

# Oblikovanje radijalnih mreža za regresiju

---

**Đakovac, Matej**

**Master's thesis / Diplomski rad**

**2019**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:200:822448>

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

*Download date / Datum preuzimanja:* **2024-11-23**

*Repository / Repozitorij:*

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU**

**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I**

**INFORMACIJSKIH TEHNOLOGIJA OSIJEK**

**Sveučilišni diplomski studij**

**Informacijske i podatkovne znanosti**

**OBLIKOVANJE RADIJALNIH MREŽA ZA REGRESIJU**

**Diplomski rad**

**Matej Đakovac**

**Osijek, 2019.**

## Sadržaj

<b>1. UVOD</b> .....	<b>1</b>
<b>2. RADIJALNE MREŽE ZA REGRESIJU</b> .....	<b>2</b>
2.1. Uvod u problem regresije .....	2
2.1.1. Primjene regresije .....	2
2.1.2 Linearna regresija.....	3
2.1.3 Nelinearna regresija .....	3
2.2. Radijalne mreže .....	4
2.2.1 Razlike u odnosu na višeslojni perceptron.....	6
2.3 Vrednovanje kvalitete mreže .....	6
<b>3. IZGRADNJA RADIJALNIH MREŽA ZA REGRESIJU</b> .....	<b>8</b>
3.1. Određivanje parametara radijalnih mreža.....	8
3.1.1 Određivanje središta radijalnih funkcija .....	8
3.1.2 Određivanje širina radijalnih funkcija .....	9
3.1.3 Određivanje težina veza između skrivenog i izlaznog sloja .....	11
3.2 Određivanje veličine radijalne mreže .....	12
<b>4. OSTVARENO PROGRAMSKO RJEŠENJE</b> .....	<b>14</b>
4.1 Način rada programskog rješenja .....	14
4.2 Prikaz i način uporabe programskog rješenja.....	15
<b>5. EKSPERIMENTALNA ANALIZA</b> .....	<b>19</b>
5.1. Postavke eksperimenta .....	20
5.2. Rezultati.....	21
<b>6. ZAKLJUČAK</b> .....	<b>36</b>

## Literatura

## Sažetak

## Životopis

## Prilozi

# 1. UVOD

Regresijska analiza je statistička metoda koja omogućuje utvrđivanje odnosa dviju ili više varijabli. Pronalazi široku uporabu u različitim znanostima gdje se koristi u svrhu predviđanja budućih vrijednosti, što može služiti, primjerice, za optimizaciju proizvodnje, predviđanja ponašanja tržišta i financija, vremenskih prilika, demografskih promjena i mnoge druge primjene. To je tehnika modeliranja koja istražuje na koji način varijable utječu na predmet interesa. Pristupi rješavanju problema regresije (izgradnje regresijskih modela) koji su obrađeni u ovom radu su linearna i nelinearna regresija. Odabir regresijskog modela ovisi o podacima na kojima se vrši regresijska analiza, ali vrlo su važne i odabrane varijable. Za složene probleme regresije moguće je korištenje neuronskih mreža kao regresijskog modela. Ovaj rad se bavi korištenjem radijalnih mreža kao regresijskog modela. Radijalne mreže su vrsta neuronske mreže i fleksibilan model koji se prilagođava obliku ulaznih podataka, a u slučaju nedovoljne prilagođenosti se može “pojačati”, to jest povećati kompleksnost i podići mogućnosti predviđanja. Radijalnim mrežama je potrebno odrediti veličinu, a uz to imaju određen broj parametara koji se određuju putem treniranja, i to je ono što predstavlja izazov kod njihovog korištenja u regresijskoj analizi. Određivanje parametara se vrši korištenjem nadziranih i nenadziranih metoda učenja, te između ostalog često uključuje korištenje algoritma za grupiranje. Grupiranje je nenadzirana metoda učenja koja služi za optimizaciju performansi mreže s ciljem povećanja kvalitete predviđanja.

