijednost između označava djelomičnu pripadnost. Tako bi se neizraziti skup vruće mogao definirati sa npr.:

$$\mu_{\text{vruće}}(x) = \left| \begin{array}{ll} 1 & \text{za } x > 34^{\circ}\text{C} \\ (x-26) / 8 & \text{za } 26^{\circ}\text{C} \leq x \leq 30^{\circ}\text{C} \\ 0 & \text{za } x < 26^{\circ}\text{C} \end{array} \right.$$

Ovakva definicija znatno je bliža načinu na koji čovjek doživljava i u govoru interpretira osjet temperature okoline. Tako bi prosječan promatrač, u određenoj klimatskoj zoni, označio temperaturu od 22°C-nije vruće, 28°C-malo je vruće, 33°C-prilično vruće, a 40°C-jako vruće. Razlika između funkcija pripadnosti neizrazitog i izrazitog skupa VRUĆE je ilustrirana na slici 19.



Slika 19. Funkcije pripadnosti izrazitog i neizrazitog skupa VRUĆE [31]

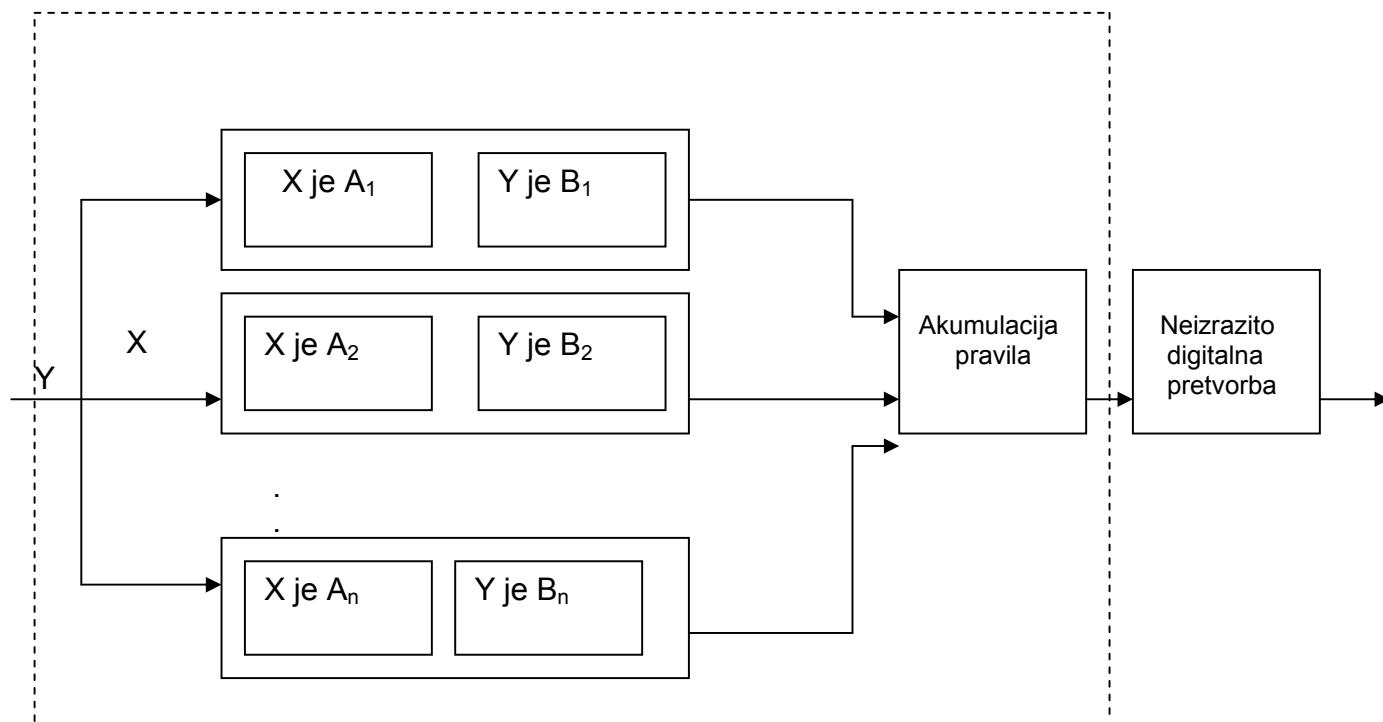
U okvirima teorije neizrazitih skupova pojmovi vruće, svjež, star, skup ili visok predstavljaju neizrazite skupove koji određuju jezične varijable ili u užem smislu vrijednosti jezičnih varijabli.

3.1.5.1. Neizraziti sustavi

Statistički ili dinamički sustav koji upotrebljava neizrazitu logiku i pripadajući matematički okoliš zove se neizraziti sustav. Ova široka definicija obuhvaća neizrazite sustave koji se primjenjuju u različitim područjima i za koje se mogu koristiti i drugi nazivi kao npr. neizraziti ekspertni sustavi (engl. Fuzzy expert system), neizraziti regulator (fuzzy logic controller), neizraziti model (fuzzy model), sustav neizrazitih pravila (fuzzy rule based system) [31]. Bez obzira na naziv ili područje primjene, neizraziti sustavi u osnovi sadrže tri komponente:

1. bazu pravila,
2. bazu podataka o funkcijama pripadnosti koje su zadane nad domenama ulaznih i izlaznih varijabli, te
3. mehanizam zaključivanja

Ilustracija neizrazitog sustava je prikazana slikom 20. Osnovna struktura u kojoj nije posebno istaknut blok digitalno - neizrazite pretvorbe je omeđena crtkanom linijom. S obzirom na mehanizam zaključivanja neizraziti sustavi se mogu načelno podijeliti u dvije grupe. Za razliku od relacijskog pristupa (koji se često naziva i globalni) u praktičnim primjenama dominira postupak lokalnog zaključivanja. Razlog leži u memorijskim i procesorskom ograničenjima računala, osobito izraženim u vrijeme prvih primjena neizrazite logike. U postupku globalnog zaključivanja koristi se neizrazita relacija koja obuhvaća čitavu bazu pravila, a točnost relacije, pohranjene u memoriji računala izravno ovisi o gustoći diskretizacije domena ulaznih i izlaznih varijabli [31].



Slika 20. Shematski prikaz neizrazitog sustava [31]

Ekspertni sustav se odnosi na problematike u medicini, farmaciji, ekonomiji, procjeni rizika i slično. Inženjeri znanja modeliraju ekspertno znanje na način da dijagnosticiraju varijable koje ulaze u obradu, utvrđuju kako te varijable smisleno povezati, oblikuju uzročno-posljedične veze, te opisuju pravila, a sve to u suradnji sa ekspertom iz domicilnog područja. Na sličan je način moguće izvršiti segmentaciju tržišta na osnovu ekspertnog znanja. Tako se, primjerice, u okviru sustava poslovnog izvješćivanja može razviti model koji će otkrivati rizične klijente za potrebe osiguravajućeg društva. Tablicom 3.5. date su neke od razlika između neuronskih mreža i ekspertnih sustava.

Tablica 3.5. Razlike između ekspertnih sustava i neuronskih mreža [31].

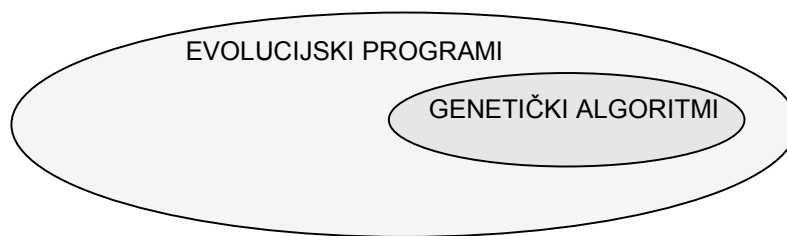
	EKSPERTNI SUSTAVI	NEURONSKE MREŽE
OSNOVE	simulacija i formalizacija ljudskog razmišljanja i logički procesi	oponašanje načina funkcioniranja ljudskog mozga
PRISTUP	makro-mozak je crna kutija: fokus-akcija, motivi.....	mikro-polazi od strukture mozga i načina njegova funkcioniranja
PROCESNA TEHNIKA	sekvencijska-znanje je kodirano u IF - THEN pravilima	paralelna-neuroni funkcioniraju paralelno
UČENJE- STATIČKO VS. DINAMIČKO	znanje se prikuplja van sustava i kodira u bazu znanja	učenje je integralni dio sustava- znanje se stvara u neuronima
METODE REZONIRANJA I UČENJA	deduktivno-opće znanje se kodira u bazu unaprijed (malo podataka)	induktivno- interna baza znanja se formira iz velike baze podataka
TEORIJSKA OSNOVA	matematička logika, objektno orijentiran pristup.....	statistika i stohastički procesi
ALGORITAM RJEŠENJA	povezivanje i ulančavanje-egzaktano uklapanje podataka u pravila	aproksimativno uklapanje- s podsjećanjem što je sustav već naučio
REPREZENTACIJA ZNANJA	eksplicitna-u obliku prepoznatljivom čovjeku	implicitna
INŽENJERSTVO ZNANJA	okrenuta prema ljudima-laki dizajn	okrenuta podacima-izbor modela i dizajn su zahtjevni

3.1.6. Genetički algoritmi

Razvoj genetičkih algoritama je značajno zabilježen posljednjih deset godina, te ima primjenu i daje prihvatljive rezultate u području učenja kod neuronskih mreža, pri traženju najkraćeg puta, problemima sličnim transportnom problemu, problemu određivanja parametara sustava, optimiranju upita nad bazom podataka, itd [35]. Živa bića se vremenom evolucije prilagođavaju uvjetima u prirodi, te je evolucija složen proces pretraživanja prostora rješenja. U genetičkim operatorima i selekciji se očituje povezanost evolucije kao procesa prirode i genetičkog algoritma - kao metode optimiranja. Kao što su uvjeti i stanje u prirodi primarni ključ selekcije nad nekom vrstom živih bića, tako je i funkcija cilja ključ selekcije nad populacijom rješenja u genetičkom algoritmu. U prirodi najveću vjerojatnost preživljavanja i razmnožavanja ima jedinka koja je najbolje prilagođena uvjetima i okolini u kojoj živi.

Genetički algoritam kao metoda optimiranja

Evolucijski program je širi pojam od genetičkog algoritma i svaka je programska implementacija ili algoritamska imitacija evolucijskog procesa u prirodi. Nad kromosomima, za evolucijski program nije važan prikaz mogućeg rješenja niti su genetički operatori definirani. Neophodno je da program raspolaže s populacijom rješenja nad kojom se obavlja selekcija. Genetički algoritam najčešće podrazumijeva niz znakova ili bitova kao prikaz rješenja, a nad nizovima su definirani klasični genetički operatori: križanje i mutacija [35].

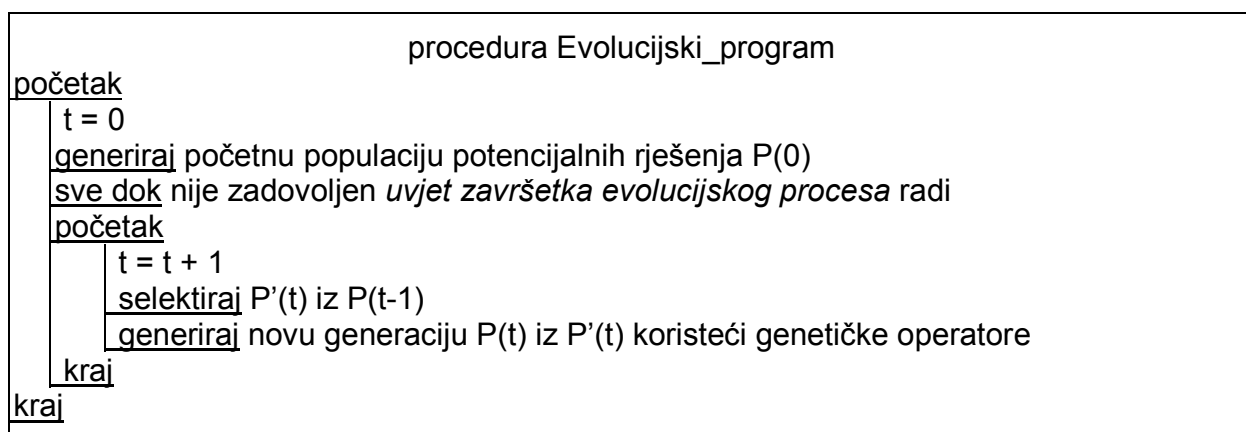


Slika 21. Evolucijski program - širi pojam od genetičkog algoritma [35]

Želi li se riješiti problem uz pomoć genetičkog algoritma, tada je problem potrebno prilagoditi genetičkom algoritmu. Moguće rješenje se najčešće prikazuje kao niz bitova u binarnom prikazu. Uobičajeni genetički operatori mogu generirati značajan postotak nemogućih rješenja (npr., za problem najkraćeg puta) koji ne donose nikakva poboljšanja, već samo usporavaju algoritam. Osim uobičajenog binarnog prikaza koriste se razni drugi prikazi (matrice, nizovi realnih brojeva, itd.) i definiraju se usko specijalizirani genetički operatori primjenjivi samo za određeni problem kako bi se izbjegla nemoguća rješenja. Promjenom prikaza i genetičkih operatora dobiva se prilagođeni genetički algoritam problemu koji se naziva evolucijski program.



Slika 22. Pristup problemu s pomoću genetičkog algoritma i evolucijskog programa



Slika 23. Struktura evolucijskog programa [35]

Prilikom inicijalizacije generira se početna populacija jedinki, te se obično početna populacija generira slučajnim odabirom rješenja iz domene. Slijedi proces koji se ponavlja sve dok ne istekne vrijeme ili je zadovoljen neki uvjet, a taj proces se sastoji od djelovanja genetičkih operatora selekcije, križanja i mutacije nad populacijom jedinki. Prilikom selekcije loše jedinke odumiru, a bolje opstaju te se u slijedećem koraku, križanju - razmnožavaju. Mutacijom se mijenjaju svojstva jedinke slučajnom promjenom gena, dok se križanjem prenose svojstva roditelja na djecu. Takovim postupkom se postiže iz generacije u generaciju sve veća i veća prosječna dobrota populacije.

Genetički algoritam simulira prirodni evolucijski proces. Za evolucijski proces kao i za genetički algoritam se može ustanoviti da [35] :

- postoji populacija jedinki;
- neke jedinke su *bolje* (bolje su prilagođene okolini);
- *bolje* jedinke imaju veću vjerojatnost preživljavanja i reprodukcije;
- svojstva jedinki zapisane su u kromosomima s pomoću genetskog koda.
- djeca naslijeđuju svojstva roditelja;
- nad jedinkom može djelovati mutacija.

Za genetički algoritam jedinke su potencijalna rješenja, a okolinu predstavlja funkcija cilja. Želi li se implementirati genetički algoritam kao metoda optimiranja, potrebno je definirati proces dekodiranja i preslikavanja podataka zapisanih u kromosomu te odrediti funkciju dobrote.

Problem: Potrebno je pomoću genetičkog algoritma pronaći optimum (minimum ili maksimum) funkcije $f(x)$ za $x \in [dg, gg]$. Funkcija je definirana za svaki x na zadanom intervalu. Dovoljno je definirati postupak za $\max\{f(x)\}$, jer ako se traži $\min\{f(x)\}$, ekvivalentan je problem riješiti $\max\{-f(x)\}$. Valja naglasiti da se nad funkcijom $f(x)$ ne postavljaju nikakva dodatna ograničenja. Primjerice nije bitno da li funkcija ima prvu, drugu ili n -tu derivaciju, niti se ne zahtjeva da je neprekidna i sl. Dakle, $f(x)$ je potpuno proizvoljna [35].

Prikaz rješenja

U jednom kromosomu zapisani su svi podaci koji obilježavaju jednu jedinku. Ukoliko se traži globalni optimum funkcije cilja jedne varijable x tako da vrijedi: $x \in [dg, gg]$ gdje su $dg, gg \in \mathcal{R}$. neka je problem jednodimenzijski, tada jedna jedinka, tj. jedan kromosom predstavlja jedno rješenje $x \in [dg, gg]$. Podaci mogu biti pohranjeni na razne načine u jedan kromosom. Zbog lakše prezentacije algoritma, podaci su spremljeni u niz bitova koji čine jedan kromosom. Ovakav prikaz naziva se binarni prikaz. Jedan kromosom može biti i realan broj, a ukoliko je problem višedimenzijski, umjesto jednog broja koji je zapisan u kromosomu, treba biti polje brojeva veličine dimenzije problema [35].

Općenito, kromosom može biti bilo kakva struktura podataka koja opisuje svojstva jedne jedinke. Za genetički algoritam je značajno da kromosom predstavlja moguće rješenje zadanog problema. Za svaku strukturu podataka valja definirati genetičke operatore. Međutim, genetički operatori trebaju biti tako definirani da oni ne stvaraju nove jedinke koje predstavljaju nemoguća rješenja, jer se time znatno umanjuje učinkovitost genetičkog algoritma.

Prikaz rješenja s pomoću prirodnog binarnog koda

Prikaz rješenja može bitno utjecati na učinkovitost genetičkog algoritma, stoga je izbor prikaza izuzetno značajan. Mnogobrojna teorija vezana uz genetičke algoritme je ograničena i vezana uz binarni prikaz [35].

Kromosom kao binarni vektor predstavlja kodiranu vrijednost $x \in [dg, gg]$. Dužina n binarnog broja utječe na preciznost i označava broj bitova, odnosno broj jedinica ili nula u jednom kromosomu. To znači da je u taj vektor moguće zapisati 2^n različitih kombinacija nula i jedinica, tj. moguće je zapisati bilo koji broj u intervalu $[0, 2^n - 1]$. Binarni vektor $v(0) = [000 \dots 0]$ predstavlja vrijednost $x = dg$, a vektor $v(2^n - 1) = [111 \dots 1]$ predstavlja vrijednost $x = gg$. Općenito, ako je binarni broj $b \in [0, 2^n - 1]$ zapisan kao binarni vektor $v(b) = [B_{n-1}, B_{n-2}, \dots, B_1, B_0]$, gdje je $B_i = 0$ ili 1 , tada vrijedi:

$$b = \sum_{i=0}^{n-1} B_i 2^i \quad (15)$$

Dekodiranje je proces pretvaranja binarnog broja u potencijalno rješenje. U ovom slučaju potencijalno rješenje je bilo koji realan broj x u intervalu $[dg, gg]$. Binarni vektor $v(b)$ predstavlja vrijednost x :

$$x = dg + \frac{b}{(2^n - 1)}(gg - dg) \quad (16)$$

Određivanje binarnog broja b , tj. *kodiranje se* za zadani realan broj x obavlja prema formuli:

$$b = \frac{x - dg}{gg - dg}(2^n - 1) \quad (17)$$

Neka je rješenje x preciznosti p i neka je $p \in \mathbb{N}$. To znači da x odstupa od točnog rješenja za maksimalno 10^{-p} . Odnosno, rješenje x je točno na p decimala. Rješenje x je preciznosti p ako je zadovoljena slijedeća nejednakost:

$$(gg - dg) * 10^p < 2^n - 1. \quad (18)$$

Za zadanu preciznost p , kromosom mora biti duljine:

$$n \geq \frac{\log[(gg - dg) * 10^p + 1]}{\log 2}. \quad (19)$$

FUNKCIJA DOBROTE [35]

Funkcija dobrote ili funkcija ocjene kvalitete jedinke se još naziva *fitness* funkcija ili funkcija cilja i u najjednostavnijoj interpretaciji ekvivalent je funkciji f koju treba optimirati :

$$dobrota(v) = f(x), \quad (20)$$

gdje binarni vektor v predstavlja realan broj $x \in [dg, gg]$. Što je *dobrota* jedinke veća, jedinka ima veću vjerojatnost preživljavanja i križanja. Funkcija dobrote je ključ za proces selekcije. Pokazalo se u praksi da se ova jednostavna transformacija može primijeniti samo uz neka ograničenja nad $f(x)$. Problem pronaći minimum funkcije $f(x)$ je ekvivalentan polaznom problemu (pronaći optimum zadane funkcije cilja). Međutim, ideja je u redu i treba je samo na drugi način iskoristiti. Nije potrebno znati $\min\{f(x)\}$ za svaki $x \in [dg, gg]$, dovoljno je znati najmanju vrijednost funkcije cilja za sve kromosome:

$$dobrota(v(x)) = f(x) - \min\{f(x_i)\}, \text{ za } i = 1, 2, \dots, VEL_POP. \quad (21)$$

Vrijednost $\min\{f(x_i)\}$ nije ista za svaku generaciju i ona se mora u svakom koraku iznova računati. To ne znači usporavanje algoritma novom potragom za najlošijom jedinkom, jer se ionako provodi pretraga za najboljom jedinkom koja se čuva i prenosi u slijedeću generaciju radi poboljšavanja učinkovitosti algoritma. Dakle, identifikacija najbolje i najlošije jedinke se obavlja u jednom koraku pretrage.