Poglavlje 2 diplomskog rada obuhvaća detaljnije objašnjenje problema regresije te detaljan opis radijalnih mreža, njihove arhitekture i primjene. Uz to, uspoređene su s višeslojnim perceptronom, drugačijim oblikom neuronskih mreža te su dani načini vrednovanja uspješnosti predviđanja. Poglavlje 3 se bavi postupkom modeliranja radijalnih mreža, odnosno određivanjem njihovih parametara i veličine. U poglavlju 4 dan je opis ostvarenog programskog rješenja koje omogućuje korištenje radijalne mreže na korisnički određenim skupovima podataka. Opisan je način rada, te je prikazan način uporabe. Poglavlje 5 predstavlja eksperimentalnu analizu. Eksperimentalna analiza se odnosi na opis korištenih stvarnih skupova podataka te dobivene rezultate. Objašnjene su postavke eksperimenta, a dobiveni rezultati su prikazani i potkrijepljeni komentarima i objašnjenjima.

## 2. RADIJALNE MREŽE ZA REGRESIJU

Radijalne mreže (engl. *radial basis function networks, RBFNs*) su općenito troslojne mreže s jednim skrivenim slojem u kojem se koriste radijalne funkcije kao aktivacijske funkcije. Primjenjuju se u regresijskoj analizi, što u biti predstavlja traženje funkcije koja najbolje aproksimira dani skup točaka. Ostale primjene uključuju klasifikaciju i predviđanje vremenskih serija.

### 2.1. Uvod u problem regresije

Regresijska analiza [1], [2] i [3] je jedna od najčešće korištenih statističkih metoda sa širokom primjenom u različitim područjima znanosti. To je postupak analize međuovisnosti zavisne varijable (najčešće označene sa  $y$ ) i jedne ili više ulaznih, nezavisnih varijabli (najčešće označene sa  $x$ ), čija je svrha [1]:

- utvrđivanje odnosa između zavisne varijable  $y$  i nezavisnih varijabli  $x_1, x_2, \dots, x_n$
- predviđanje vrijednosti zavisne varijable na temelju vrijednosti nezavisnih varijabli
- analiza nezavisnih varijabli s ciljem identifikacije koje od njih imaju najveći utjecaj na vrijednost zavisne varijable kako bi se učinkovitije i preciznije odredila međuovisnost ulaznih i izlaznih varijabli

Proces regresijske analize se sastoji od slijedećih koraka [2]:

1. analiza problema
2. odabir varijabli
3. prikupljanje podataka
4. odabir regresijskog modela
5. odabir metode procjene regresijskih koeficijenata
6. treniranje modela
7. ocjena uspješnosti modela
8. korištenje modela za rješavanje danog problema

#### 2.1.1. Primjene regresije

Regresijska analiza se primjenjuje u ekonomiji, financijama, poslovanju, pravu, meteorologiji, medicini, biologiji, kemiji, fizici, edukaciji, povijesti, sociologiji itd. Chatterjee i Hadi u [2] navode i neke konkretne primjere korištenja regresije:

- analiza podataka o proizvodnji mlijeka, što uključuje razvoj pogodnog modela za predviđanje količine proizvedenog mlijeka u ovisnosti o trenutnoj i prijašnjoj proizvodnji, postotku masti i bjelančevina u mlijeku i slično
- analiza podataka o unutarnjim migracijama, što se odnosi na predviđanje migracija unutar države promatranjem čimbenika kao što su vremenski uvjeti, stope kriminala i nezaposlenosti, porezi itd.
- procjena starosti povijesnih pronalazaka na temelju povijesnih činjenica (primjerice, procjena starosti pronađene lubanje na temelju širine, visine, duljine i dimenzija nosne šupljine)

### 2.1.2 Linearna regresija

Prema [1], pretpostavka linearne regresije je da je zavisna varijabla linearna funkcija regresijskih koeficijenata. Općeniti oblik linearnog regresijskog modela je dan s [1]:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \varepsilon, \quad (2-1)$$

gdje su  $x_1, x_2, \dots, x_i$  nezavisne varijable,  $y$  je zavisna varijabla,  $\beta_0, \beta_1, \beta_2, \dots, \beta_i$  su regresijski koeficijenti, a  $\varepsilon$  pogreška s pretpostavkom  $E(\varepsilon) = 0$  te  $\text{Var}(\varepsilon) = \sigma^2$ , što znači da  $\varepsilon$  ima normalnu razdiobu. Ovakav model se naziva višestruka linearna regresija, iz razloga što ima više od jedne nezavisne varijable. Nasuprot tome, ako postoji jedna nezavisna varijabla onda se govori o jednostavnoj linearnoj regresiji. Prema [1], regresijski koeficijenti u linearnoj regresiji se najčešće procjenjuju korištenjem metode najmanjih kvadrata, koja je detaljnije objašnjena u poglavlju 3.

### 2.1.3 Nelinearna regresija

Prema [3], pretpostavka nelinearne regresije je da je odnos između zavisne i nezavisnih varijabli nelinearan s obzirom na regresijske koeficijente. Prema tome, nelinearan regresijski model je svaki regresijski model koji nije u obliku danom formulom (2-1). Xin i Xiao u [1] navode da su nelinearni regresijski modeli složeniji od linearnih u smislu procjene regresijskih koeficijenata, odabira i vrednovanja modela, odabira varijabli i slično. Jedna od modernijih metoda procjene regresijskih koeficijenata za nelinearnu regresiju je primjena radijalnih mreža, za što je detaljno objašnjenje dano u nastavku. Neke od ostalih metoda opisanih u [3] su metoda nelinearnih najmanjih kvadrata, linearna aproksimacija, Gauss-Newtonov algoritam, korištenje prijašnjeg znanja za postavljenje početnih vrijednosti i slično.

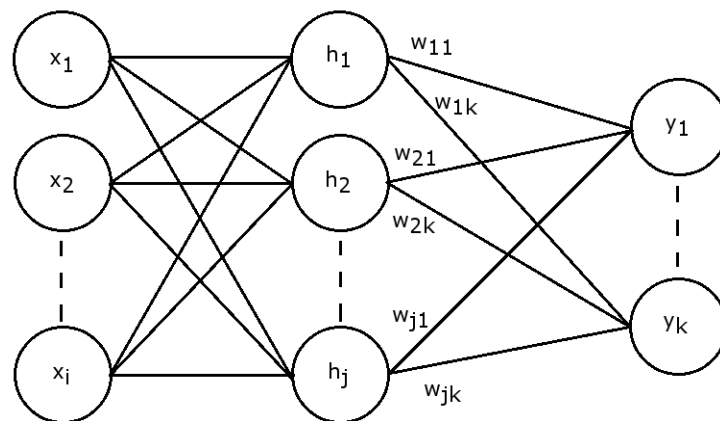
Jedan od primjera nelinearnog regresijskog modela je model rasta u biologiji [1]:

$$y = \frac{\alpha}{1+e^{\beta t}} + \varepsilon, \quad (2-2)$$

gdje je  $y$  mjera rasta određenog organizma i funkcija vremena  $t$ ,  $\alpha$  i  $\beta$  su regresijski koeficijenti, a  $\varepsilon$  pogreška.

## 2.2. Radijalne mreže

Broomhead i Lowe u [4] opisuju arhitekturu radijalnih mreža kao troslojnu, a sastoji se od ulaznog, skrivenog i izlaznog sloja, kako je prikazano na slici 2.1. Ulazni sloj je skup od  $i$  čvorova koji prima komponente  $i$ -dimenzionalnog ulaznog vektora  $\mathbf{x}$ . Svaki od čvorova u ulaznom sloju je povezan sa svakim čvorom skrivenog sloja. Skriveni sloj čini skup od  $j$  čvorova, po jedan za svaku radijalnu funkciju. Svaki od čvorova skrivenog sloja je spojen sa svim čvorovima u izlaznom sloju. Broj čvorova u izlaznom sloju odgovara broju komponenta  $k$  izlaznog vektora  $\mathbf{y}$ . Ulazna vrijednost u svaki čvor izlaznog sloja je težinska suma svih izlaza iz čvorova skrivenog sloja, gdje su težine veza od skrivenog čvora  $j$  do izlaznog čvora  $k$  označene sa  $w_{jk}$ .