Opisani postupak naziva se translacija ili pomak funkcije cilja. Na taj način funkcija dobrote niti može poprimiti negativne vrijednosti, niti se translacijom dobivaju približno iste vrijednosti.

3.2. Softverska rješenja poslovnog izvješćivanja na inozemnom i domaćem tržištu

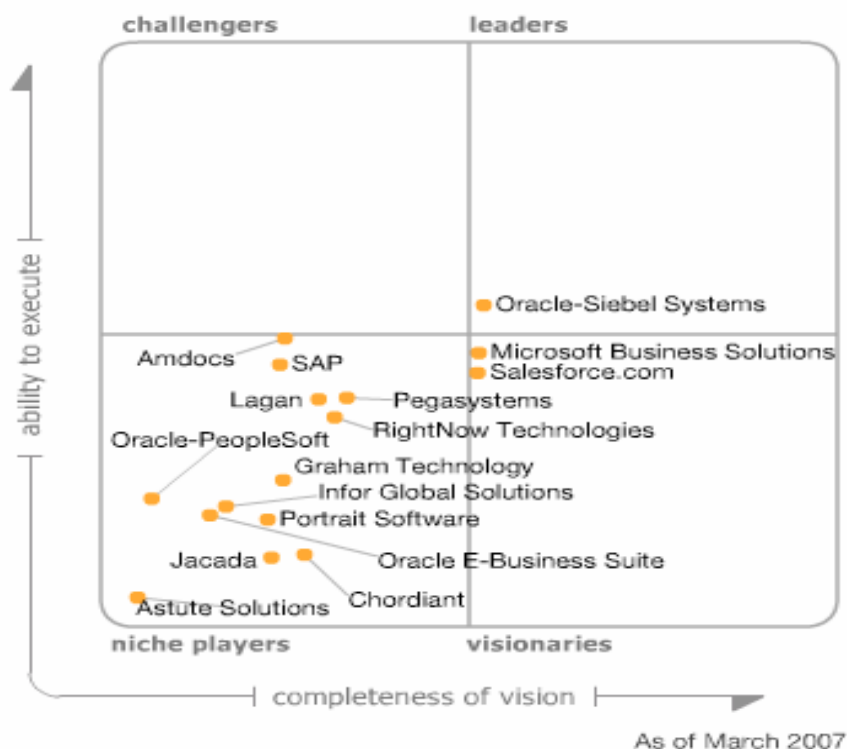
Prema istraživanjima koje je provela tvrtka Gartner Inc. napravljen je prikaz vodećih svjetskih ponuđača alata poslovnog izvješćivanja [39]. Prema istraživanjima iz kolovoza 2003. i ožujka 2007. godine dat je i grafički prikaz tvrtki i njihovih pozicija u tzv. *magičnom kvadrantu* kao i predviđanja stanja na tržištu od strane Gartnerovih analitičara.

Promjena pozicija u svijetu istih tvrtki može se vidjeti usporedbom kvadranta iz kolovoza 2003. godine (slika 24.) i ožujka 2007. godine (slika 25.), ali i pojava novih ponuđača u kvadrantu, i njihova pozicija u odnosu na druge tvrtke koje su već "zauzele" mjesto u kvadrantu. Prema istraživanjima Gartner Inc. iz ožujka 2007. godine, tržište tehnologije poslovnog izvješćivanja nastavilo je s razvojem u odnosu na početak 2003. godine. Prema tom istraživanju, zapažen je lagani rast u istovremenu vidljivu konsolidaciju tvrtki ponuđača tehnologije poslovnog izvješćivanja. Razvijanje vlastite vizije i infrastrukture za svoju arhitekturu poslovnog izvješćivanja bila je preporuka Gartner-ovih analitičara kompanijama. Na taj način će smanjiti utjecaj „nereda“ koji može nastati daljnjim oporavkom tržišta, odnosno tvrtki ponuđača alata poslovnog izvješćivanja [39]. Analitičari su predvidjeli stabilizaciju tržišta poslovnog izvješćivanja i preuzimanje malih i manje stabilnih tvrtki od strane velikih tržišnih tvrtki. U kvadrant su uključene i prikazane samo one tvrtke za koje se u kolovozu 2003. godine smatralo da su najznačajniji i najaktivniji predstavnici ponuđača tehnologije poslovnog izvješćivanja na svjetskom tržištu. Gartnerovi analitičari su dali prognozu da će planovi integracije predstavljati ključ daljnjeg uspjeha opstanka i rasta pojedinih tvrtki ponuđača alata i platformi poslovnog izvješćivanja. Na slici 24. u magičnom kvadrantu prikazani su ponuđači tehnologije poslovnog izvješćivanja koja uključuje samo alate poslovnog izvješćivanja (EBIS - The Enterprise BI Suites), bez ponude platformi poslovnog izvješćivanja (BI Platforms) jer je to tržište aktivnije u odnosu na dva tržišta zajedno (EBIS i BI platforms) [39].



Slika 24. Ponuđači alata poslovnog izvješćivanja 2002./2003. godine [39]

Tvrtnka Gartner, analizom prikazanih ocjena dolazi do zaključka da danas postoje četiri grupe ponuđača rješenja za poslovno izvješćivanje -1. mali i ranjivi, 2. fokusirani i brzi, 3. jaki i fokusirani i 4. veliki i oportunistički. U grupi malih i ranjivih proizvođača spadaju mali proizvođači koji se tek bore za svoje mjesto na tržištu, te su nemoćni i ranjivi u napadu konkurencije. Grupa fokusiranih i brzih se sastoji od srednjevelikih i jako fokusiranih ponuđača poput Cognos-a, Hyperion Solutionsa, Crystal Decisions-a i dr. Skupinu jakih i fokusiranih čine ponuđači koji dobro poznaju tržište i imaju snage odgovoriti na sve hirove i pritiske, a čine ju kompanije SAS Institute i NCR Teradata Division. U grupu velikih i oportunističkih spadaju ponuđači čija je tržišna snaga imponzantna i velika, poput Oracle-a, Microsofta, IBM-a i Computer Associates [3].



Slika 25. Ponuđači alata poslovnog izvješćivanja u 2007. godini [41]

3.2.1. SAP - Vodeći svjetski proizvođač poslovnih softvera

SAP je vodeći svjetski proizvođač poslovnih softverskih rješenja. Danas, više od 29.800 korisnika u više od 120 zemalja koristi više od 100.600 instalacija SAP-ova softvera – od rješenja koja odgovaraju na potrebe malih i srednjih tvrtki do poslovnih rješenja za velike i globalne korporacije **mySAP™ Business Suite** rješenja, pokretana **SAP NetWeaver™** platformom koja omogućuje inovaciju i poslovne promjene, pomažu korporacijama diljem svijeta u poboljšavanju odnosa s korisnicima, povećanju suradnje s poslovnim partnerima i u kreiranju učinkovitosti diljem svojih nabavnih lanaca i poslovnih operacija [37].

Tvrtka SAP je rješenjem mySAP ERP pokazala da joj pripada vodeća uloga na tržištu rješenja za planiranje kapaciteta tvrtki. Rješenje mySAP ERP koje je postavilo standarde u svojoj kategoriji, uvodi nove razine integracije poslovnih procesa u heterogena systemska okruženja. Opsežna funkcionalnost mySAP ERP, njegove fleksibilne

mogućnosti dogradnje i globalna orijentacija opravdavaju njegovu ulogu najcjelovitijega ERP rješenja na tržištu.

Bitne prednosti rješenja mySAP ERP uključuju [37]:

- *novu učinkovitost u sklopu integriranih end-to-end poslovnih rješenja* – omogućuje cjelovite, potpuno integrirane poslovne procese za financije, upravljanje ljudskim kapitalom, operacijama i korporativnim uslugama uz optimalizaciju povezanih informatičkih ulaganja i povećanje poslovnih koristi.
- *poslovnu agilnost* - rad u realnom vremenu radi bržega donošenja odluka i bržega reagiranja na tržišne prilike.
- *niži ukupni trošak vlasništva* - iskorištavanje aktualnih informatičkih ulaganja, smanjenje integracijske složenosti i minimaliziranje potrebe za programiranjem po mjeri, zahvaljujući snazi SAP NetWeavera. Rješenje se lako instalira i nudi fleksibilne mogućnosti koje pojednostavnjuju implementaciju i dogradnju te dodatno smanjuju trošak vlasništva. Osiguravajući interoperabilnost svojih rješenja s tehnologijama trećih strana, SAP pomaže tvrtkama da zaštite svoja postojeća informatička ulaganja.
- *postupnost instalacije* - poboljšanje financijskoga tijeka i smanjenje skupih pozajmica korištenjem ključnih funkcija onda kad ih trebate, u skladu sa zahtjevima vašega poslovanja. Postupna dogradnja prati promjenu vaših potreba, uz dodavanje novih rješenja i poslovnih funkcija.
- *brži povrat ulaganja* - eliminacija troškova nabave i integracije softverskih dodataka

trećih ponuđača te skraćivanje vremena do ostvarivanja stvarnih prednosti primjenom definiranih projekata u kratkim, upravljivim vremenskim okvirima.

MySAP ERP lako se proširuje naprednim mogućnostima dodavanjem rješenja kao što je mySAP CRM (Customer Relationship Management), mySAP SRM (Supplier Relationship Management), mySAP PLM (Product Lifecycle Management), te mySAP SCM (Supply Chain Management). Alternativno, s lakoćom je moguće prijeći na mySAP Business Suite koji korisnicima pruža sveobuhvatnu funkcionalnost te podiže sustav na višu razinu poslovanja. MySAP Business Suite sveobuhvatna je cjelina prilagodljivih poslovnih rješenja koja optimiziraju kritične poslovne procese. Optimizacija je omogućena putem neograničene integracije s gotovo bilo kojim sustavom. Tehnološka osnova SAP-ovih rješenja je integracijska i aplikacijska platforma SAP NetWeaver. Ona polaže temelje međufunkcijskim poslovnim procesima te snižava ukupni trošak vlasništva smanjenjem potreba za pojedinačnom integracijom i ponudom cjelovitog upravljanja životnim vijekom rješenja [37].

MySAP ERP osigurava i najbolje tipske i najbolje specifične funkcije za pojedinu djelatnost i najbolju praksu na temelju dugogodišnjih SAP-ovih iskustava. To znači da je moguće znatno poboljšati upravljanje tvrtkom pretvaranjem podataka u željene informacije po kojima se pravodobno odlučuje. Tvrtke moraju vrednovati svaku informatičku investiciju prema mogućnosti ostvarivanja brzih povrata i smanjenja troškova poslovanja. Istodobno, moraju procjenjivati zahtjeve koji se postavljaju pred njihove informatičke sustave, od potpore novim modelima poslovanja do proširenja poslovnih procesa na internet i mobilno poslovanje.

U to spadaju ovi zahtjevi [37]:

- fleksibilnost informatičkih struktura koje omogućavaju tvrtkama brzu prilagodbu poslovnih procesa;

- jednostavnost integracije novih funkcija i procesa, bez negativnog utjecaja na temeljnu tehnologiju ili ranija ulaganja;
- manja složenost informatičkih struktura i racionalizirani poslovni procesi;
- personalizirani podaci, procesi i rješenja za svakog zaposlenika u bilo koje vrijeme, bilo gdje,
- integracija partnera, ponuđača i korisnika u sklopu poslovnoga intraneta ili ektraneta.

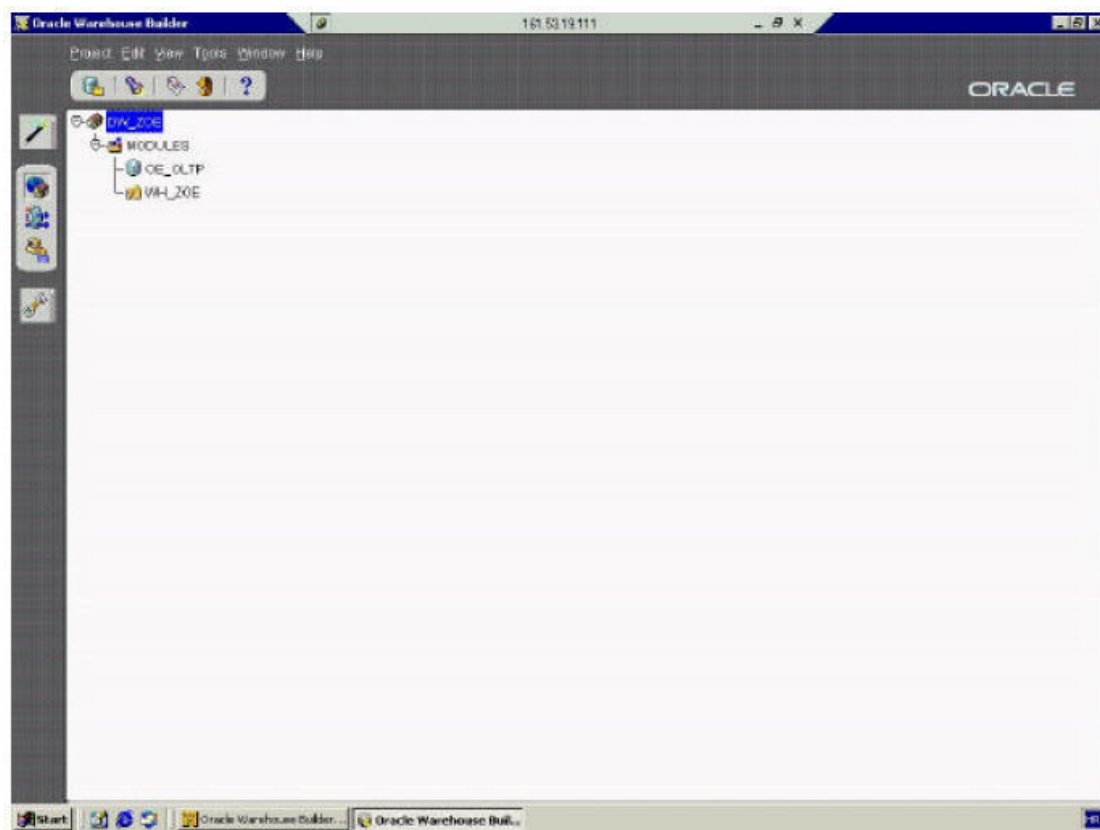
Kao cjelovito rješenje, mySAP ERP izlazi na kraj sa svim tim izazovima. Njegove napredne integracijske mogućnosti posredstvom SAP NetWeavera i fleksibilne mogućnosti dogradnje, zajedno sa širokim rasponom funkcija, globalnim obuhvatom i orijentacijom na konkretnu djelatnost potkrepljuju tvrdnju kako je mySAP ERP najcjelovitije rješenje za planiranje kapaciteta tvrtke na tržištu.



Slika 26. Tvrtka SAP [37]

3.2.2. Oracle warehouse builder (programski paket)

Oracle Warehouse Builder je softverski paket tvrtke Oracle Corporation, a trenutno najnovija inačica je verzija 3i. OWB je programski paket koji služi za definiranje logičkog modela, implementaciju skladišta podataka kao i za nadgledanje i kontrolu rada skladišta podataka. Radi se o integriranom skupu programskih rješenja koji omogućava daleko lakše dizajniranje i izgradnju skladišta podataka, ali i kontrolu rada i nadgledanje skladišta podataka [40].



Slika 27. Glavni prozor Oracle Warehouse Buildera [40]

Programski paket OWB se sastoji od OWB repozitorija, OWB klijenta i OWB *Runtime*-a. Osim tih proizvoda za potrebe skladišta podataka neophodan je i *Oracle Enterprise*

Manager, te neki alati za generiranje izvješća. Da bi instalacija bila uspješna slijedećim redoslijedom potrebno je instalirati ove produkte: prvo treba instalirati bazu (ako već ne postoji), potom se instalira repozitorij u tu odgovarajuću bazu podataka, te se klijent instalira na korisnikovo osobno računalo (može biti i više korisnika koji rade u OWB-u). OWB *Runtime* se instalira posljednji i služi za poslove nadgledanja. Korisnik pristupa repozitoriju preko OWB klijenta. Prilikom pokretanja OWB klijenta prvi put potrebno je dati informacije o imenu računala na kojem je repozitorij, broju port-a, te Oracle SID. Također je potrebno dati svoje korisničko ime i lozinku i izabrati projekt na kojem će se raditi. Sva akcija korisnika se odvija u grafičkom sučelju koje je standardno za današnje aplikacije. Korisnikove akcije i promjene koje on unese na ekranu, ne zapisuju se automatski u repozitoriju već je te promjene i akcije potrebno potvrditi pritiskom na tipku *Commit*. Tek tada napravljene promjene postaju važeće i unose se u repozitorij [40].

Osnovni elementi Oracle Warehouse Buildera

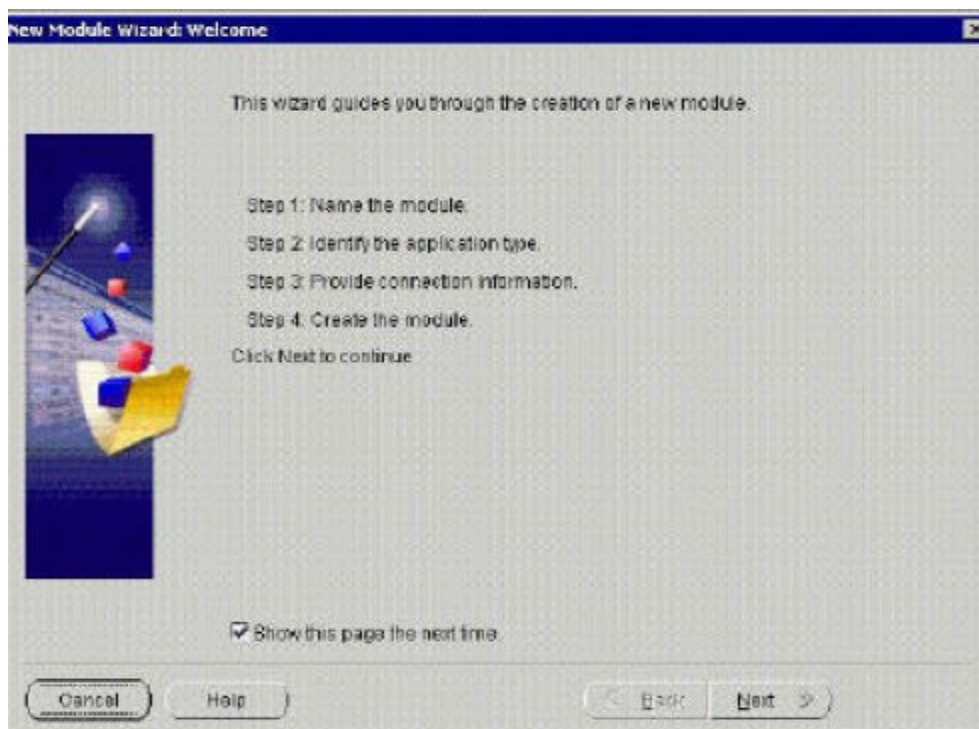
Projekt je osnovni element Oracle Warehouse Buildera, a definira se kao struktura repozitorija u kojoj se čuvaju formalni opisi koji definiraju skladište podataka i u kojoj OWB sprema generirane skripte korištene pri implementaciji i učitavanju podataka. Projekt je, dakle, osnovna jedinica u Oracle Warehouse Builderu, a svaki se sastoji od jednog ili više izvorišnih modula (engl. Source module) i jednog ili više odredišnih ili skladišnih modula (engl. target module, warehouse module). Mjesto unutar OWB projekta koje organizira i sprema definicije potrebne za logičku shemu skladišta zove se *odredišni ili skladišni modul*. Odredišni modul sadrži definicije za dimenzije, materijalizirane pogledе, obične pogledе, tablice, te za mapiranja i transformacije. Izvorišni modul je mjesto unutar OWB projekta koje organizira i sprema definicije relacijskih baza ili običnih datoteka (engl. flat files) koje služe kao izvori podataka za skladište podataka. Definicije relacijskih baza se mogu uvesti (engl. import) iz bilo koje baze podataka (ne samo Oracle-ove). OWB koristi takozvane softverske integratore (engl. software integrators) za čitanje definicija i izvlačenje podataka iz izvora. Ovisno o izvoru koristit će se odgovarajući integrator [40].