Sl. 2.1. Arhitektura radijalne mreže

Svrha ulaznog sloja je da prosljeđuje ulazne podatke (primjere iz skupa podataka) skrivenom sloju, gdje je svaki od atributa ulaznog podatka komponenta ulaznog vektora. U [4] se također navodi da su radijalne mreže takve mreže s propagacijom prema naprijed (engl. *feedforward*). Na ovaj način, svakom od čvorova u skrivenom sloju je prikazan cijeli ulazni vektor.

Svaki od čvorova u skrivenom sloju kao aktivacijsku funkciju koristi radijalnu funkciju  $\varphi$ . Prema tome, izlaz  $h_j$  iz skrivenog čvora  $j$  je dan sa [5]:

$$h_j(x) = \varphi(\|x - \mu\|_2), \quad (2-3)$$

gdje je  $x$  ulazna vrijednost,  $\varphi$  je radijalna funkcija,  $\mu$  je središte radijalne funkcije, a  $\|x - \mu\|_2$  označuje euklidsku udaljenost između ulazne vrijednosti  $x$  i središta  $\mu$  radijalne funkcije. Radijalne funkcije (engl. *radial basis function*, *RBF*) su radijalno simetrične funkcije čija vrijednost ovisi o udaljenosti ulazne vrijednosti od središta  $\mu$ . Kao takvim, vrijednost im je najviša u središtu i opada prema nuli udaljavanjem od središta. U [5] se navodi da je najčešća mjera udaljenosti euklidska udaljenost. Prema tome, izlaz čvora skrivenog sloja je skalarna vrijednost i mjera udaljenosti  $x$  od središta radijalne funkcije. Dodatno, neke od radijalnih funkcija koriste parametar  $\sigma$ , koji karakterizira širinu receptivnog polja. Neke od funkcija koje se koriste u radijalnim mrežama su [5]:

- kubna funkcija:

$$\varphi(\|x - \mu\|_2) = \|x - \mu\|_2^3 \quad (2-4)$$

- *thin-plate-spline* funkcija:

$$\varphi(\|x - \mu\|_2) = \|x - \mu\|_2^3 \ln \|x - \mu\|_2 \quad (2-5)$$

- multikvadratna funkcija:

$$\varphi(\|x - \mu\|_2, \sigma) = \sqrt{(\|x - \mu\|_2^2 + \sigma^2)} \quad (2-6)$$

- inverzna multikvadratna funkcija:

$$\varphi(\|x - \mu\|_2, \sigma) = \frac{1}{\sqrt{\|x - \mu\|_2^2 + \sigma^2}} \quad (2-7)$$

- Gaussova funkcija:

$$\varphi(\|x - \mu\|_2, \sigma) = e^{-\frac{\|x - \mu\|_2^2}{2\sigma^2}} \quad (2-8)$$

U ovom radu se kao aktivacijsku funkcija čvorova skrivenog sloja koristi Gaussova funkcija, dana formulom (2-8) koja je prema [5] jedna od najčešće korištenih aktivacijskih funkcija u radijalnim mrežama.



Svaki od čvorova izlaznog sloja koristi linearnu aktivacijsku funkciju. Sukladno tome, izlaz radijalne mreže je linearna kombinacija radijalnih funkcija [5]:

$$y_k = \sum_j w_{jk} h_j, \quad (2-9)$$

gdje je  $y_k$  izlaz  $k$ -tog čvora u izlaznom sloju,  $w_{jk}$  je težina veze između  $j$ -tog čvora skrivenog sloja i  $k$ -tog čvora izlaznog sloja, a  $h_j$  je izlaz  $j$ -tog čvora skrivenog sloja.