Način rada Oracle Warehouse Buildera

Izgradnja skladišta podataka je podjeljena u tri dijela [40]:

1. logička definicija koja obuhvaća i logičku definiciju mapiranja podataka iz izvora.
2. konfiguracija svih objekata definiranih na logičkoj razini.
3. generiranje i pokretanje skripti za stvaranje logičkog modela (dimenzija, tablica činjenica,...), te generiranje i pokretanje skripti za izvlačenje, transformaciju i učitavanje podataka iz izvora u skladište podataka.

Stvaranjem izvorišnih i odredišnih modula započinje logička definicija. Nakon definiranja izvorišnog modula, neophodno je uvesti definicije relacijske baze koja služi kao izvor podataka. Ukoliko postoji više izvora, potrebno je stvoriti više izvorišnih modula (za svaki izvor podataka, potreban je jedan izvorišni modul). Poslije definiranja odredišnog modula, potrebno je definirati, unutar samog modula, dimenzije, tablice činjenica, materijalizirane poglede,... prema logičkom modelu podataka. Također unutar odredišnog modula definiraju se transformacije i mapiranja. U OWB-u već postoje standardne funkcije i procedure koje se koriste u mapiranjima. Logička definicija je zapisana u repozitorij, ali još nije stvoren nijedan objekt, nijedna tablica, niti je stvorena ijedna skripta. Da bi se stvorila fizička instanca skladišta, prvo je potrebno konfigurirati fizička svojstva svakog modula, objekta, tablice, svakog mapiranja i operatora unutar mapiranja. Na taj način je definirano kako će se logički model fizički kreirati itd. Poslije konfiguracije fizičkih parametara potrebno je generirati skripte za izvlačenje, transformaciju i učitavanje podataka iz izvora u skladište, te skripte za kreiranje raznih tablica. Skripte zatim treba spremiti u bazu podataka i nakon toga pokrenuti. Na taj način se stvara fizička instanca skladišta podataka i učitavaju se podaci u njega. Skladište podataka je „izgrađeno“ učitavanjem podataka [40].

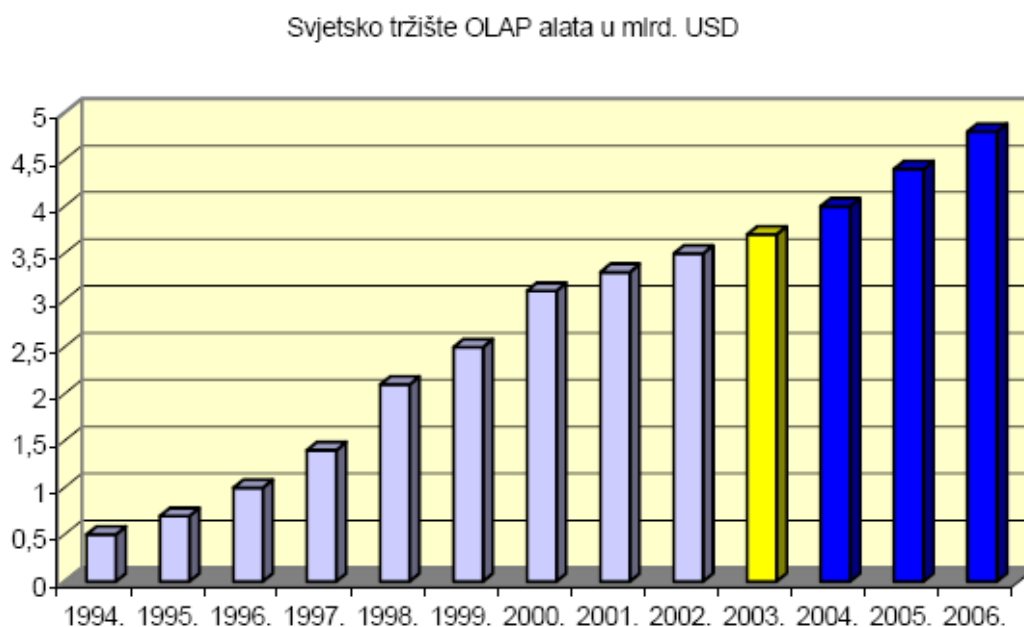


Slika 28. Uvodni prozor New Module Wizarđa u ORACLE-u [40]

3.2.3. Analiza tržišta OLAP alata

Prema provedenoj analizi udjela na tržištu OLAP alata ažuriranoj 8. travnja 2004. godine, i izvješću autora Nigel Pendsea, započela je u 2003. godini dugoročna prognoza o konsolidaciji industrije uporabom poslovnog izvješćivanja s tvrtkama Business Objects, Cognos i Hyperion koje značajnim akvizicijama proširuju svoju paletu proizvoda [39]. Prema citiranom izvoru, navedene tvrtke su značajne za tržište poslovnog izvješćivanja u cijelosti, ali imaju relativno malen utjecaj na udio u OLAP tržištu. Poteškoće koje se javljaju u određivanju precizne veličine cijelog OLAP tržišta i individualnih udjela u njemu, nastaju zbog toga što veliki prodavači kao Microsoft, Oracle, SAP, Business Objects ne mogu uvijek izmjeriti OLAP udio sam za sebe zbog toga jer se njihovi alati ne isporučuju i

prodaju samostalno već često kao dio većih paketa proizvoda. Poteškoće u dobivanju preciznih podataka za određivanje OLAP tržišta javljaju se i zbog mnogih drugih razloga, npr. mnoge velike "non-BI" kompanije preuzimaju prave OLAP vendore, neke od njih ne razdvajaju OLAP proizvode od ostalih (Microsoft), neke manje kompanije npr. na tržištu SAD nemaju obvezu publicirati financijske rezultate, neke kompanije su koncentrirane na direktnu prodaju, dok druge to čine uz pomoć preprodavača itd. Na grafikonu br. 1. dat je prikaz veličine svjetskog tržišta OLAP alatima iz uvodno citiranog izvješća, u razdoblju od 1994. do 2006. godine izražen u USD.



Grafikon 1. Prikaz veličine svjetskog tržišta OLAP alatima [39]

Iz grafikona 1. je vidljivo da je vrijednost svjetskog OLAP tržišta zajedno sa uslugama implementacija, sredinom devedesetih bila oko 1 milijardu USD, i imala trend rasta preko 2 mlrd. USD u 1998. godini, uz daljnji trend rasta [39]. Stopa rasta u 2003. godini je mala u odnosu na 2002. godinu. Međutim od 2004. do 2006. je uočljiv značajan trend rasta prodaje. Kao i u većini ostalih SW sektora 2002. godina bila je teška godina, kao i početak 2003., uz poboljšanje situacije do kraja 2003. godine. S daljnjim razvojem OLAP aplikacija očekuje se i znatan porast prodaje OLAP aplikacija. Na kraju dat je i tablični

prikaz vodećih tvrtki na svjetskom tržištu OLAP alata odnosno njihov udio u tom tržištu u razdoblju od 2001. – 2003. godine [39].

Tablica 3.6. Prikaz vodećih tvrtki na svjetskom OLAP tržištu i njihov udjel u tom tržištu u 2001., 2002. i 2003. godinu [39]

Godina	2003.		2002.		2001.	
Tvrtka	Red. br.	Udio u %	Red. br.	Udio u %	Red. br.	Udio u %
Microsoft	1	26,1	1	24,4	2	21,1
Hyperion Solutions (uklj. Brio S.)	2	21,9	2	23,3	1	27,4
Cognos (uklj. Adaytum)	3	14,2	3	14,7	3	13,7
Busniess Objects (uklj. Crystal)	4	7,7	4	7,4	4	7,6
MicroStrategy	5	6,2	5	5,4	6	6,8
SAP	6	5,8	6	5,2	7	5,4
Oracle	7	4,0	7	4,7	5	7,0
Cartesis	8	3,1	9	2,6	9	2,4
Applix	9	3,0	8	2,6	8	2,5
MIS AG	10	3,0	11	2,1	11	2,1
Geac	11	2,0	10	2,2	10	2,3
SAS Institute	12	0,9	12	1,1	12	1,2

3.2.4. Inozemne tvrtke i njihovi partneri na hrvatskom tržištu

Prema iznesenim informacijama konzultanta, na hrvatskom tržištu javlja se više kategorija ponuđača SW koji nude alate i / ili usluge poslovnog izvješćivanja [39]:

- **Prva kategorija** koja se pojavljuje su sistem integratori koji nude sve varijante usluge sistema integracije, pa između ostalog i implementaciju sustava poslovnog izvješćivanja.

Dakle ova kategorija nudi HW, SW + usluge. To su velike svjetske tvrtke proizvođači informatičke opreme te konzultantske kuće kao: IBM, Siemens, Deloitte, KPMG.

• **Drugu kategoriju** čine dobavljači softvera i njihovi distributeri u RH, koji su koncentrirani uglavnom na distribuciju SW, a pružaju i usluge implementacije njihovog softvera. Takav odnos na hrvatskom tržištu imaju sljedeće inozemne tvrtke:

- **SAS**, sa registriranim uredom i timom stručnjaka u RH,
- **Cognos** koji ima partnerski odnos s hrvatskim tvrtkama Megatrend, Poslovna inteligencija i Optima,
- **SAP**, također ima registriran ured i stručnjake u RH ali i dodatne partnere,
- **MicroStrategy** - partnerski odnos imaju s hrv. tvrtkom MCS Group,
- **Oracle**, ima vlastiti ured plus partnere: In2 i Infoopus,
- **Microsoft**, za alate BI surađuje sa 3 front-end partnera u RH: SoftPro (proizvodi BI front-end platforme za Microsoft OLAP), ProClarity i Panorama,
- **Panorama**, partner je Poslovna inteligencija, a servis radi mreža parntera
- **ProClarity** sa partnerom Incendo,
- **Hyperion Solutions** i **Informatica** kao poznate svjetske BI tvrtke nemaju zasad predstavnika na tržištu RH.

Hrvatske tvrtke korisnici sustava poslovnog izvješćivanja

Poznatije hrvatske tvrtke korisnici koje su implementirale SW i u praksi primjenjuju koncept poslovnog izvješćivanja između ostalih su [39]:

1. SW BI tvrtke Cognos ugradile su:

- Lura,
- Agrokor

- Kraš
- Zagrebačka pivovara
- Nexe Grupa (Našice Cement).

2. SW BI tvrtke Business Objects ugradile su:

- VIPnet
- Croatia Airlines
- HT Mobile.

3. SW BI tvrtke SAS koriste:

- Zagrebačka banka
- Privredna banka
- INA.

4. SW BI tvrtke SAP koriste:

- Pliva
- Podravka
- Ministarstvo financija RH.

5. SW BI tvrtke Oracle koriste:

- VIPnet
- Croatia osiguranje.

6. SW BI tvrtke SoftPro koriste:

- Alastor
- Medika
- Magma.

7. SW BI tvrtke Panorama koriste:

- MSan
- Farmacija
- Institut za turizam.

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Da bi došli do određenih spoznaja o načinu održavanja vagona potrebno je kroz određeno razdoblje prikupiti podatke o vagonima u vožnji i zastojima u radionici. U tim evidencijama (snimkama radova i vagonskim kartonima) se vode i podaci o zastojima zbog kvarova i popravcima na samim vagonima. Evidencije se vode na svim putničkim vagonima, pri čemu se u vagonске kartone upisuju sve promjene stanja praćenih parametara bez obzira da li je vagon u radionicu došao zbog izvanrednog kvara ili redovitih pregleda i popravaka.

Pored toga u svim radionicama održavanja, tehničkoj i operativnoj pripremi također se vode odgovarajuće evidencije o radovima na održavanju i popravku svakog vagona pojedinačno, odnosno, njihovim kvarovima, zamijenjenim dijelovima izvršenim kontrolama i sl. Sva ta dokumentacija dobar je izvor za prikupljanje podataka koji su neophodni za realizaciju istraživanja. Iz tih evidencija su prikupljeni podaci o zastojima po serijama klimatiziranih putničkih vagona. Podaci su prikupljeni u razdoblju od 1.siječnja 2005. godine do 1. siječnja 2007. godine i evidentirani na način kako je to prikazano u tablici 4.1.

Kriteriji podjele klimatiziranih putničkih vagona

Kriterij podjele klimatiziranih putničkih vagona Hrvatskih željeznica prikazan je tablicom 4.1, Vagoni se dijele s obzirom na:

- brzinu vagona (do 160 km / h i 200 km / h),
- tip i proizvođač uređaja opskrbe električnom energijom (statički pretvarač)
- red i klasu razreda (oznaka A-1. razred, B-2. razred, AB-1. + 2. razred)

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Tablica 4.1. Podjela modernih putničkih vagona Hrvatskih željeznica s obzirom na seriju, tip statičkog pretvarača i brzinu vagona

Vagoni serije	Broj vagona	Tip statičkog pretvarača						Brzina vagona (km/h)
		Siemens EVA / B	Končar VIS 50-1	Končar VIS 50	Pintch 15	Pintch 75	MGVG 3000	
Aeelt (10-70)	12		X					160
Aeelmt (10-90)	6	X						200
Bee (20-70)	26		X		X			160
Beelmt (20-90)	5	X						200
Beelt (20-70)	4			X				160
Beel (20-70)	5						X	160
Bee (29-70)	9				X			160
ABeel (39-70)	4						X	160
WRee (88-70)	2					X		160
WReelmt (87-90)	2	X						200

Da bi se pojasnilo prema slovničkim oznakama serije o kojim je vagonima riječ:

A- je vagon 1. klase vožnje

B- je vagon 2. klase vožnje,

AB- je vagon 1. i 2. klase vožnje,

WR- je vagon restoran,

ee- vagon se može opskrbiti svim naponima prema RIC-u (vagoni ovih oznaka imaju ugrađeni statički pretvarač električne energije i klimatiziran je)

Za provedbu analize o kvarovima korišteni su arhivski kartoni prema slici 29., gdje se za svaki vagon pojedinačno vode i bilježe evidencije o razlogu isključenja vagona iz prometovanja tj. vrsti kvaru, datum isključenja itd. U svrhu istraživanja, analiza je obuhvatila serije klimatiziranih putničkih vagona serija, na uzorku :

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

- 10-90 (Aeelmt), 20-90 (Beelmt) i 87-90 (WReelmt) – ukupno 11 vagona
- 10-70 (Aeelt), 20-70 (Bee), WRee – ukupno 40 vagona
- 20-70 (Beelt) – ukupno 4 vagona
- 39-70 (ABeel), Beel (20-70)- ukupno 9 vagona
- 29-70 (Bee) – ukupno 9 vagona

Vagoni su podijeljeni prema tipovima proizvođača uređaja za opskrbu električnom energijom – statičkim pretvaračima, a ukupno ih je u razmatranju obuhvaćeno 73.

SERIJA I BROJ VAGONA

50-48 Bf

21-00 006-8

REVIZIJA 80 1 95 / 30.2.96
26.9.96 / 26.3.97
14.8.96 / 14.2.97

TIP OSOVINSKOG POSTOLJA

TIP KOČNICE

TIP (SNAGA GENERATORA RASVJETE U KW)

TIP REGULATORA NAPONA

TIP ELEKTRIČNOG GRIJANJA

Datum izljudivanja	Datum ulaza	Datum istaza	Vrsta popravka	Uzrok izbacivanja iz prometa	Primjedba o stanju vagona i mjestu upućivanja
4.2.95	6.2.95	6.2.95	TPSO	Zarubica dno	depravacija prozora
2.9.95	2.9.95	1.95	TPSO	PRAVLICA OBAJE STRANE KAPAKA	-II-
4.5.95	16.2.95	4.5.95	TPSO	POKREZ BOJILJANYA	-I-
1.3.95	13.95	3.35	TPSO	IZNOSNA BR. LADIA 258	-I-
3.9.95	3.9.95	'	TPSO	E.G.	Garnit u TSV
12.6.95	12.6.95	6.95	Pso	P-2	-I-
15.8.95	15.8.95	15.8.95	Pso	POP. SOCH. VEOSTA P-2	-II- GARANT. TSV
9.5.95	9.5.95	1.9.95	Pso	P-2	-II-
15.8.95	15.8.95	15.8.95	TPSO	BOCNI DOCKERINIG	MEDPOSTATE -II-
10.9.95	10.9.95	10.9.95	Pso	P-2 E.C.	-II-
5.9.95	5.9.95	11.4.95	TPSO	BOCNIK V.I.S.T.A.	-II-
26.11.95	26.11.95	11.45	TPSO	PRIVODNI BOČNOSTI STROJ	-II-
3.12.95	4.12.95	4.12.95	TPSO	E.G.	-II-

[illegible]

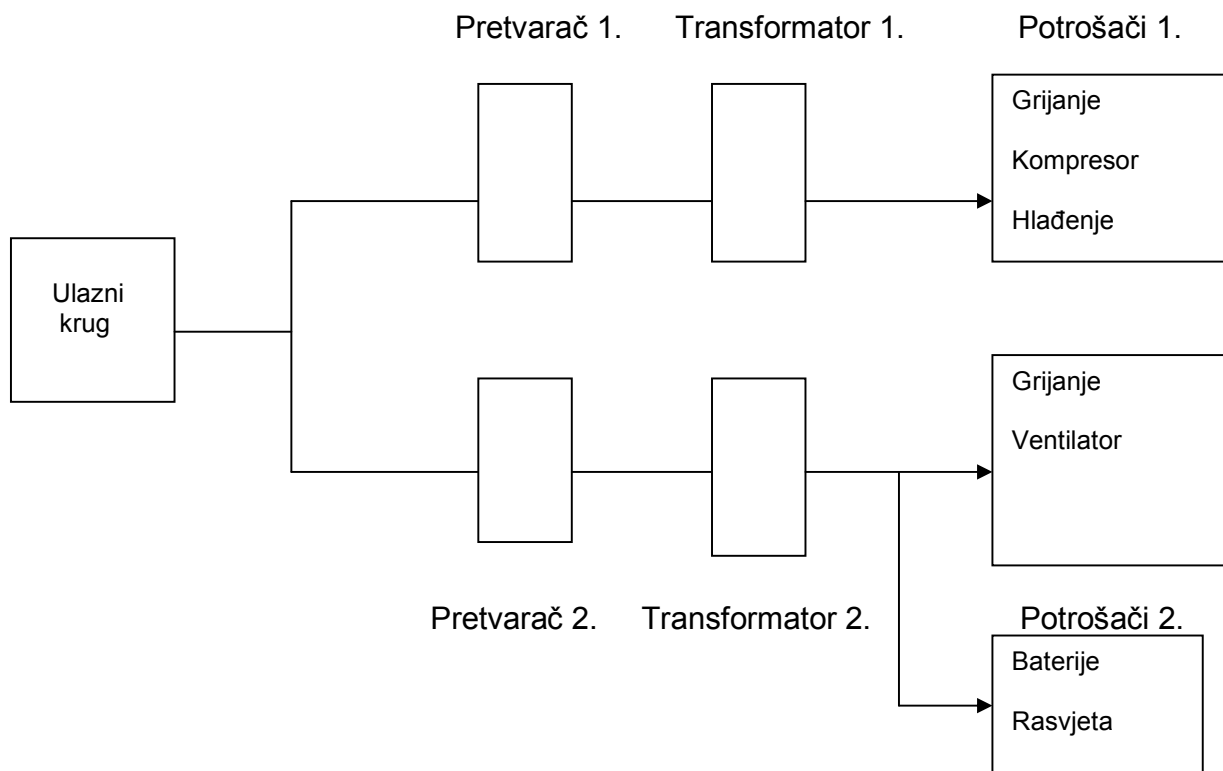
Slika 29. Vagonski kartoni sa evidencijom datuma kvara, te vrstom kvara