### 2.2.1 Razlike u odnosu na višeslojni perceptron

Postoje određene sličnosti između radijalnih mreža i višeslojnih perceptrona (engl. *multilayer perceptron*, *MLP*). Prema [5], oboje su vrste umjetnih neuronskih mreža (engl. *artificial neural networks*, *ANN*) s propagacijom prema naprijed korištene za aproksimaciju proizvoljnih nelinearnih višedimenzionalnih funkcija. Međutim, postoje bitne razlike između dva pristupa, a neke od njih su [6]:

- izlazi čvorova skrivenih slojeva u MLP su težinska suma ulaza transformirana nekom aktivacijskom funkcijom, dok RBFN koriste udaljenost ulaznog vektora od središta radijalne funkcije transformiranu korištenjem radijalne funkcije.
- kod MLP uobičajeno više skrivenih čvorova doprinose određivanju izlazne vrijednosti, dok kod RBFN samo nekoliko čvorova u skrivenom sloju daju značajan odziv.
- MLP često ima više skrivenih slojeva sa složenim načinima spajanja, to jest mreža nije uvijek u potpunosti spojena. RBFN obično ima jednostavnu arhitekturu s jednim skrivenim slojem, te su svi slojevi potpuno spojeni.
- kod MLP se mogu koristiti više različitih aktivacijskih funkcija, dok su kod RBFN iste za sve čvorove skrivenog sloja.
- svi parametri MLP se obično određuju u isto vrijeme u sklopu jednostruke strategije nadziranog učenja. RBFN se trenira u dvije faze: prvo se određuju parametri radijalnih funkcija postupcima nenadziranog učenja koristeći samo ulazne podatke, te nakon toga težine veza koje povezuju skriveni i izlazni sloj koristeći metode nadziranog učenja.

### 2.3 Vrednovanje kvalitete mreže

Kako bi se ocijenila uspješnost modela, potrebno je izmjeriti njegovu prilagođenost podacima za treniranje. Prema [7], potrebno je razdvojiti podatke na dva podskupa, od kojih se prvi koristi za treniranje, a drugi za validaciju. Prvi podskup je veći i koristi se za određivanje parametara

radijalne mreže, a na preostalim podacima u drugom podskupu se model ocjenjuje. Jednostavna mjera je zbroj kvadratne pogreške (engl. *Sum Squared Error, SSE*) [5]:

$$SSE = \sum_v \sum_k (t_{kv} - y_{kv})^2, \quad (2-10)$$

gdje je  $v$  ukupan broj točaka u podskupu za validaciju,  $k$  je broj izlaznih čvorova,  $t_{kv}$  i  $y_{kv}$  predstavljaju ciljanu i izlaznu vrijednost na čvoru  $k$  za točku  $p$  u podskupu podataka za validaciju. Mjera SSE je zbroj udaljenosti ciljanih i dobivenih vrijednosti, a udaljenost se kvadrira kako bi sve vrijednosti bile pozitivne (jer se mogu poništavati u slučaju različitih predznaka) te kako bi se kažnjavale veće pogreške. Međutim, prema [5], SSE nije pogodna mjera kad se uspoređuje uspješnost na skupovima različitih veličina. U takvim slučajevima moguće je koristiti srednju kvadratnu pogrešku (engl. *Mean Squared Error, MSE*) danu s [5]:

$$MSE = \frac{\sum_v \sum_k (t_{kv} - y_{kv})^2}{vk}, \quad (2-11)$$

koja koristi iste varijable kao SSE. U ovom slučaju izraz za SSE (u brojniku) se normalizira umnoškom broja točaka u podskupu za validaciju i broja izlaznih čvorova mreže.

Orr u [7] navodi da je podskupove za treniranje i validaciju potrebno odrediti više puta, izvršiti treniranje i validaciju za svaku podjelu i uzeti prosjek dobivenih pogrešaka. Ovo se može postići korištenjem algoritma  $k$ -rezova (engl. *k-fold*), čija je implementacija dana s [8], koji funkcionira na slijedeći način: skup podataka se dijeli na  $d$  dijelova (rezova) jednake veličine koji sadrže nasumične točke. Slijednim postupkom se redom uzima po  $d-1$  rezova za treniranje i jedan za validaciju, te se pri validaciji zabilježi MSE. Nakon  $d$  iteracija se uzima srednja vrijednost svih zabilježenih MSE. Prema [8], preporučena vrijednost za  $d$  je 10.

### 3. IZGRADNJA RADIJALNIH MREŽA ZA REGRESIJU

Izgradnja radijalnih mreža za regresiju je procedura modeliranja mreže određivanjem pojedinih parametara. Neuronske mreže, uključujući radijalne mreže su neparametarski modeli, što znači da njihovi parametri (težine veza i slično) nemaju posebno značenje u problemima na koje se primjenjuju. Radijalne mreže imaju nešto više parametara za određivanje od tipičnih neuronskih mreža. Uz to, određivanje parametara se vrši korištenjem i nadziranih i nenadziranih metoda učenja. Nenadziranim metodama učenja se određuju parametri radijalnih funkcija, a nadziranim se određuju težine veza između skrivenog i izlaznog sloja.

#### 3.1. Određivanje parametara radijalnih mreža

Uz odabir radijalne funkcije koja će se koristiti u mreži za regresiju i određivanje broja tih funkcija, potrebno je odrediti slijedeće parametre (za Gaussovu radijalnu funkciju):

- Standardna devijacija –  $\sigma$  – određuje širinu radijalne funkcije
- Središte radijalne funkcije –  $\mu$  – određuje položaj radijalne funkcije

Osim parametara radijalnih funkcija, potrebno je odrediti i vrijednosti težina veza između skrivenog i izlaznog sloja.

##### 3.1.1 Određivanje središta radijalnih funkcija

Nakon odabira radijalne funkcije i broja čvorova u skrivenom sloju, prvi korak procesa treniranja radijalne mreže je određivanje središta radijalnih funkcija. Bishop u [6] navodi proceduru koja koristi sve točke iz skupa za treniranje kao središta radijalnih funkcija. Međutim, ovakav pristup za velike skupove podataka ima i velik broj čvorova u skrivenom sloju. Prema tome, treniranje takve mreže koristi veliku količinu računalnih resursa i vremena, a uz to je moguća pojava pretjerane prilagođenosti podacima za treniranje, što je objašnjeno u poglavlju 3.2. Iz ovih razloga, u [6] se predlaže drugačiji pristup, koji predstavlja postavljanje središta radijalnih funkcija u samo nekoliko točaka nasumično odabranih iz skupa za treniranje. Ova procedura također nije optimalna što se tiče procjene gustoće, te također može rezultirati korištenjem nepotrebno visokog broja radijalnih funkcija, što je pokazano u poglavlju 5.

U skladu sa spomenutim problemima, u [5] i [6] se navodi procedura određivanja središta radijalnih funkcija korištenjem algoritma za grupiranje (engl. *clustering*) *k-means*. Sve točke iz ulaznog skupa podataka potrebno je grupirati u  $j$  grupa, gdje je  $j$  broj čvorova u skrivenom sloju mreže. Prema [9], za točke koje treba grupirati algoritam *k-means* pronalazi takvu particiju gdje

je SSE (zbroy kvadratne pogreške) između središta grupa i točaka koje pripadaju toj grupi minimalan. Središte grupe je aritmetička sredina svih točaka koje joj pripadaju. Algoritam se izvršava u koracima:

1. Odabirati  $j$  početnih središta. Ponavljati korake 2 i 3 sve dok ima promjena u pridruživanju točaka grupama ili dok nije zadovoljen predefrirani kriterij zaustavljanja (primjerice maksimalan broj iteracija).
2. Generirati novu particiju pridruživanjem svake točke najbližem centru grupe.
3. Izračunati nova središta grupa.

Kvadratna pogreška uvijek opada s povećanjem broja grupa  $j$  (ima vrijednost 0 ako je  $j$  jednak broju točaka koje treba grupirati), pa ju je stoga moguće minimizirati samo ako se broj grupa ne mijenja tokom izvođenja algoritma.

Algoritam *k-means* zahtjeva tri korisnički specificirana parametra: broj grupa  $j$ , početna središta grupa i mjeru udaljenosti. Način odabira broja grupa  $j$  u slučaju radijalnih mreža je opisan u poglavlju 3.2. Postavljanje početnih središta grupa se može izvršiti na tri načina [9]:

1. nasumičan odabir
2. *k-means++*
3. korisnički definiran

Ako se koristi način 1, gdje su početna središta odabrana nasumično, prema [9], različiti odabiri mogu rezultirati različitim završnim particijama jer *k-means* konvergira u lokalne minimume. Jedan način obilaska ovog problema je pokretanje algoritma za dani  $j$  više puta s različitim odabirima početnih središta i odabiranje iteracije s najmanjom kvadratnom pogreškom. Korištenje *k-means++* načina za određivanje početnih središta grupa se radi na način da se prvo središte odredi nasumično uniformnom razdiobom, a nakon toga sva ostala središta se također određuju nasumično, gdje udaljenije točke od postojećih odabranih središta imaju veću šansu odabiranja. Što se tiče mjere udaljenosti, u [9] se navodi da se najčešće koristi euklidska udaljenost između točaka i središta grupa.