4.1. Opis ključnih agregata klimatiziranih putničkih vagona Hrvatskih željeznica

a.) Statički pretvarač

Prema slici 30. statički pretvarač pretvara u napone prilagođene potrošačima u vagonu četiri napona mreže (1000V 16 2/3 Hz, 1500 V 50 Hz, te 1500 V i 3000 V istosmjerne struje). Za napajanje statičkog pretvarača električna energija se isporučuje iz generatora lokomotive ili izravno iz električne vučne mreže. Statički pretvarač isporučuje napone za tri grupe potrošača električne energije u vagonu, prema [9]:

1. potrošače visokog napona (380 / 220 V) i promjenljive frekvencije,
2. potrošače visokog napona (380 / 220 V) stabilne frekvencije i
3. potrošače niskog istosmjernog napona akumulatorskih baterija 24 V



Slika 30. Blok shema uređaja za opskrbu vagona električnom energijom (statički pretvarač) [9]

Ulazni napon u pretvarače 1 i 2 koji je i izlazni napon međukruga od ukupno 570 V, podijeljen je na dva jednaka dijela U_{zk} (+285 V) i U_{zk} (- 285 V) tzv. virtualnom nulom. Naponi $+ U_{zk}$ i $- U_{zk}$ se dovode tranzistorima u pretvaraču koji imaju funkciju prekidača. Tranzistori se naizmjenično otvaraju i zatvaraju prema unaprijed određenom programu upisanom u upravljački uređaj pretvarača. Širinsko – impulsno modulirani napon se generira naizmjeničnim otvaranjem i zatvaranjem tranzistora čija je visina jednaka naponu međukruga (570 V je zbroj napona $+ U_{zk}$ i $- U_{zk}$). Neposredno dovođenje širinsko – impulsno moduliranog napona na primar glavnog transformatora izazvalo bi velike gubitke snage u transformatoru, pa se zato širinsko – impulsno modulirani napon privodi sinus filteru iz kojeg se dobije izlazni napon približno sinusoidalnog valnog oblika. Međusobni odnos vremena otvorenosti i zatvorenosti energetskih tranzistora je nepromjenljiv ali je promjenljiv takt otvaranja i zatvaranja tranzistora. Ako se takt ne mijenja, napon na izlazu sinus filtera bit će nepromjenljive visine i konstantne frekvencije. Ako se takt uspori, izlazni napon iza sinus-filtera bit će manji i niže frekvencije, a ako se takt ubrza, rast će i izlazni napon i frekvencija. Maksimalna frekvencija izlaznog napona je 50 Hz-a, a toj frekvenciji odgovara međufazni napon od 400 V [9].

b.) Vrata s elektropneumatskim upravljanjem

Na klimatizirane visokokomforne vagone ugrađena su tri tipa elektropneumatski upravljanih vrata i to bočna, čelna i prolazna vrata [9]:

1.) Bočna vrata su namijenjena za primjenu na putničkim vagonima velikih brzina (iznad 160 km / h). Mikroprocesorski sustav za protukliznu zaštitu povezan je s elektroničkim uređajem za upravljanje bočnim vratima koji mu šalje signale o brzini vožnje. Upravljački elektronički uređaj odlikuje se paralelnim djelovanjem na električne i pneumatske elemente, instalacije, čime se postiže dvostruko osiguranje od nekontroliranog otvaranja tijekom vožnje. Bočna vrata se otvaraju i zatvaraju stlačenim zrakom pomoću pneumatskog cilindra preko elektromagnetskih ventila pritiskom na tipkala za otvaranje i zatvaranje.

2.) Čelna vrata imaju elektropneumatski sustav upravljanja koji je namijenjen za primjenu na vlakovima velikih brzina. Ova vrata se otvaraju pritiskom na tipkala, a zatvaraju se automatski nakon određenog vremena. Pritiskom na tipkalo (vanjsko ili unutarnje) aktivira se magnetni ventil za otvaranje i zatvaranje, ubacuje se zrak u stražnju stranu pneumatskog cilindra i vrata se otvaraju, a istodobno se zrak iz prednje strane pneumatskog cilindra ispušta u atmosferu.

3.) Prolazna vrata odjeljka se otvaraju pomoću triju optičkih osjetila. Magnetni ventil za otvaranje vrata se aktivira pobudom bilo kojega od triju osjetila preko upravljačkih releja i elektroničnog vremenskog sklopa. Zrak ulazi u pneumatski cilindar i vrata se otvaraju. Vrata ostaju u otvorenom položaju do isteka namještenog vremena oko deset sekundi a zatim elektronički vremenski preklapa magnetni ventil u prvobitni položaj. Vrata se zatvaraju pomoću poluge, dok se preko magnetnog ventila zrak iz cilindra ispušta u atmosferu.

c.) Klima uređaj

Visokokomforni putnički vagoni opremljeni su klima uređajima koji imaju ulogu održavanja ugodne temperature u vagonu u svim godišnjim dobima [9]. Zrak je medij za transport energije za grijanje ili hlađenje u prostoru za putnike a prethodno je toplinski obrađen u klima uređaju. Sustav za regulaciju daje nalog za određeno pogonsko stanje ovisno o temperaturnim uvjetima. U vagonima su moguća tri pogonska stanja - predgrijavanje, grijanje i hlađenje. Klima uređaj je podijeljen u nekoliko komponenti, koje se mogu do određene granice smatrati samostalnim, ali im je funkcioniranje strogo određeno međusobnim vezama. Električne i mehaničke međusobne veze su u funkciji toplinske obrade zraka kao konačnog prenositelja energije u procesu klimatiziranja vagona. Klima agregat, rashladni agregat, kanalna razvodna mreža i elektronika za upravljanje i regulaciju su četiri osnovne komponente klima uređaja.

d.) Vakuum WC

Primjena zatvorenog sustava WC-a bila je neophodna zbog velikih brzina vožnje, te je ovakvom koncepcijom izbjegnuto izlaganje putnika djelovanjima zraka pri mimoilaženjima vlakova [9]. Osim gore navedenih poboljšanja, spriječeno je i prljanje donjeg postroja vagona (okretna postolja, kočno polužje, razne instalacije itd.) što olakšava održavanje. Fekalije se skupljaju u središnji spremnik smješten pod vagon i iz njega prazne pomoću posebnog uređaja. Za normalno funkcioniranje sustava moraju se ispuniti uvjeti napajanja instalacija stlačenim zrakom od preko 6 bara, napajanje vodom za ispiranje, napajanje električnom energijom od 220 V, dok temperatura u prostoriji WC-a mora biti viša od 5°C. Magnetni ventil ejektora se uključuje aktiviranjem tipkala za ispiranje. Stlačeni tlak prolaskom kroz ejektor stvara podtlak u spremniku fekalija i spojnom cjevovodu prema WC školjkama, a u trenutku postizanja dovoljnog podtlaka (0,2 bara podtlaka) se uključuje vakuumski prekidač. Uključenjem vakuumskog prekidača nalog za ispiranje WC školjke daje upravljačka automatika. Uključuje se brizgalica za vodu te kroz sapnice ispire WC školjku. Nakon 1,5 sekunde otvore se zaklopke za pražnjenje fekalija, a sadržaj WC školjke stvorenim podtlakom se usisava u spremnik fekalija.

4.2. ODRŽAVANJE PUTNIČKIH VAGONA HRVATSKIH ŽELJEZNICA

Za visokokomforne putničke vagone Hrvatskih željeznica postoji redovito (preventivno) i izvanredno (korektivno) održavanje:

a.) Redovito održavanje

Redovito održavanje putničkih vagona obuhvaća:

- kontrolne preglede (P1 - mjesečni i MP - mali popravak)
- redovite popravke (SP - srednji popravci i VP – veliki popravci)

Mjesečni pregledi P1 i Mali popravci MP (godišnji pregled) propisani su Uputom za kontrolne preglede i redovite popravke željezničkih vozila (241-1), s radovima prema tablici 4.2.

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Tablica 4.2. Održavanje putničkih vagona prema uputi 241-1 Hrvatskih željeznica

	OPIS RADOVA	P1	MP
1. OSOVINSKI SKLOPOVI			
	1. pregled monoblok ploča (kotača),	X	X
	2. mjerenje geometrijskih karakteristika kotača (debljina vijenca, visina vijenca i Qr veličina, razmak između kotača)	X	X
	3. pregled kućišta ležajeva (vizualno)	X	X
	4. pregled ispravnosti položaja lisnatih opruga	X	X
	5. ispitivanje tijela osovine ultrazvukom	-	X
	6. pregled kočnih diskova	X	X
	7. mjerenje geometrije osovinskog sklopa prema mjernoj listi	-	X
2. OKRETNOSTI POSTOLJA			
	1. pregled okvira okretnog postolja (posebna pozornost na stanje zavarenih spojeva, mogućnost pukotina),	X	X
	2. pregled pripadajućih elemenata okretnog postolja (središnji svornjak- veza s sandukom vagona, amortizeri, vodilice (opruge), torzione poluge (opruge), opruge kolijevke,....)	X	X
	3. mjerenje veličina okretnog postolja prema mjernoj listi	-	X
3. RUČNA KOČNICA I KOČNO POLUŽJE			
	1. pregled istrošenosti trošenja kočnih papuča i umetaka	X	X
	2. pregled polužja kočnih cilindara, čahura, svornjaka i nosača kočnih umetaka	X	X
	3. ispitivanje funkcionalnosti i sigurnosti dijelova ručne kočnice	X	X
	4. podmazivanje dijelova kočnog polužja	X	X
	5. pregled ispravnosti rada mjenjača vrste kočenja i uređaja za isključivanje kočnice	X	X
	6. provjera rada magnet elektromagnetske kočnice	X	X

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

4. ZRAČNA KOČNICA			
	1. ispitivanje kočnice (prema uputi 245 HŽ-a)	-	X
	2. ispitivanje kočnice „A“ prema uputi 233 HŽ-a	X	-
	3. pregled pričvršćenosti i ispravnosti ventila, spremnika zraka, gumenih crijeva, elektrovodova magnetske kočnice itd.	X	X
5. VRATA I UREĐAJI ZA ZATVARANJE VRATA			
	1. ispitivanje funkcionalnosti i sigurnosti vanjskih vrata (sa automatskim zatvaranjem) i brave na vratima	X	X
	2. podmazivanje kliznih dijelova vrata	X	X
	3. ispitivanje pogonske sigurnosti brave	X	X
	4. ispitivanje elektropneumatskog upravljanja vratima, te dijelova	X	X
	5. ispitati ispravnost tipke za pomicanje vrata	X	X
6. UNUTRAŠNJI PROSTOR VAGONA			
	1. pregled ispravnosti unutrašnjosti odjeljaka, predulaza i hodnika	X	X
	2. čišćenje predulaza vagona, odjeljaka i hodnika	X	X
	3. pregled unutrašnjosti aparata za gašenje požara	X	X
	4. pregled ispravnosti mehanizama sjedala i ležajeva	X	X
	5. pregled i čišćenje preklopnih stolova	X	X
7. UREĐAJI OPSKRBE ELEKTRIČNOM ENERGIJOM (statički pretvarači)			
	1. provjera otvora za ispuštanje kondenzata statičkog pretvarača napona	X	X
	2. provjera i čišćenje rashladnih tijela energetskih modula, zračnih kanala i magnetskih komponenti statičkog pretvarača napona	X	X
	3. vizualna kontrola ugrađenih modula, sklopova, prigušnica, transformatora i ventilatora statičkog pretvarača napona	X	X
	4. provjera otvora za ispuštanje kondenzata statičkog pretvarača napona	X	X
	5. provjera podešenosti strujnih zaštitnih automatskih sklopki statičkog pretvarača napona	X	X
	6. provjera releja i tiskanih veza relejnih kartica statičkog pretvarača	X	X
	7. mjerenje veličina stratičkog pretvarača prema mjernoj listi	-	X

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

8. RASVJETNA TIJELA			
	1. provjera upravljanja glavnom i pomoćnom rasvjetom vagona	X	X
	2. provjera rasvjetnih tijela i prekidača električne rasvjete, svjetiljki završnog signala, rasvjete za čitanje, nužne rasvjete, mjernih instrumenata	X	X
	3. provjera minimalnog releja i ograničavača napona rasvjete	X	X
	4. provjera električne instalacije i spojnih mjesta	X	X
	5. ispitivanje rasvjete i upisivanje u mjernu listu	-	X
	6. čišćenje ormara elektrorasvjete	X	X
	7. provjera stanja rezervnih dijelova rasvjete (žarulje, fluo –cijevi, osigurači....)	X	X
	8. provjera spojeva od izvora energije do razvodnog ormara	X	X
	9. Pregled kompletnog sustava rasvjete prema mjernoj listi	-	X
9. ELEKTRIČNO GRIJANJE			
	1. pregled, čišćenje i podmazivanje spojnih naprava grijanja	X	X
	2. ispitivanje rada birača napona	X	X
	3. pregled osigurača strujnih krugova i strujnih krugova	X	X
	4. ispitivanje rada grijalica	X	X
	5. ispitivanje rada glavnog prekidača, termostatske regulacije	X	X
	6. čišćenje kanala toplozračnog grijanja	-	X
	7. ispitivanje električnog grijanja jednim od RIC napona	X	X
	8. pregled četkica na motor ventilatoru	X	X
	9. pregled kompletnog sustava grijanja prema mjernoj listi	-	X
10 Vakuum WC i instalacije			
	1. ispitivanje rada vakuum WC-a	X	X
	2. pregled ispravnosti svih ventila vakuum WC-a	X	X

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

11.RAZGLASNI UREĐAJI			
	1. ispitivanje ispravnosti razglasa	X	X
	2. pregled ispravnosti zvučnika, regulatora razglasa i sklopova	X	X
	3. čišćenje spojnih mjesta i sprava	X	X
	4. pregled kompletnog razglasnog uređaja prema mjernoj listi	-	X
12. AKUMULATORSKE BATERIJE			
	1. pregled akumulatorskih baterija	X	X
	2. punjenje baterija	X	X
	3. mjerenje gustoće elektrolita i napona baterije	X	X
	4. pregled akumulatorskih baterija prema mjernoj listi	-	X
	5. pregled ispravnosti statičkog punjača akumulatorskih baterija	X	X
13. PROZORI			
	1. pregled ispravnosti prozora	X	X
	2. ispitivanje zabrtvljenosti prozora prema mjernoj listi	-	X
	3. podmazivanje vodilica prozora		
14. KLIMA UREĐAJI			
	1. zamjena filtera za zrak	X	X
	2. pregled cjelokupnog klima uređaja (klima agregat)	X	X
	3. pregled isparivača i ventila u klima agregatu	-	X
	4. provjera razine freona u rashladnom agregatu	X	X
	5. provjera razine ulja u kompresoru rashladnog agregata	X	X
	6. pregled ventilatorskog kruga, crijeva visokog i niskog tlaka rashladnog agregata	X	X
	7. pregled klima uređaja prema mjernoj listi	-	X

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Troškovi održavanja za pojedinu seriju putničkih vagona pojedinih serija pri kontrolnim pregledima P1 (mjesečni redovni pregled), MP (godišnji redovni pregled) i IP-izvanredni popravak prikazani su u tablici 4.3.

Tablica 4.3. Troškovi održavanja za pojedinu seriju vagona (u kunama) bez cijene utrošenog materijala i dodatnih radova

Serijski broj vagona	P1	MP	IP
Aeelt (10-70)	7259,42	87377,16	3326,43
Aeelmt (10-90)	7012,97	122634,60	3374,17
Bee (20-70)	8489,64	101403,48	3950,10
Beelmt (20-90)	7012,97	122585,66	3374,17
Beelt (20-70)	7363,62	91975,95	3374,17
Beel (20-70)	6944,25	120841,05	2948,32
Bee (29-70)	7041,36	123944,23	2745,07
ABeel (39-70)	6959,11	123944,23	2589,69
WReelmt (87-90)	7259,76	126500,67	3375,55

Prema UIC (međunarodni željeznički pravilnik) objavi svaki vagon se sastoji od konstruktivnih grupa prema tablici 4.4., te na osnovu tih podjela se razmatraju i analiziraju kvarovi.

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

4.3. ANALIZA REZULTATA UPORABOM METODA POSLOVNOG IZVJEŠĆIVANJA

Tablica 4.4. Broj kvarova na klimatiziranim putničkim vagonima u razdoblju 1.1.2005.-
1.1.2007.godine

	Agregati na vagonu	Serije vagona						Zbroj kvarova
		Aeelmt, Beelmt, WReelmt	Aeelt Bee	Beelt (20-70)	Bee WRee	ABeel Beel	Wlee Bee (29-70)	
1	Osovinski sklopovi	4	2	3	4	2	2	17
2	Okretna postolja	14	18	3	1	9	3	48
3	Ručna kočnica i kočno polužje	3	11	0	4	12	7	37
4	Zračna (pneumatska) kočnica	20	54	6	11	17	10	118
5	Vrata i uređaji za zatvaranje vrata	77	136	10	22	13	11	269
6	Unutrašnji inventar	2	12	1	7	0	4	26
7	Uređaji opskrbe (statički pretvarači)	45	133	15	3	13	10	219
8	Rasvjetna tijela	14	19	1	3	14	8	59
9	Električno grijanje	20	53	8	30	26	12	149
10	Vakuum WC i Vodovodne instalacije	95	165	15	10	6	16	307
11	Razglasni uređaji	14	55	2	0	4	4	79
12	Akumulatorske baterije	12	27	2	2	6	5	54
13	Prozori	5	17	0	13	20	7	62
14	Klima uređaji	59	102	17	0	6	9	193
	PROIZVOĐAČ STATIČKOG PRETVARAČA	Siemens	Končar VIS 50-1	Končar VIS 50	Pintch 75	MGVG 3000	Pintch Bamag 15	

4.3.1. PRIMJENA KLASTER POSTUPKA U ANALIZI KVAROVA

Za klasifikaciju grupa kvarova na serijama klimatiziranih putničkih vagona odabrana je klaster analiza. Cilj klaster analize je grupiranje subjekata u klastere, tako da je svaki klaster što je

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

više moguće homogen s obzirom na varijable klasteriranja. Prvi korak u klaster analizi je odabir mjere sličnosti, a zatim slijede donošenje odluke o tipu tehnike klasteriranja koju ćemo koristiti (hijerarhijska ili nehijerarhijska). Treći korak je odabir tipa metode klasteriranja za odabranu tehniku (npr. metoda najbližeg susjeda ili prosječne udaljenosti).