### 3.1.2 Određivanje širina radijalnih funkcija

Drugi korak procesa treniranja se odnosi na određivanje širina radijalnih funkcija. Jednostavan pristup je da se širine fiksiraju, to jest da svaka od radijalnih funkcija ima istu širinu  $\sigma$  određenu s [10]:

$$\sigma = \frac{d_{max}}{\sqrt{2j}}, \quad (3-2)$$

gdje je  $d_{max}$  najveća udaljenost između bilo kojeg para središta radijalnih funkcija, a  $j$  broj radijalnih funkcija i čvorova u skrivenom sloju mreže. Prema [11], ovakvo rješenje za problem određivanja širine  $\sigma$  je optimalno samo u slučaju ravnomjerne prostorne raspodjele ulaznih podataka, što znači da su i središta radijalnih funkcija ravnomjerno raspoređena. Iz tog razloga ova metoda je nepogodna za korištenje u praksi zato što je velika većina podataka u stvarnom svijetu neravnomjerno raspoređena.

Ako su udaljenosti između susjednih središta radijalnih funkcija različite, bolji je pristup dodijeliti svakoj od njih vlastitu širinu. K tome, drugi način određivanja širina se bavi određivanjem parametra  $\sigma$  za svaku radijalnu funkciju posebno. Jedan od načina rješavanja ovog problema je korištenjem algoritma p-najbližih susjeda [11]:

$$\sigma_j = \frac{1}{p} \left( \sum_p \|\mu_j - \mu_p\|^2 \right)^{\frac{1}{2}}, \quad (3-3)$$

gdje je  $\sigma_j$  širina radijalne funkcije  $j$ ,  $p$  je broj susjeda koji se uzima u obzir,  $\mu_j$  je središte radijalne funkcije  $j$ , a  $\mu_p$  je središte  $p$ -tog najbližeg susjeda. Moody i Darken u [11] preporučaju korištenje dva najbliža susjeda.

Korištenjem jednog najbližeg susjeda se pojednostavljuje izračun, a samim time i povećava brzina izvođenja. Prema [12], u slučaju da se u formuli (3-3) koristi jedan najbliži susjed moguće je smanjiti pogrešku tako da se dobivena udaljenost pomnoži s faktorom skaliranja širine  $q$ :

$$\sigma_j = q(\|\mu_j - \mu_p\|) \quad (3-4)$$

Benoudjit et al. u [13] navode da je ubacivanjem faktora skaliranja širine proces generalizacije učinkovitiji jer se omogućava optimalno preklapanje radijalnih funkcija. Također, u [13] se navodi da faktor  $q$  ovisi o funkciji koju se aproksimira, dimenzijama ulaznog skupa podataka te razdiobi podataka. Općeniti pristup se ostvaruje tako da se zada skup faktora skaliranja širina  $Q$ . Za svaki  $q_l \in Q$  je potrebno odrediti pogrešku (odabrati način vrednovanja, za ovaj primjer je to MSE), gdje optimalni faktor  $q_{opt}$  odgovara najmanjoj pogrešci [13]:

$$\forall l, MSE(q_{opt}) \leq MSE(q_l), \quad (3-5)$$

Kad se pojavi više od jednog minimuma, u [13] se preporuča odabrati onaj s najmanjom vrijednosti  $q$ .