4.3.1.1. Metoda najbližeg susjeda (Kvadratni Euklid)

Prema formuli (3) za izračun udaljenosti između grupa klastera po Euklidovom kvadratu:

$$d(x, y) = \sum_{i=1}^p (x_i - y_i)^2$$

gdje je $d(x, y) = d$ (broj kvarova konstruktivne grupe 1, broj kvarova konstruktivne grupe 2),
a prema tablici 4.4.

dobivamo izračun udaljenosti između grupa (npr. 1 i 2):

$$d(1, 2) = (4-14)^2 + (2-18)^2 + (3-3)^2 + (4-1)^2 + (2-9)^2 + (2-3)^2 = 415$$

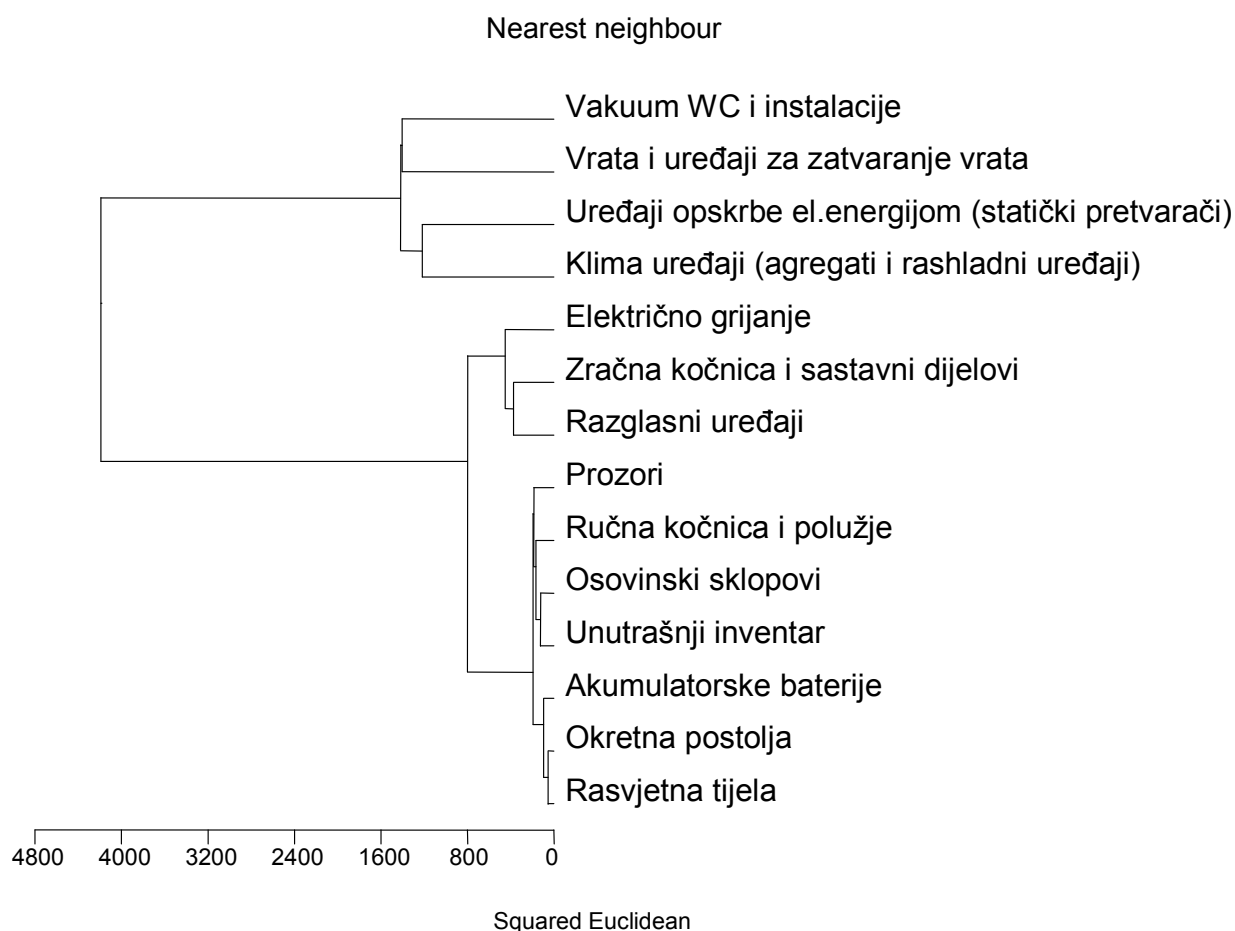
Sve ostale izračunate vrijednosti navedene su u matrici na slici 31.

	Konstruktivne grupe (1-14)													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	415													
3	216	213												
4	3307	1554	2257											
5	23860	18463	21542	10127										
6	125	302	165	2454	21525									
7	19172	14399	16884	7027	1420	16907								
8	574	59	192	1363	18110	421	14158							
9	4234	2497	3014	451	10376	3323	7976	2130						
10	35242	28573	32558	18185	1408	32443	3658	28250	19034					
11	2934	1397	2150	379	11208	2059	7340	1422	1524	19078				
12	719	100	385	1036	16655	388	12569	143	2009	26287	797			
13	665	372	185	1652	19591	480	15439	223	1935	30515	1963	474		
14	13302	9323	11762	4189	2066	11715	1220	9244	5312	5418	4488	8079	10799	

Slika 31. Matrica udaljenosti prema kvadratnom Euklidu za grupe kvarova

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Prema rezultatima izračuna udaljenosti između grupa klastera na slici 31., vidljivo je da je, a što prikazuje i dendrogram na slici 32. najmanja izračunata udaljenost između grupa kvarova 2 i 8, a iznosi 59. Najmanja udaljenost znači i da su grupe najbližije, tj. blisko povezane.



Slika 32. Dendrogram kvarova (metoda najbližeg susjeda, kvadratni Euklid)

Dendrogram na slici 32. prikazuje rezultat klasteriranja četrnaest grupa kvarova na klimatiziranim putničkim vagonima koristeći metodu najbližeg susjeda i kvadratne Euklid udaljenosti. Na početku, svaki od četrnaest vrsta kvarova oblikuje separatne (posebne)

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

klastera. Prvi klaster se je stvorio spajajući grupe okretna postolja (2) i *rasvjetna tijela* (8) gdje je izmjerena najmanja udaljenost. Dendrogram prikazuje podjelu kvarova u dvije grupe:

- 1.grupa – kvarovi na statičkim pretvaračima, klima uređajima, vakuum WC-u i vratima,
- 2.grupa – kvarovi na okretnim postoljima, unutrašnjosti, osovinskim sklopovima, rasvjeti, ručnoj i zračnoj kočnici, akumulatorskim baterijama, prozorima i električnom grijanju.

Grupa 1 sadrži komponente sa najviše kvarova, a to su: vrata i uređaji za zatvaranje vrata (5), statički pretvarači (7), vakuum WC i vodovodne instalacije (10) i klima uređaji (14)

Aglomeracijski raspored grupa klastera

Tablica 4.5. Aglomeracijski raspored klastera udaljenosti pojedinih grupa kvarova
metodoma klastera (najbliži susjed)

Node	Group 1	Group 2	Koeficijent fuzije	Objects in group
1	Okretna postolja	Rasvjeta	59	2
2	Node 1	Akumulatorske baterije	100	3
3	Osovinski sklopovi	Unutrašnjost	125	2
4	Node 3	Ručna kočnica	165	3
5	Node 4	Prozori	185	4
6	Node 5	Node 2	192	7
7	Zračna kočnica	Razglasni uređaji	379	2
8	Node 7	El.grijanje	451	3
9	Node 6	Node 8	797	10
10	Statički pretvarači	Klima uređaji	1220	2
11	Vrata i uređaji za zatv.	Vakuum WC	1408	2
12	Node 11	Node 10	1420	4
13	Node 9	Node 12	4189	14

Tablica 4.5. pokazuje aglomeracijski raspored spajanja klastera. U stupcu *koeficijent fuzije* je data vrijednost statističke udaljenosti (ili sličnosti grupa) korištene za formiranje klastera. Od navedenih vrijednosti u tablici, manji koeficijent objašnjava da postoji velika sličnost između

grupa (npr. grupe okretna postolja (2) i rasvjetna tijela (8)), dok najveći koeficijent objašnjava da ne postoji tako velika povezanost između grupa (grupe osovinski sklopovi (1) i vrata i uređaji za zatvaranje vrata (5)). Aglomeracijski raspored počinje koracima povezivanja između dva klastera, a to su u prvom koraku grupe okretna postolja (2) i rasvjetna tijela (8). Nakon navedenog spajanja ova grupa prelazi na korak 2. gdje se klaster spaja sa grupom grupom akumulatorske baterije (12) itd., pri čemu raste koeficijent fuzije prema tablici 4.5.

Izračun koeficijenta fuzije sličnosti grupa kvarova metodom najbližeg susjeda (Euklid udaljenost)

Za izračun koeficijenta fuzije sličnosti grupa kvarova neophodno je imati matricu sličnosti $n \times n$ (slika 29. 14×14) koja sadrži udaljenosti grupa. Algoritam izračuna koeficijenta fuzije se sastoji od slijedećih koraka:

1. Inicijalizacija

Konstrukcija matrice veličine 14×14

2. **Konstrukcija vektora v** , pri čemu je $v_r = \text{minimum člana retka } u_r \text{ i vektor } c$, gdje je c odgovarajući indeks stupca minimuma u_r . Na kraju je potrebno konstruirati vektor n brojem objekta u svakom klasteru.

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

V =	415	c =	1	n =	1
	213		2		1
	1554		2		1
	10127		4		1
	125		1		1
	7027		4		1
	59		2		1
	451		4		1
	1408		5		1
	379		4		1
	100		2		1
	185		3		1
	1220		7		1

Korak 1. Potrebno je odrediti minimalnu vrijednost udaljenosti u matrici (min = 59)

	Konstruktivne grupe (1-14)													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	415													
3	216	213												
4	3307	1554	2257											
5	23860	18463	21542	10127										
6	125	302	165	2454	21525									
7	19172	14399	16884	7027	1420	16907								
8	574	59	192	1363	18110	421	14158							
9	4234	2497	3014	451	10376	3323	7976	2130						
10	35242	28573	32558	18185	1408	32443	3658	28250	19034					
11	2934	1397	2150	379	11208	2059	7340	1422	1524	19078				
12	719	100	385	1036	16655	388	12569	143	2009	26287	797			
13	665	372	185	1652	19591	480	15439	223	1935	30515	1963	474		
14	13302	9323	11762	4189	2066	11715	1220	9244	5312	5418	4488	8079	10799	

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Iz priložene matrice sličnosti najmanja udaljenost iznosi 59, a izračunata je između grupa (2, 8). U drugom koraku se formira klaster (2, 8)

Korak 2.

	Konstruktivne grupe (1-14)													
	1	(2,8)	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	415													
3	216	213												
4	3307	1554	2257											
5	23860	18463	21542	10127										
6	125	302	165	2454	21525									
7	19172	14399	16884	7027	1420	16907								
8														
9	4234	2497	3014	451	10376	3323	7976							
10	35242	28573	32558	18185	1408	32443	3658		19034					
11	2934	1397	2150	379	11208	2059	7340		1524	19078				
12	719	100	385	1036	16655	388	12569		2009	26287	797			
13	665	372	185	1652	19591	480	15439		1935	30515	1963	474		
14	13302	9323	11762	4189	2066	11715	1220		5312	5418	4488	8079	10799	

Pomoću formule za postupak *pojedinačne veze (single linkage)* prema tablici 3.3. [8]

$d = 0.5 * \{d(i, p) + d(i, q) - |d(i, p) - d(i, q)|\}$ izračunaju se nove vrijednosti stupca spajanja klastera (2, 8).

$$d((2, 8), 3) = 0,5 * (213 + 192 - |213 - 192|) = 0,5 * (405 - 21) = 192$$

$$d((2, 8), 4) = 0,5 * (1554 + 1363 - |1554 - 1363|) = 0,5 * (2917 - 191) = 1363 \text{ itd.}$$

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Korak 3. Ponovo se traži minimalna vrijednost matrice (min =100)

	Konstruktivne grupe (1-14)													
	1	(2, 8)	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	415													
3	216	192												
4	3307	1363	2257											
5	23860	18110	21542	10127										
6	125	302	165	2454	21525									
7	19172	14158	16884	7027	1420	16907								
8														
9	4234	2130	3014	451	10376	3323	7976							
10	35242	28250	32558	18185	1408	32443	3658		19034					
11	2934	1397	2150	379	11208	2059	7340		1524	19078				
12	719	100	385	1036	16655	388	12569		2009	26287	797			
13	665	223	185	1652	19591	480	15439		1935	30515	1963	474		
14	13302	9244	11762	4189	2066	11715	1220		5312	5418	4488	8079	10799	

I dalje se postupak ponavlja prema formuli za pojedinačnu vezu (single linkage) pri čemu se dobiva konačni redoslijed spajanja klastera. Ovdje je vidljivo da prvo dolazi do spajanja klastera (2,8) na koji se nadovezuje grupa 12, što je vidljivo i iz aglomeracijskog rasporeda klastera prema tablici 4.5.

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

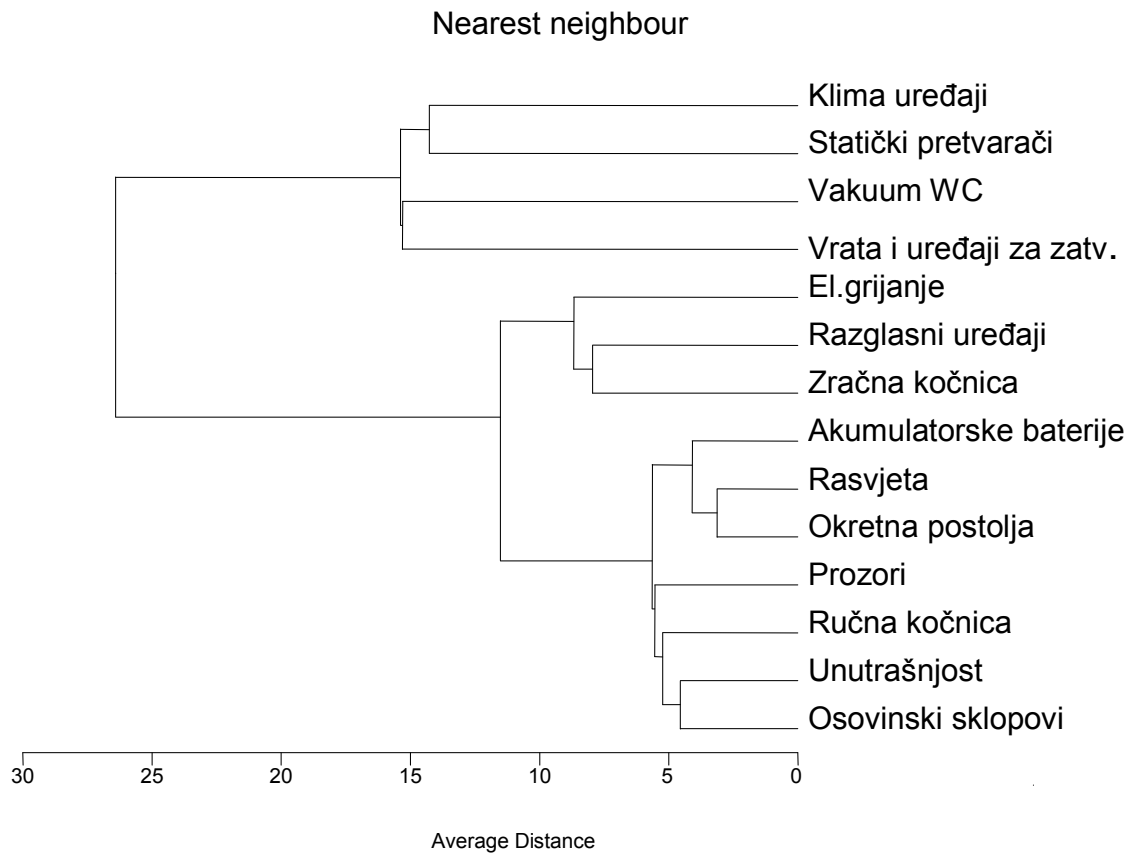
4.3.1.2. Metoda najbližeg susjeda, prosječna udaljenost (average distance)

Sve ostale izračunate vrijednosti navedene su u matrici na slici 33.

	Konstruktivne grupe (1-14)													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	8,32													
3	6,00	5,96												
4	23,48	16,09	19,35											
5	63,06	55,47	59,92	41,08										
6	4,56	7,09	5,24	20,23	59,89									
7	56,53	48,99	53,05	34,22	15,38	53,08								
8	9,78	3,14	5,66	15,07	54,94	8,38	48,58							
9	26,56	20,40	22,41	8,67	41,59	23,54	36,46	18,84						
10	76,64	69,01	73,67	55,05	15,32	73,53	24,69	68,62	56,32					
11	22,11	15,26	18,93	7,95	43,22	18,53	34,98	15,40	15,94	56,39				
12	10,95	4,08	8,01	13,14	52,69	8,04	45,77	4,88	18,30	66,19	11,53			
13	10,53	7,87	5,55	16,59	57,14	8,94	50,73	6,10	17,96	71,32	18,09	8,88		
14	47,09	39,42	44,28	26,42	18,56	44,19	14,26	39,25	29,75	30,05	27,35	36,70	42,42	

Slika 33. Matrica udaljenosti prema metodi prosječne udaljenosti (average distance) za grupe kvarova

Prema rezultatima izračuna udaljenosti između grupa klastera metodom najbližeg susjeda i prosječnom (average) udaljenosti, vidljivo je da je, a što prikazuje i dendrogram najmanja udaljenost izračunata između grupa kvarova 2 i 8, a iznosi 3,14. Najmanja udaljenost znači i da su grupe najsličnije, tj. blisko povezane.



Slika 34. Dendrogram kvarova (metoda najbližeg susjeda, average distance)

Dendrogram na slici 34. prikazuje rezultat klasteriranja četrnaest grupa kvarova na klimatiziranim putničkim vagonima koristeći metodu najbližeg susjeda i prosječne udaljenosti. Na početku, svaki od kvarova oblikuje separatne (posebne) klastere. Prvi klaster se je stvorio spajajući grupe kvarova 2-okretna postolja i 8 - rasvjetna tijela gdje je izmjerena najmanja udaljenost od 3,14. Nakon toga se grupa 12 - akumulatorske baterije spaja sa oblikovanim klasterom (2, 8) itd.

Dendrogram prikazuje podjelu kvarova u dvije grupe:

- 1.grupa – kvarovi na statičkim pretvaračima, klima uređajima, vakuum WC-u i vratima,
- 2.grupa – kvarovi na okretnim postoljima, unutrašnjosti, osovinskim sklopovima, rasvjeti,

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

ručnoj i zračnoj kočnici, akumulatorskim baterijama, prozorima i električnom grijanju.

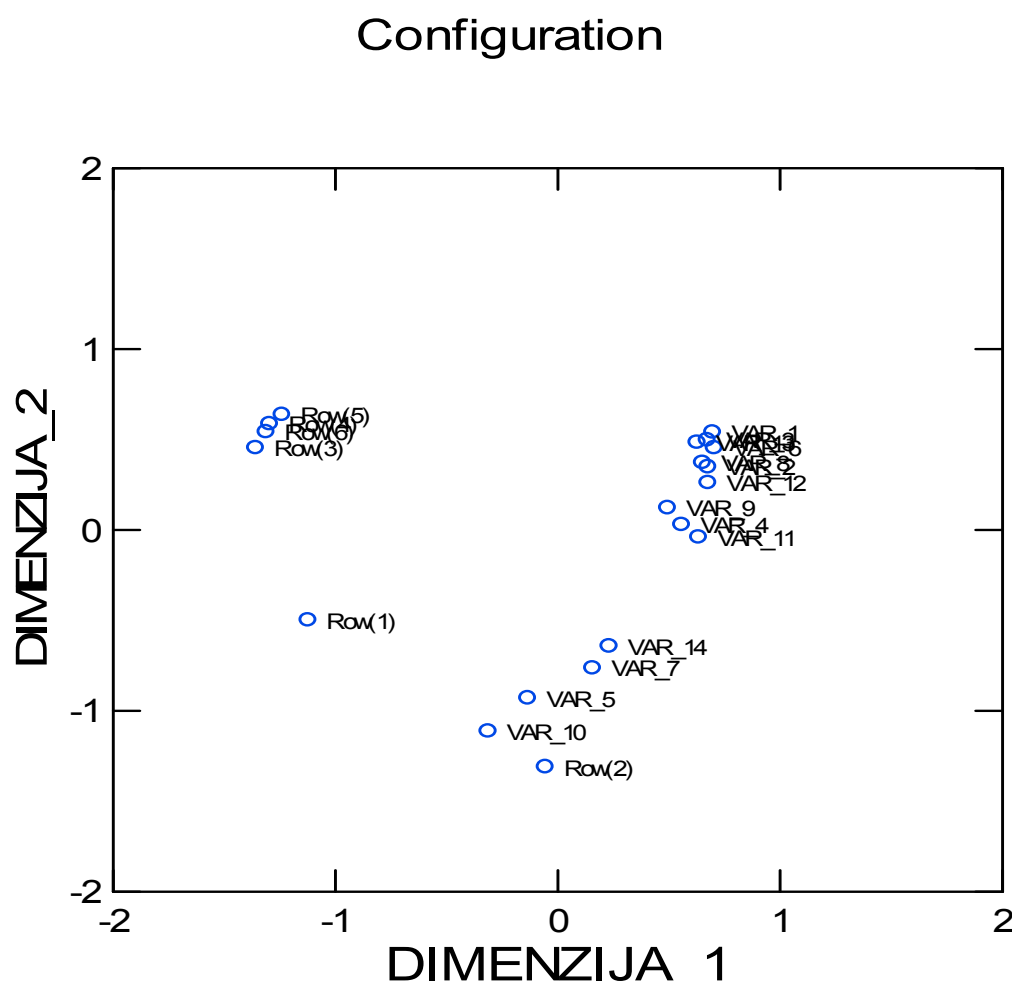
Tablica 4.6. Aglomeracijski raspored klastera udaljenosti pojedinih grupa kvarova metodom prosječne udaljenosti (average distance)

Node	Group 1	Group 2	Koeficijent fuzije	Objects in group
1	Okretna postolja	Rasvjeta	3,136	2
2	Node 1	Akumulatorske baterije	4,082	3
3	Osovinski sklopovi	Unutrasnjost	4,564	2
4	Node 3	Ručna kočnica	5,244	3
5	Node 4	Prozori	5,553	4
6	Node 5	Node 2	5,657	7
7	Zračna kočnica	Razglasni uređaji	7,948	2
8	Node 7	El.grijanje	8,670	3
9	Node 6	Node 8	11,525	10
10	Statički pretvaraci	Klima uređaji	14,259	2
11	Vrata i uređaji za zatv.	Vakuum WC	15,319	2
12	Node 11	Node 10	15,384	4
13	Node 9	Node 12	26,423	14

Tablica 4.6. pokazuje aglomeracijski raspored spajanja klastera metodom najbližeg susjeda i prosječnom udaljenosti. U stupcu *koeficijent fuzije* je data vrijednost statističke udaljenosti (ili sličnosti grupa) korištene za formiranje klastera. Od navedenih vrijednosti u tablici, manji koeficijent objašnjava da postoji velika sličnost između grupa (npr. grupe 2 – okretna postolja i 8 - rasvjetna tijela), dok najveći koeficijent objašnjava da ne postoji tako velika povezanost između grupa (grupe 1-osovinski sklopovi i 5- vrata i uređaji za zatvaranje vrata). Aglomeracijski raspored počinje koracima povezivanja između dva klastera, a to su u prvom koraku grupe 2 – okretna postolja i 8 - rasvjetna tijela. Nakon navedenog spajanja ova grupa prelazi na korak 2. gdje se klaster spaja sa grupom 12 –akumulatorske baterije itd. Rezultati su isti kao i u prethodnom postupku istraživanja koristeći metodu najbližeg susjeda i kvadratnog Euklida.

4.3.2. Primjena postupka multidimenzijanskog skaliranja u analizi kvarova

Pomoću multidimenzionalnog skaliranja (MDS analysis) i primijenjene Kruskal metode, dobili smo rješenja prema slici 35. i tablici. Rezultati su dobiveni uporabom software paketa SYSTAT 12.



Slika 35. Multidimenzionalni prikaz grupa kvarova (u logaritamskom mjerilu)

Slika 35. prikazuje konfiguraciju klastera četrnaest grupa kvarova klimatiziranih putničkih vagona, pri čemu su se izdvojile u jedan tzv. klaster grupu kvarovi 5- vrata i uređaji za

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

zatvaranje vrata, 10 - vakuum WC i vodovodne instalacije, 7- statički pretvarači (uređaji opskrbe električnom energijom, i 14 - klima uređaji. Rezultati su pokazali rješenja ista kao i klaster analiza u točkama 4.3.1.1 i 4.3.1.2. Na slici je vidljivo da se ove četiri grupacije kvarova komponenti najviše javljaju kod vagona serije 10-90, 20-90 i 87-90 (Row 1) koji imaju ugrađeni statički pretvarač tvrtke Siemens EVA/B i 10-70 i 20-70 (Row 2) koji imaju ugrađeni statički pretvarač tvrtke Končar VIS 50-1. Za Kruskal metodu upotrebljavaju se STRESS formule 1 :

$$STRESS\ 1 = \sqrt{\frac{\sum_k \sum_l (d_{kl} - \hat{d}_{kl})^2}{\sum_k \sum_l d_{kl}^2}}$$

Tablica 4.7. Prikaz rješenja metodom multidimenzijskog skaliranja

Kruskal Metoda

Podaci su analizirani na bazi sličnosti uporabom kvadratne Euklidove udaljenosti
Minimizirani Kruskal-ov STRESS (formula 1) u 2 dimenzije

Povijest iteracije		Koordinate u 2 dimenzije		
Iteracija	STRESS	Variable (grupe agregata)	Dimenzije	
			1	2
0	0,187	VAR_1 – osovinski sklopovi	0,700	0,538
1	0,115	VAR_2 – okretna postolja	0,678	0,347
2	0,085	VAR_3 – ručna kočnica i kočno polužje	0,676	0,496
3	0,070	VAR_4 – zračna (pneumatska) kočnica	0,560	0,028
4	0,061	VAR_5 – vrata i uređaji za zatvaranje vrata	-0,132	-0,931
5	0,055	VAR_6 – unutrašnji inventar	0,707	0,452
6	0,050	VAR_7 – statički pretvarači (uređaji opskrbe električnom energijom)	0,160	-0,765
7	0,047	VAR_8 – rasvjetna tijela	0,654	0,371
8	0,044	VAR_9 – električno grijanje	0,497	0,121
9	0,041	VAR_10 – vakuum WC i vodovodne instalacije	-0,310	-1,115
0	0,045	VAR_11 – razglasni uređaji	0,636	-0,041
1	0,036	VAR_12 – akumulatorske baterije	0,679	0,260
2	0,034	VAR_13 - prozori	0,629	0,482
3	0,033	VAR_14 – klima uređaji	0,234	-0,644
		Row(1) – grupa vagona serije Aeelmt, Beelmt i WReelmt	-1,121	-0,500
		Row(2) – grupa vagona serije Aeelt (10-70) i Bee (20-70)	-0,052	-1,311
		Row(3) – grupa vagona serije Beelt (20-70)	-1,356	0,452
		Row(4) – grupa vagona serije Bee i WRee	-1,293	0,584
		Row(5) – grupa vagona serije ABeel i Beel	-1,237	0,636
		Row(6) – grupa vagona serije Wlee i Bee (29-70)	-1,309	0,540

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

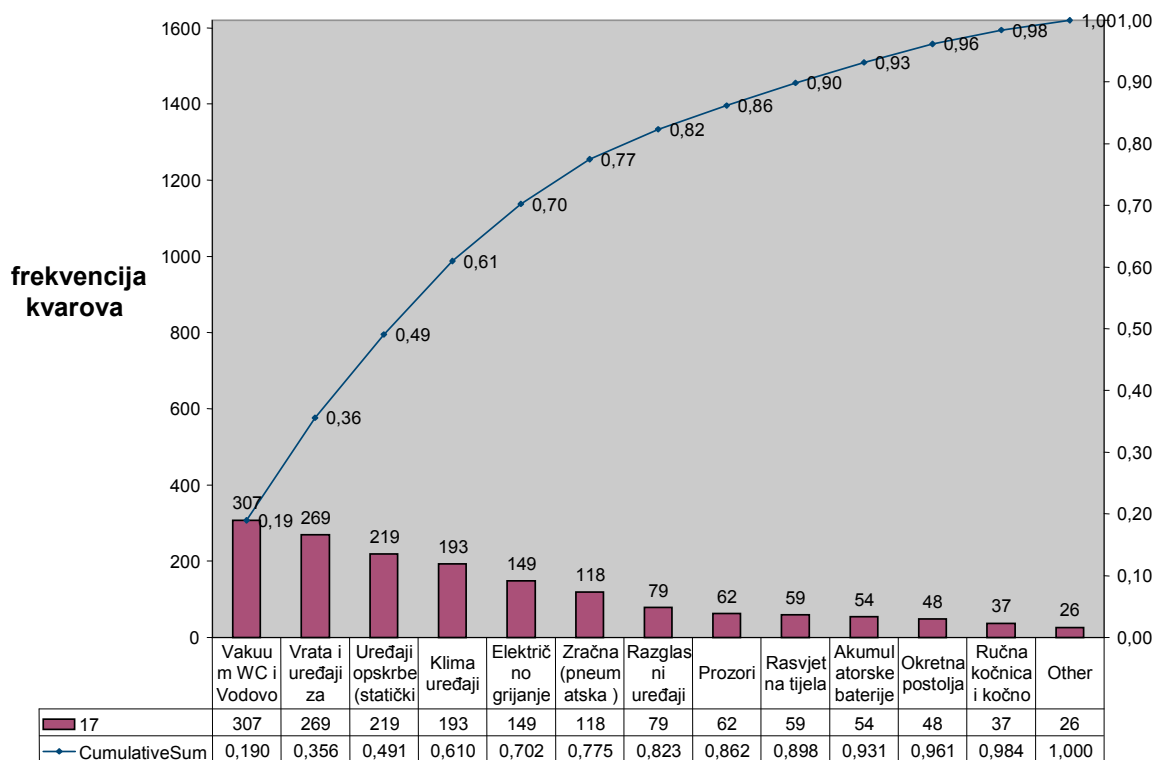
Tablica 4.7. (lijevi dio) prikazuje postupke iteracije za funkciju STRESS (standardno odstupanje (sume) kvadrata prema formuli Kruskala STRESS1, pri čemu je dobivena vrijednost nakon treće iteracije 0.033. U idealnom slučaju prihvatanja, ova vrijednost iznosi 0. Ukoliko je funkcija STRESS vrlo mala, npr. Ukoliko je manja od 0.01 tada se objekti (grupe) grupiraju oko centra (središta) koordinatnog sustava. Primjer izračuna STRESS veličine analitičkim putem prikazana je tablicom 4.8

Tablica 4.8. Prikaz izračuna funkcije STRESS1 multidimenzijskog skaliranja

Grupa kvarova	Udaljenost grupa prema kvadratnom Euklidu	disparitet grupa	kvadratna razlika udaljenosti i dispariteta	Udaljenost na kvadrat
	d_{kl}	$\hat{d}_{kl}$	$(d_{kl} - \hat{d}_{kl})^2$	d_{kl}^2
(1,2)	415	315,5	9900,25	172225
(1,3)	216		9900,25	46656
.....				
.....				
(2,3)	213	883,5	449570,25	45369
(2,4)	1554		449570,25	2414916
.....				
.....				
(8,12)	143	183	1600	20449
(8,13)	223		1600	49729
.....				

$$STRESS\ 1 = \sqrt{\frac{\sum_k \sum_l (d_{kl} - \hat{d}_{kl})^2}{\sum_k \sum_l d_{kl}^2}} = 0,033$$

4.3.3. PRIMJENA PARETO POSTUPKA U ANALIZI KVAROVA

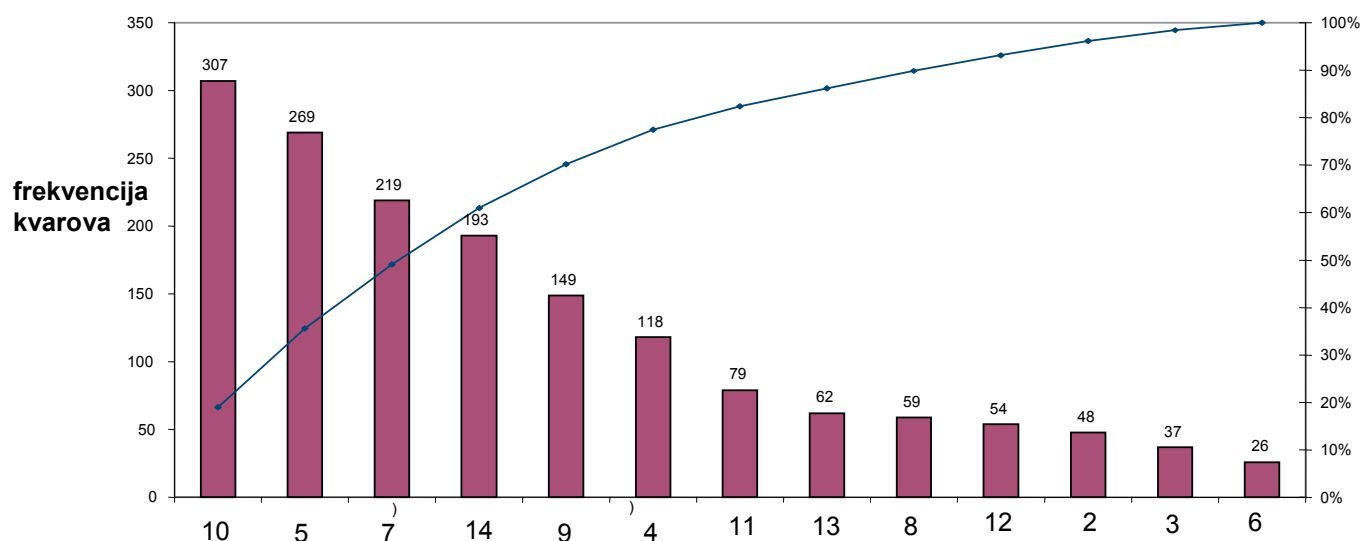


Grupe kvarova i kumulativne sume po konstruktivnim grupama prema tablici 4.4.

Slika 36. Pareto analiza frekvencije kvarova na klimatiziranim putničkim vagonima

Slike 36. i rezultati prikazani slikom 37. daju analizu kvarova četrnaest grupa komponenti na klimatiziranim putničkim vagonima Hrvatskih željeznica, gdje su grupe poredane po najvećim količinama i postotnim udjelima kvarova. Najveća količina kvarova je u grupi 10 – vakuum WC i vodovodne instalacije sa količinom od 307 kvara i kumulativnim postotkom od 19% . Nakon toga, slijede grupe kvarova komponenti 5 - vrata i uređaji za zatvaranje sa 269 kvara i kumulativnim postotkom od 16,6 % , 7- statički pretvarači, te 14 - klima uređaji itd.

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA



Konstruktivne grupe agregata prema tablici 4.4

grupe -agregati	Rang	frekvencija	kumulativni zbroj	postotak	kumulativni postotak
10 -Vakuum WC	1	307	307	19,00	19,00
5 -Vrata i uređ. za zatv.	2	269	576	16,60	35,60
7- Statički pretvarači	3	219	795	13,50	49,10
14- Klima uređaji	4	193	988	11,90	61,00
9 -Električno grijanje	5	149	1137	9,02	70,02
4- Zračna kočnica	6	118	1255	7,48	77,50
11- Razglasni uređaji	7	79	1334	4,80	82,30
13- Prozori	8	62	1396	3,78	86,20
8 - Rasvjetna tijela	9	59	1455	3,60	89,80
12 - Aku. baterije	10	54	1509	3,30	93,10
2- Okretno postolje	11	48	1557	3,00	96,10
3 -Ručna kočnica	12	37	1594	1,30	97,40
6- Unutrašnjost	13	26	1620	1,12	98,52
1- Osovinski sklopovi	14	17	1637	1,06	100,0
UKUPNO		1637			

Slika 37. Prikaz rezultata frekvencije kvarova na klimatiziranim putničkim vagonima

Iz dobivenih rezultata se vidi potvrda dobivenih rješenja Pareto analizom koje su također izlučile kvarove grupa 5, 7, 10 i 14 sa najvećom frekvencijom.

Zaključak istraživanja kvarova metodama poslovne inteligencije.

Metode klaster analize, multidimenzijskog skaliranja i kontrolna Pareto metoda primijenjene u istraživanjima su dale kao rezultat kvarove istih grupa uređaja 5, 7, 10 i 14, te se uređaji smatraju kritičnima i ključnima za daljnja razmatranja zbog najučestalijih kvarova. S obzirom da su komponente vrata, vakuum WC i klima unificirani na svim vagonima i imaju jednog proizvođača, u daljnjem nastavku rada će se razmatrati pojava kvarova na *statičkim pretvaračima* zbog više različitih proizvođača, te dati prijedlog rješenja za poboljšanje pri čemu bi se smanjili troškovi održavanja i vremena zadržavanja vagona na popravcima.

4.3.4. PRIMJENA REGRESIJSKE ANALIZE

Regresijskom analizom pokušava se dovesti u korelaciju broj kvarova statičkih pretvarača različitih proizvođača koje smo dobili prethodnim metodama sa ukupnim prijeđenim kilometrima pojedinih serija klimatiziranih vagona u razdoblju od 1.1.2005. do 1.1.2007. godine. Tablica 4.9. prikazuje broj kvarova i prijeđene kilometre po pojedinim serijama vagona u navedenom vremenskom razdoblju.

Tablica 4.9. Prikaz prijeđenih kilometara i kvarova na pojedinim komponentama klimatiziranih putničkih vagona u razdoblju od 1.1.2005.- 1.1.2007. godine

Serije vagona	Proizvođač statičkog pretvarača	Prijeđeni kilometri (km)	Broj kvarova Na statičkim pretvaračima
Aeelmt, Beelmt, WReelmt	Siemens	3387626	45
Aeelt (10-70), Bee (20-70)	Končar VIS 50-1	12738696	133
Beelt (20-70)	Končar VIS 50	985864	15
Bee (29-70),	Pintch Bamag 15	2493110	3
ABeel, Beel (20-70)	MGVG 3000	1250280	13

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Tablica 4.10. Rezultati regresijske analize (ovisnost broja kvarova statičkih pretvarača o broju prijeđenih kilometara grupa pojedinih serija vagona)

Zavisna varijabla	Prijeđeni km.
N – broj serija putničkih vagona	5
R	0,972
R kvadrat	0,945
Prilagođeni R kvadrat	0,927
Standardna greška procjene	1.324.603,007

Utjecaj	koeficijent	Standardna greška	Standardni koeficijent (beta)	t-test	signifikantnost
konstanta	448.871,657	787.582,299	0,000	0,570	0,609
Broj kvarova	89.048,889	12.416,504	0,972	7,172	0,006

Utjecaj	koeficijent	95,0% razina povjerenja	
		donja	gornja
CONSTANT	448.871,657	-2.057.566,721	2.955.310,036
VAR_2	89.048,889	49.534,031	128.563,747

ANALIZA VARIJANCE

Izvor	suma kvadrata	srednja vrijednost kvadrata	F- test	signifikantnost
regresija	9,025E+013	9,025E+013	51,545	0,006
odstupanja	5,264E+012	1,755E+012		

Prema podacima iz tablice 4.10. F-test se računa kao:

$$F_{\text{empir.}} = \frac{R^2 / J}{(1 - R^2) / (N - J - 1)} = \frac{0,945 / 1}{(1 - 0,945^2) / (5 - 1 - 1)} = 51,545$$

gdje je: N – broj promatranih veličina (N=5),

J – broj objašnjivih varijabli (J=1, broj kvarova)

Prema F-tablici (95% razina povjerenja) za J=1 i N-J-1=3 slijedi vrijednost 10,10.

Tada je $F_{\text{empir.}} > F_{\text{tablice}}$ tj. $51,545 > 10,10$ i znakovita je ovisnost dvije varijable.

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Izračun međurezultata linearne regresije analitičkim putem dat je tablicom:

Promatrane veličine (k)	Y_k = broj kilometara	X_k = broj kvarova	XY	X^2
1.	3387626	45	152443170	2025
2.	12738696	133	1694246568	17689
3.	985864	15	14787960	225
4.	2493110	3	7479330	9
5.	1250280	13	16253640	169
	$\bar{Y} = 4171115,2$	$\bar{X} = 41,8$	$\Sigma = 1885210668$	$\Sigma = 20117$

Vrijednost procjena parametara linearne regresije (N=5)

a.) parametar b_1

$$b_1 = \frac{5 * (\sum_{k=1}^5 X_k Y_k) - (\sum_{k=1}^5 X_k) (\sum_{k=1}^5 Y_k)}{5 * (\sum_{k=1}^5 X_k^2) - (\sum_{k=1}^5 X_k)^2} = \frac{5 * 1885210668 - 209 * 20855576}{5 * 20117 - 2092} = 89049$$

b.) parametar b_0

$$b_0 = \bar{Y} - b_1 \bar{X} = 4171115,2 - 89049 * 41,8 = 448867$$

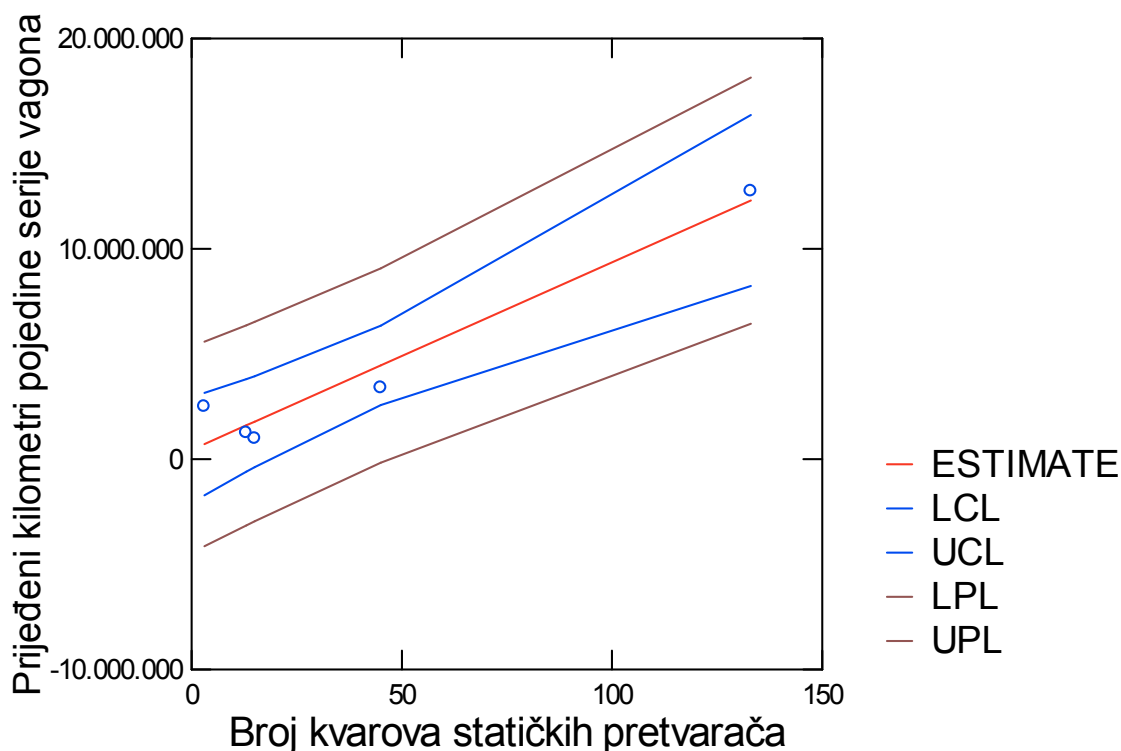
Jednadžba linearne regresije za ovaj slučaj:

$$Y_i = 448867 + 89049 * X_i$$

gdje je: Y_i – broj prijeđenih kilometara za pojedinu seriju vagona

X_i – broj kvarova na statičkim pretvaračima

Confidence Interval and Prediction Interval



Slika 38. Prikaz linearne regresije ovisnosti broja kvarova statičkih pretvarača i broja prijeđenih kilometara za pojedinu grupu serija vagona

Koeficijent korelacije $R = 0.97$ ili $R^2 = 94,5\%$ pokazuje da se 94,5 % rasutih podatak može objasniti i dovesti u povezanost, dok je 5,5 % podataka ostalo neobjašnjeno. Jednadžba linearne regresije pokazuje da se 1 kvar na klimatiziranim putničkim vagonima javlja na svakih 537921 km po vagonu. Pošto je između dva redovita pregleda (SP – srednji popravak) potrebno da vagon prijeđe 600000 km, vidljivo je da je opravdana ugradnja statičkih pretvarača, koji su zamijenili rotirajuće izvore. Pojava jednog kvara u prosjeku između dva velika redovita pregleda ukazuje da su uređaji dosta pouzdani, ali bez obzira na sve, neophodno je pri mjesečnim pregledima ispitati i kontrolirati rad pretvarača.

4.4. PRIJEDLOG POBOLJŠANJA TRAJNOSTI POJEDINIH AGREGATA PREMA PROIZVOĐAČU

Jedna od osnovnih veličina kojima se prati stanje kvarova je srednje vrijeme između kvarova (SVIK). Srednje vrijeme između kvarova je funkcija broja i trajanja kvarova. Cilj je da SVIK bude što je moguće veće tj. da kvarovi traju što kraće. Predloženo poboljšanje ima upravo to za cilj. Srednje vrijeme između kvarova, ukoliko se promatra N elemenata (u našem slučaju se to odnosi na količinu vagona serije, tj. broj statičkih pretvarača...) [12]:

$$SVIK = \frac{\sum_{j=1}^N \sum_{i=1}^{k_j} t_{ij}}{\sum_{j=1}^N k_j} \quad (22)$$

gdje je: k_j = ukupan broj kvarova j-tog elementa u tijeku vremena promatranja,

t_{ij} = vrijeme rada j-tog uređaja između i-1 i i-tog kvara,

N = ukupan broj promatranih elemenata

GRUPA A - vagoni Aeelmt, Beelmt i Wreelmt (ukupno 11 kom.) – statički pretvarač tvrtke Siemens (SR Njemačka),

GRUPA B – vagoni Aeelt, Bee, WRee (ukupno 40 kom.) – statički pretvarač tvrtke Končar VIS 50-1 (Hrvatska),

a.) Srednje vrijeme rada statičkih pretvarača između dva kvara

1.) Statički pretvarači tvrtke Siemens

Vagoni serija Aeelmt, Beelmt i WReelmt (pojedinačno) prijeđu u dvije godine prosječno 307966 km. S obzirom da im je izračunata prosječna brzina vožnje bila 180 km/h uzimajući u razmatranje sve relacije na kojima su vagoni prometovali u razdoblju od 1.1.2005. do 1.1.2007. godine, 11 vagona ovih serija godišnje (između dva kontrolna pregleda godišnjeg malog popravka MP) je u vožnji 18821 h, što iznosi po jednom vagonu 1711 h u dvije godine, a prema formuli (22) srednje izračunato vrijeme između kvarova iznosi:

$$SVIK_{Siemens} = 18821 \text{ h} / 45 \text{ kvara} = 418,225 \text{ h}$$

2.) Statički pretvarači tvrtke Končar

Vagoni serija Aeelt, Bee i WRee (pojedinačno) prijeđu u dvije godine prosječno 318467,4 km. S obzirom da im je izračunata prosječna brzina vožnje bila 140 km/h uzimajući u razmatranje sve relacije na kojima su vagoni prometovali u razdoblju od 1.1.2005. do 1.1.2007. godine, 40 vagona ovih serija godišnje (između dva kontrolna pregleda godišnjeg malog popravka MP) je u vožnji 90991 h, što iznosi po jednom vagonu 2274,8 h u dvije godine, a prema formuli (22) srednje izračunato vrijeme između kvarova iznosi:

$$SVIK_{Končar} = 90991 \text{ h} / 133 \text{ kvara} = 684,14 \text{ h}$$

Rezultati pokazuju da je vrijeme $SVIK_{Končar} > SVIK_{Siemens}$, te se zaključuje da je unutar dvije godine vožnje pojedine serije vagona statički pretvarač tvrtke Končar ima vremenski duži rad između kvarova, nego statički pretvarač tvrtke Siemens. Uspoređivanjem srednjeg vremena između kvarova za Siemensove statičke pretvarače i srednjeg vremena između kvarova za

statičke pretvarače domaćeg proizvođača Končar, eksperimentalno je utvrđeno da zamjena Končarevih statičkih pretvarača Siemensovim dovodi do povećanja postotnog SVIK-a za:

$$\eta = [1 - (SVIK_{Siemens} / SVIK_{Konč.})] * 100 \quad (23)$$

$$\eta = 0,389 = 38,9\%$$

Na osnovu rezultata istraživanja preporučuje se da se u budućnosti pri kupnji i izgradnji novih modernih vagona Hrvatskih željeznica ugrade uređaji opskrbe električnom energijom domaćeg proizvođača KONČAR.

b.) Srednje vrijeme popravaka vagona u radionici zbog kvarova statičkih pretvarača

Ako se promatra više elemenata koji se mogu više puta kvariti u tijeku ispitivanja, tada se srednje vrijeme popravka računa prema izrazu:

$$T_{sred} = \frac{\sum_{j=1}^N \sum_{i=1}^{k_j} t_{ij}}{\sum_{j=1}^N n_j} \quad (24)$$

gdje je: n_j = broj popravaka j-tog elementa

t_{ij} = vrijeme popravka (zadržavanja vagona u radionici)

N = ukupan broj elemenata koji se popravljaju

U radionici za popravke putničkih vagona u razdoblju od 1.1.2005. do 1.1.2007. najčešće su se popravljale komponente (elementi) prema tablici 4.11.:

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

Tablica 4.11. Kvarovi komponenti statičkog pretvarača Siemens EVA/B

Komponente (elementi)	Broj komponenti u kvaru	Vrijeme popravka (zamjene) (h)	Ukupno (h)
1. Isključen zbog sumnje u kvar (provjera)	25	4	100
2. Zamjena elektroničkog modula	2	8	16
3. Popravak elektroinstalacija	4	2	8
4. Zamjena SIBAS elektromotora	5	3	15
5. Nedostatak rashladnog sredstva	4	4	16
6. Zamjena utičnog modula punjača	1	2	2
7. Zamjena čopera	1	6	6
8. Zamjena prigušnice mrežnog filtera L3	2	40	80
9. Zamjena visokonaponskog osigurača	1	1,5	1,5
10. Popravak CPU jedinice	1	10	10
11. Zamjena osigurača F1	1	1,5	1,5
12. Zamjena ulaznog kondenzatora	1	5	5
13. Zamjena VN oscilatora	2	2,5	5
14. Zamjena punjača baterija	2	6	12
UKUPNO	52		278

Srednje vrijeme zadržavanja se dobije iz formule (24):

$$T_{\text{sred}} = \frac{278}{52} = 5,35 \text{ h / vagonu}$$

Uvrštavajući vrijednosti iz tablice 4.11., dobije se vrijeme zadržavanja 5,35 h po vagonu zbog statičkih pretvarača tvrtke Siemens EVA/B.

Tablica 4.12. Kvarovi komponenti statičkog pretvarača Končar VIS 50-1

Komponente (elementi)	Broj komponenti u kvaru	Vrijeme popravka (zamjene) (h)	Ukupno (h)
1. Isključen zbog sumnje u kvar (provjera)	23	4	92
2. Zamjena startera	6	8	48
3. Zamjena zaštitnog prekidača Q3	2	1,5	3
4. Zamjena osigurača F1, UNS	34	1,5	51
5. Zamjena VN sklopova, modula	68	5	340
6. Popravak dozemnih spojeva	4	2	8
7. Popravak spojeva (konektora)	5	2	10
8. Zamjena VN transformatora	1	10	10
9. Zamjena punjača baterija	1	8	8
10. Popravak CPU (elektronika, kartice)	14	0,5	7
11. Zamjena mjernog pretvarača A5 i A11	28	8	224
12. Zamjena čopera	1	6	6
13. Zamjena sklopnika pretvarača	1	3	3
14. Zamjena napajanja VN sklopova	2	1,5	3
UKUPNO	190		816

Srednje vrijeme zadržavanja se dobije iz iz formule (24):

$$T_{sred} = \frac{816}{190} = 4,28 \text{ h / vagonu}$$

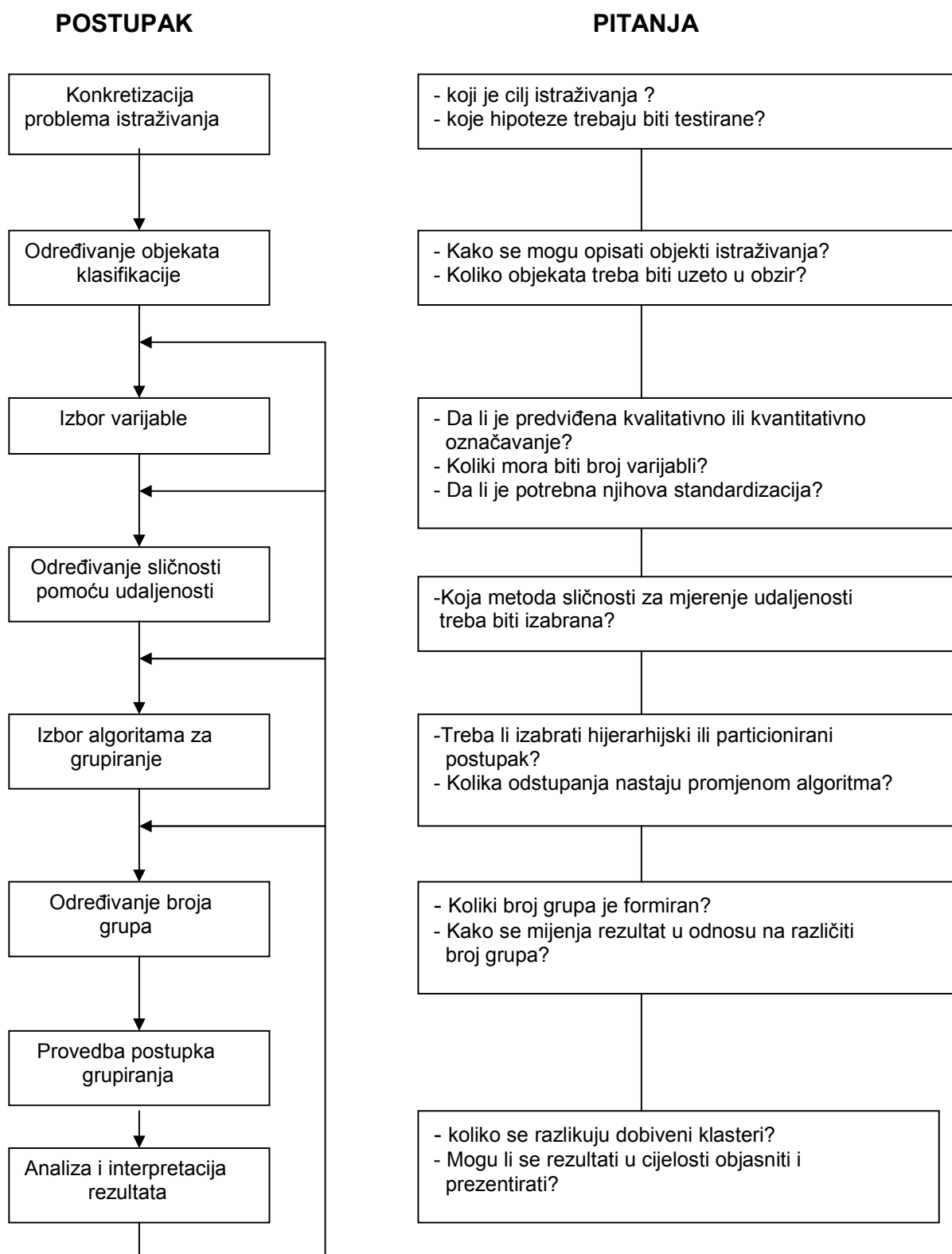
Uvrštavajući vrijednosti iz tablice 4.12, dobije se vrijeme zadržavanja 4,28 h po vagonu za statičke pretvarače tvrtke Končar VIS 50-1. Usporedbom dobivenih vremena popravaka statičkih pretvarača Siemens EVA/B i Končar VIS 50-1, dokazano je da se vagoni najmanje zadržavaju pri popravcima statičkih pretvarača tvrtke Končar, te je ovo još jedan razlog za njihovom ugradnjom u budućnosti, kao i zamjenom postojećih Siemensovih.

4.5. Odabir optimalne metode u rješavanju problema

Rješavajući problem kvarova na klimatiziranim putničkim vagonima metodama dubinskih analiza kao optimalna se pokazala klaster analiza zbog jednostavne mogućnosti prepoznavanja diskontinuiranih podgrupa. Za razliku od drugih primijenjenih metoda, prvi korak u klaster analizi je mjerenje „sličnosti“ između grupa promatranja (ili objekata) pomoću statističkih alata. Prije same uporabe klaster analize potrebno je znati koji su nam bili ciljevi, a u ovom radu je bilo neophodno istražiti kvarove sa najvećom frekvencijom te kvarove sa najmanjom frekvencijom, pri čemu smo Euklid metodom kao i metodom prosječne udaljenosti dobili iste podjele grupa kvarova sa najvećom učestalošću – statičke pretvarače, klima uređaje, vrata i uređaje za zatvaranje vrata, te vakuum WC i njegove instalacije i ostale kvarove. Klaster analiza je optimalna metoda u odnosu na ostale metode i zbog:

- mogućnosti izračuna udaljenosti između promatranih grupa ili objekata različitim metodama (Ward, Euklid, Minkovski...) na temelju čega se određuje sličnost između njih (što je manja udaljenost - sličnost je veća i obrnuto),
- jasnog grafičkog prikaza rezultata, te interpretacija istih (pomoću dendrograma),
- određivanja broja klastera,
- potvrde i objašnjenja formiranih grupa ili objekata,
- razmatranja ograničenja i različitosti među grupama ili objektima.

Matematički cilj klaster analize je potreba za minimiziranjem multivarijabilnosti između promatranih grupa ili objekata. Uporabom više metoda klasteriranja (u našem slučaju analize kvarova) možemo za razliku od ostalih metoda odabrati put po kojem možemo grupirati grupe (objekte) na temelju njihove međusobne sličnosti. Prema algoritmu na slici 39. dat je prikaz koraka uvjeta uspješnog provođenja klaster analize u odnosu na ostale metode dubinskih analiza. Lijeva strana dijagrama tijekom sadrži osam znanstvenih koraka važnih u procesu grupiranja podataka. Zaključno, navedena pitanja prema slici 39. su vodič za uspješno rješavanje centralnog problema klaster analizom, pri čemu postoji mogućnost više alternativa.



Slika 39. Koraci za rješavanje problema klaster analizom [8]

4. ISTRAŽIVANJA UČESTALOSTI KVAROVA NA SERIJAMA VAGONA

U odnosu na metode (Pareto analiza, multidimenzijско skaliranje) koje smo primijenili u istraživanjima kvarova, klaster analiza ima više mogućnosti mjerenja sličnosti i to:

a.) udaljenosti između dve promatrane veličine x_i i x_j :

- pomoću Euklidove udaljenosti $d_{i,j} = \left[\sum_{k=1}^k (x_{i,k} - x_{j,k})^2 \right]^{1/2}$
- korelacije $d_{i,j} = 1 - r^2$, gdje je r = koeficijent korelacije, te
- kuta između parova vektora: $\Theta = \cos^{-1} \left[\frac{x^T * y}{\|x\| * \|y\|} \right]$

b.) udaljenosti između klastera (pojedinačna veza, potpuna veza, prosječna veza, centroid...)

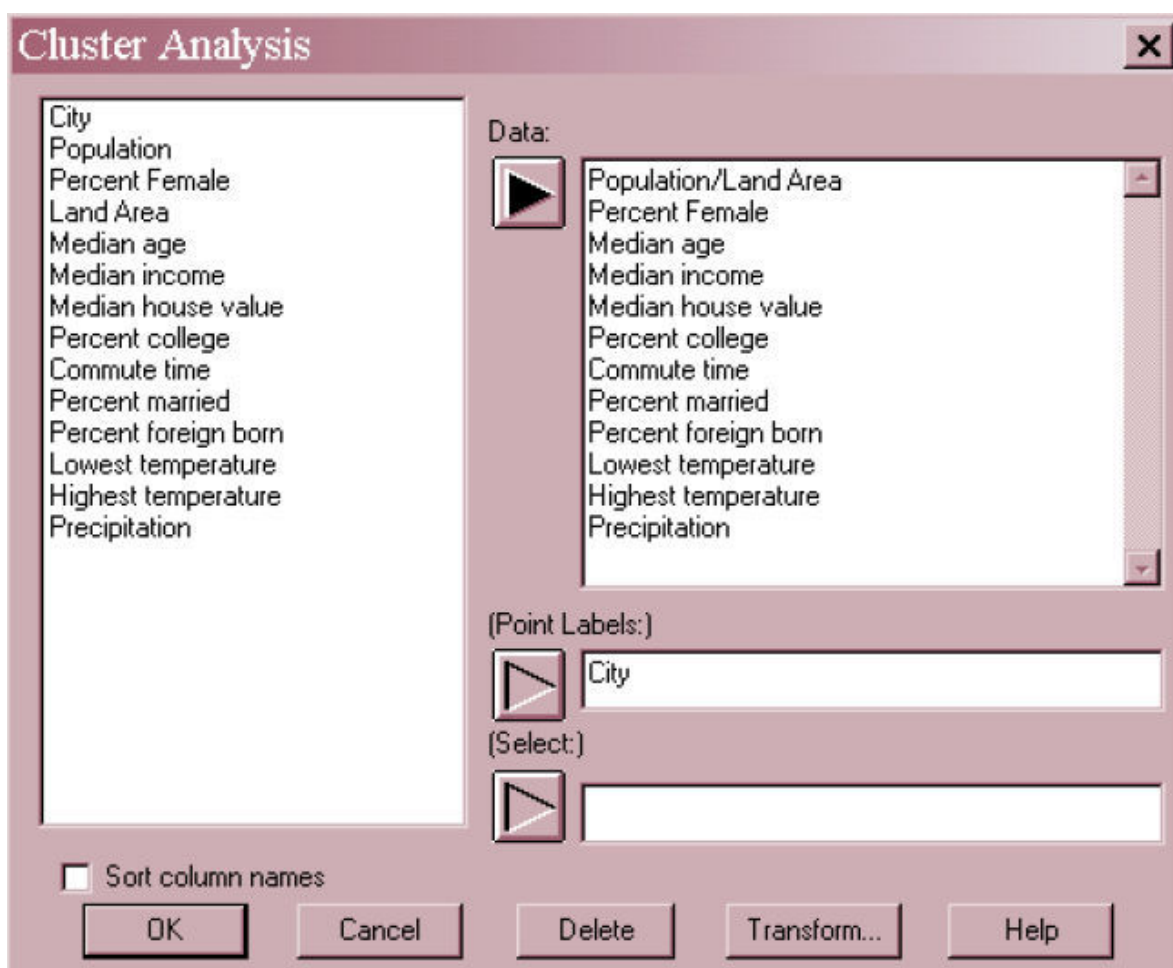
5. PRIMJENA STATISTIČKIH ALATA U ANALIZI KVAROVA

Model rješavanja problema klaster analizom uporabom STATGRAPHICS (verzije Centurion XV i 5.1.) i SYSTAT verzija 12

Programski paket Statgraphics ima mogućnost obrade podataka klaster analizom. Rezultati analize su ispisani na nekoliko načina, uključujući dendrogram, tablice i dijagrame.

Primjer: Na slici pod Data se unose imena stupaca koji sadrže ulazne podatke (slika 40.)

Korak 1.

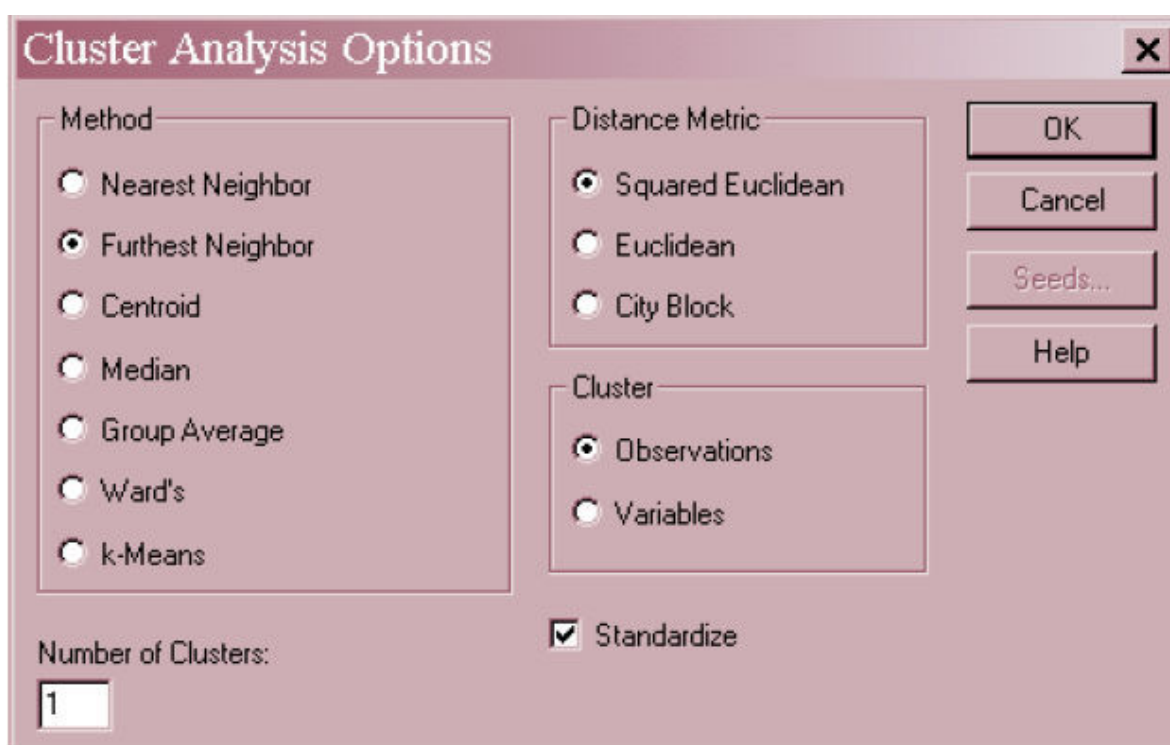


Slika 40. Polje podataka klaster analize u software-u Statgraphics verzija 5.1.

Data: ukoliko će se promatrani stupci klasterirati, imena od p ulaznih varijabli sadrži vrijednosti za n slučajeva, ili $(n \times n)$ matrica sadržava udaljenost između svakog slučaja. Ukoliko su varijable bile klasterirane, imena od p ulaznih varijabli sadrže vrijednosti za n slučajeva, ili $(p \times p)$ matrica sadrži udaljenost između svakog para varijabli.

Point labels: optimalne granice za svaki stupac u polju unosa podataka

Korak 2. OPCIJE ANALIZE



Slika 41. Polje odabira metoda klaster analize u Software-u Statgraphics 5.1.

METHOD: metoda korištena za stvaranje klastera

NUMBER OF CLUSTERS: krajnji broj željenih klastera

DISTANCE METRIC: metrika koja se koristi za mjerenje udaljenosti između klastera.

STANDARDIZE: ukoliko su potvrđene, varijable će biti standardizirane prije nego se izvede klasteriranje.

6. ZAKLJUČAK

Metode poslovnog izvješćivanja su suvremene metode koje su primijenjene u ovom radu, te su svaka posebno pomogle u rješavanju identifikacije kvarova pojedinih agregata i povezanosti grupa kvarova na klimatiziranim putničkim vagonima Hrvatskih željeznica. Svaka od primijenjenih metoda je kontrolirana Pareto metodom koja je potvrdila iste rezultate kao i metode poslovnog izvješćivanja. U radu je opisan niz suvremenih metoda poslovnog izvješćivanja kao što su klaster analiza, multidimenzijske analize, stablo odlučivanja i dr. te njihove najvažnije karakteristike. Na osnovu dobivenih rezultata istraživanja dokazano je da se poslovno izvješćivanje može u budućnosti primijeniti u sustavu tvrtke Održavanje vagona d.o.o. u analizama kvarova novih vagona, te pri preispitivanju odluka o ugradnji novih agregata (uređaja) koji imaju duži vijek trajnosti, tj. veće srednje vrijeme između kvarova (SVIK). Ciljevi i rezultati istraživanja su prihvatljivi, ali ovo nije kraj istraživanjima, već bi se uz pomoć navedenih metoda poslovnog izvješćivanja uz poslovnu politiku Hrvatskih željeznica moglo omogućiti još niz obrada podataka koji u ovom radu nisu obuhvaćeni.

Danas je, međutim problem kako pretvoriti informaciju u znanje, jer se tvrtke vode na osnovi znanja: o konkurenciji, kupcima, dobavljačima, procesima. Poslovno izvješćivanje je proizvodnja znanja koje je osnovica za donošenje poslovnih odluka. Uspjeh projekta ovisi ponajviše o korporacijskoj kulturi i spremnosti organizacije za prihvaćanje takvog sustava te o tri ključna faktora – jasnim ciljevima projekta, podršci vrhovnog menadžmenta i osiguranom odgovarajućem financiranju za opremu, programsku podršku i znanje. Tehnološki, sve je moguće napraviti, ali projekt je uspješan jedino ako se informacije iz tog sustava svakodnevno koriste u poslovanju. Kada govorimo o primjeni poslovnog izvješćivanja u Hrvatskim željeznicama, ali i u Hrvatskoj, moramo imati na umu nekoliko ograničenja. Prvo, investicije u informatičku tehnologiju se tretiraju kao trošak, a ne kao investicija, što bitno otežava mogućnosti za provođenjem ovih sustava. Ovdje dolazimo do apsurdne situacije da većina današnjih hrvatskih menadžera, koji su ujedno i informatički prosječno ili vrlo slabo obrazovani, stavljaju ovu tehnologiju u drugi plan poduzeća, iako bi upravo njima pomogla u poslovanju i kvalitetnijem odlučivanju. Pojmovi troškova i investicija se u informatičkoj tehnologiji moraju tretirati različito jer je koncept troška nužnost (primjerice, kupnja računala),

dok je investicija okrenuta i nužnosti razvoju (ulaganje u software, korisničku podršku, koncepte, modifikacija poslovnih procesa i slično). Znanje, na žalost, najčešće nije dovoljan argument da se dobije veliki projekt, budući da iza velikih konzultantskih kuća koje dovode strance koji odrađuju projekte stoji jaka marketinška i prodajna infrastruktura.

Ekonomski aspekti poslovnog izvješćivanja

Nekoliko financijskih pokazatelja je moguće koristiti na temelju kojih će se moći dobiti uvid u stanje i perspektive projekta prilikom pokretanja i provedbe projekata poslovnog izvješćivanja, kao i nakon njihova dovršetka. Primarni su: neto sadašnja vrijednost, unutarnja stopa povrata, povrat ulaganja i vrijeme povrata. Izračun povrata ulaganja je pokazatelj koji u sebi uključuje i neto sadašnju vrijednost i unutarnju stopu povrata i vrijeme povrata. Troškovi projekta mogu se podijeliti u tri glavne grupe [3]:

- troškovi hardware-a
- troškovi software-a
- troškovi rada čovjeka

Troškove uređaja i oprema korištene u informacijskom sustavu potrebne za realizaciju projekta se odnose na troškove Hardware-a. Troškovi software-a uključuju troškove aplikacijskog software-a nužnost u funkciji potpore projektu i pri njegovoj realizaciji. Osnovne komponente troškova tog software-a čine naknade za licence aplikacija poslovnog izvješćivanja i relacijskog sustava za upravljanje bazom podataka. Troškovi ljudskog rada obuhvaćaju naknade za rad zaposlenika i vanjskih suradnika angažiranih na projektu. Ukupni troškovi implementacije projekta poslovnog izvješćivanja mogu se podijeliti u dva segmenta koje tvore:

- početni (inicijalni) troškovi,
- ponavljajući (operativni) troškovi.

Početni troškovi su jednokratni i ovdje se ubrajaju troškovi hardware-a potrebnog za realizaciju projekta, te troškovi ljudskog rada nužnog za konfiguriranje i implementaciju

sustava. Ovom treba pridodati i troškove edukacije i uvježbavanje korisnika za rad sa sustavom poslovnog izvješćivanja. Ponavljajući (operativni) troškovi su oni koji se počinju javljati tijekom projekta poslovnog izvješćivanja, koji će se nastaviti, ali ponavljati i nakon što projekt bude dovršen. Poznavanje i razumijevanje ukupnih troškova implementacije projekta poslovnog izvješćivanja je imperativ ako tvrtka želi ostati u okvirima predviđenog projektnog budžeta. To će također pomoći i pri utvrđivanju faza projekta implementacije projekta poslovnog izvješćivanja [3]. Potencijalne koristi modela poslovnog izvješćivanja (definiranog u širem smislu) su:

- **povećanje prihoda** – pronalaženje ili otvaranje novih tržišta, učinkovitija sugestivna prodaja, brže prepoznavanje novih tržišnih mogućnosti, brže pojavljivanje na tržištu;
- **povećanje profita** – učinkovitije ciljno promotivno oglašavanje, rano prepoznavanje nepovoljnih tržišnih kretanja, utvrđivanje unutarnjih slabosti i nedostataka, učinkovito upravljanje promjenama;
- **unapređenje zadovoljstva klijenata** – bolje razumijevanje potreba, želja i preferencija klijenata, bolje povezivanje potreba klijenata s osobinama proizvoda i usluga, ponuda sličnih proizvoda kojima bi se mogle zadovoljiti potrebe klijenata, ponuda posebnih pogodnosti, brže rješavanje reklamacije klijenata, poboljšanje usluga nakon prodaje;
- **uštede** – bolja ulazna i izlazna logistika, smanjenje količine škarta i neprodane robe, manji povrat kupljene robe od klijenata;
- **povećanje tržišnog udjela** – ponuda robe koja bolje zadovoljava potrebama klijenata, povećanje broj klijenata koje je poduzeće preotelo konkurentima.

7. LITERATURA

- [1] R. Zelenika, *Metodologija i tehnologija izrade znanstvenog i stručnog djela*, Ekonomski fakultet u Rijeci 2000.
- [2] D. Taboršak, N. Šakić, *Predavanja iz kolegija Metode znanstveno – istraživačkog rada*, školska godina 2002. / 2003., Fakultet strojarstva i brodogradnje Zagreb.
- [3] Ž. Panian, G. Klepac, *Poslovna inteligencija*, Masmedia, Zagreb 2003.
- [4] P. Sikavica, B. Bebek, H. Skoko, D. Tipurić, *Poslovno odlučivanje*, II. izmijenjeno i dopunjeno izdanje, Informator, Zagreb 1999.
- [5] V. Srića, J. Muller, *Put k elektroničkom poslovanju*, Sinergija, Zagreb 99.
- [6] S. Kudyba, R. Hoptroff, *Dubinske analize and Business Intelligence: A Guide to Productivity*, Idea Group Publishing, Hershey, PA USA
- [7] Barry de Ville, *Decision Trees for Business Intelligence and Dubinske analize-Using SAS Enterprise Miner*, SAS Institute, Cary, NC, USA 2006.
- [8] K. Backhaus, B. Erichson, W. Plinke, R. Weiber, *Multivariate Analysemethoden - Eine anwendungsorientierte Einführung 11. Auflage*, Springer Berlin 2005.
- [9] Uputa za održavanje i rukovanje opremom putničkih vagona tipa Z1 serije, Aeelmt (10-90), Beelmt (20-90) i WReelmt (87-90), Hrvatske željeznice, Zagreb 1992.
- [10] B. Preprotić, *Poslovna inteligencija u gospodarenju tehničkim sustavima*, magistarski rad, Fakultet strojarstva i brodogradnje, Zagreb 2005.
- [11] <http://users.informatik.haw-hamburg.de/ubicomp/projekte/master2005/elvers/slides.pdf>
Microsoft Power Point presentation-Business Intelligence.pdf
- [12] N. Majdančić, *Strategije održavanja i informacijski sustavi održavanja*, SI.Brod 1999.
- [13] www.uni-mo.ba/ef/down/fris_fil/FRIS_03052006_9.pdf
Prezentacija teme–Financijsko računovodstveni informacijski sustav –FRIS, Tema- Poslovna inteligencija
- [14] www.eizg.hr/Adminlite/FCKeditor/PKIEP1C5-Bakaric.pdf
- [15] [web.efzg.hr/mvarga/Skladištepodataka/rudarenjepodataka\(2007\).pdf](http://web.efzg.hr/mvarga/Skladištepodataka/rudarenjepodataka(2007).pdf)
- [16] www.sve-mo.ba/ef/PDS_materijal/MIS.pdf
- [17] www.mathworks.com/access/helpdesk/help/toolbox/stats/f55357.html

- [18] en.wikipedia.org/wiki/Decision_tree
- [19] www.mindtools.com/dectree.html
- [20] www.cse.unsw.edu.au
- [21] dms.irh.hr/tutorial/tut_dtrees.php
- [22] www.corp.at/Download/CORP2006
- [23] www.doc.ic.ac.uk/report.html
- [24] www.clustan.com/what_is_klaster_analysis.html
- [25] www.pitt.edu/~csna/Milligan/readme.html
- [26] www.informs-cs.org/wsc97papers/1307.pdf
- [27] www.icsr.agh.edu.pl/~dzwinel/ramki/pdf-publ/EarthplanO.pdf
- [28] www.statsoft.com/textbook/glosn.html
- [29] Joao F.G. de Freitas- *Bayesian Methods for Neural Networks-Disertation, Trinity College, University of Cambridge* –Thesis.pdf
- [30] *Inženjerski priručnik IP4 –proizvodno strojarstvo, 3. svezak - Organizacija proizvodnje*, Školska knjiga, Zagreb 2002.
- [31] [http:// powerlab.fsb.hr/dloncar/ch_2.pdf](http://powerlab.fsb.hr/dloncar/ch_2.pdf) (neizrazita logika)
- [32] www.thearling.com/text/dmwhite/dmwhite.htm (An Introduction to data mining)
- [33] kreativa.efzgforum.net/index (Poslovna_inteligencija_Hrvatskoj_Vedran_Majstorović pdf)
- [34] whitepapers.techrepublic.com (30 ideas of using OLAP)
- [35] Marin Golub, Vrednovanje uporabe genetskih algoritama za aproksimaciju vremenskih nizova, magistarski rad, FER Zagreb
- [36] http://www.fthm.hr/e-student/files/christis/predavanje_opi2.pdf (Determinante modela)
- [37] <http://www.times.hr/modules/news/article.php?storyid=119> (SAP.....)
- [38] <http://www.trend.hr/clanak.aspx?BrojID=39&KatID=5&ClanakID=480> (mySAPerp)
- [39] www.skladistenje.com/download (Ljubetić - magistarski rad, Ekonomski fakultet u Zagrebu)
- [40] Ante Laušić, Izgradnja skladišta podataka, diplomski rad br. 2149, Fakultet elektrotehnike i računarstva, Zagreb travanj 2002.
- [41] www.contentmanager.eu.com

Životopis

Rođen sam 23.06.1973 u Zagrebu, gdje sam završio i osnovnu školu. Od 1988.-1992. pohađao sam X. Gimnaziju u Klaićevoj u Zagrebu. Godine 1992. sam upisao studij Strojarstva na Fakultetu strojarstva i brodogradnje u Zagrebu gdje sam i diplomirao 2001. Nakon završetka studija 2001. započinjem raditi u tvrtci Održavanje vagona (tvrtka kćer Hrvatskih željeznica) kao inženjer na održavanju putničkih vagona. Od zimskog semestra 2007. godine sudjelujem u izvođenju vježbi na Fakultetu strojarstva i brodogradnje iz kolegija Elementi konstrukcija kao honorarni asistent. Služim se engleskim i njemačkim jezikom. Nisam oženjen.

Curriculum Vitae

I was born on 23 June 1973 in Zagreb where i finished elementary school. From 1988 to 1992, i was a student at the Tenth High School in Klaićeva Street in Zagreb. In 1992, i enrolled for a course of studies at the Faculty of Mechanical Engineering and Naval Architecture in Zagreb. In 2001, i started working for a company engaged in maintenance of railway carriages which is an affiliated company of the Croatian Railways as an engineer responsible for maintenance of passenger carriages. Since the winter semester of 2007, i have participated in the carrying out of exercise courses in "Construction Elements" at the Faculty of Mechanical Engineering and Naval Architecture as freelance assistant professor. I speak English and German. I am not married.

Lebenslauf

Ich wurde am 23.06.1973 in Zagreb geboren, wo ich auch die Grundschule beendete. Von 1998 – 1992 besuchte ich das X. Gymnasium in der Klaićeva Strasse in Zagreb. 1992 habe ich das Maschinenbaustudium an der Fakultät für Maschinen- und Schiffbau in Zagreb immatrikuliert, wo ich 2001 mein Diplom erworben habe. Nach dem Abschluss des Studiums, habe ich 2001 als Ingenieur in der Firma Wartung von Eisenbahnwagen, Tochtergesellschaft der Kroatischen Eisenbahnen, angefangen zu arbeiten, wo ich mit der Wartung der Reisezugwagen beauftragt wurde. Seit dem Wintersemester 2007 bin ich als Honorarassistent an der Fakultät für Maschinen- und Schiffbau an den Übungen des Studienkollegs Konstruktionselemente tätig. Ich kann Englisch und Deutsch. Ich bin ledig.