### 3.1.3 Određivanje težina veza između skrivenog i izlaznog sloja

Težine veza između skrivenog i izlaznog sloja se određuju metodama nadziranog učenja. U radu su objašnjene dvije takve metode: gradijentni spust (engl. *gradient descent*) i metoda najmanjih kvadrata. Metodom gradijentnog spusta težine se računaju na slijedeći način [14]:

1. Odabrati nasumične početne vrijednosti za težine veza između skrivenog i izlaznog sloja.
2. Izračunati nagib (engl. *gradient*) funkcije pogreške (engl. *error function*) u točki određenoj vrijednostima težina.
3. Promijeniti vrijednosti težina korištenjem dobivenog nagiba.
4. Ponavljati drugi i treći korak dok se ne pronađe lokalni minimum funkcije pogreške ili dok se ne zadovolji neki drugi kriterij zaustavljanja (primjerice maksimalan broj iteracija).

Za funkciju pogreške se koristi neka od mjera prilagođenosti mreže podacima za treniranje (MSE ili SSE). Nagib funkcije pogreške se računa korištenjem parcijalne derivacije po svakoj od težina [14]:

$$\nabla_{\mathbf{w}} e = \frac{\partial e}{\partial \mathbf{w}} = \left( \frac{\partial e}{\partial w_{11}}, \frac{\partial e}{\partial w_{12}}, \dots, \frac{\partial e}{\partial w_{jk}} \right), \quad (3-6)$$

gdje je  $e$  funkcija pogreške, a  $w_{jk}$  težina veze između  $j$ -tog čvora skrivenog sloja i  $k$ -tog čvora izlaznog sloja. Promjena težine  $w_{jk}$  se dobiva množenjem odgovarajućeg nagiba funkcije pogreške i stope učenja (engl. *learning rate*) te dodaje na prijašnju vrijednost  $w_{jk}$ . Stopa učenja je faktor koji odlučuje za koliko se mijenjaju težine u svakoj iteraciji algoritma. Ako je premali, potrebno je previše iteracija kako bi se došlo do lokalnog minimuma funkcije pogreške. Ako je prevelik i promjene na težinama su prevelike i nikad ne poprima vrijednost za koju funkcija pogreške postiže lokalni minimum. Prema [14], preporučena vrijednost za stopu učenja je 0.2.

Tri su moguća načina izvođenja algoritma gradijentnog spusta [14]: korištenjem skupina (engl. *batch*), mini-skupina i stohastički. Korištenjem skupina se cijeli skup podataka za treniranje uzima u obzir prije promjene težina. Stohastički pristup mijenja vrijednosti težina za svaki primjerak iz skupa podataka. Pristup korištenjem mini-skupina kombinira dva nabrojana pristupa na način da podijeli skup podataka za treniranje na podskupove te mijenja vrijednosti težina za svaki podskup.

Prema formuli (2-9) izlaz radijalne mreže je linearna kombinacija radijalnih funkcija. Ako se težine veza između skrivenog i izlaznog sloja promatraju kao regresijski koeficijenti u modelu

višestrukih linearnih regresija danog formulom (2-1), tada se njihove vrijednosti mogu procijeniti korištenjem metode najmanjih kvadrata. Formula (2-1) se može zapisati u matričnom obliku:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3-7)$$

gdje su:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} \\ x_{21} & x_{22} & \dots & x_{2j} \\ \dots & \dots & \dots & \dots \\ x_{a1} & x_{a2} & \dots & x_{aj} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_j \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \dots \\ \varepsilon_a \end{bmatrix},$$

za  $a$  ulaznih podataka, te  $j$  čvorova u skrivenom sloju. Prema [14], članove matrice  $\boldsymbol{\beta}$  je potrebno procijeniti tako pogreška bude što manja zato što ne postoji uvijek rješenje koje svodi pogrešku na nulu. Procjena  $\mathbf{B}$  regresijskih koeficijenata  $\boldsymbol{\beta}$  po metodi najmanjih kvadrata za model višestrukih linearnih regresija jest [14]:

$$\mathbf{B} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_t \quad (3-8)$$

uz pretpostavku da  $(\mathbf{X}^T \mathbf{X})$  nije singularna matrica. Primjenjujući formulu (3-8),  $\mathbf{X}$  postaje matrica koja sadržava vrijednosti izlaza iz skrivenog sloja,  $\mathbf{Y}_t$  matrica ciljanih izlaza iz mreže određenih skupom podataka za treniranje, a  $\mathbf{B}$  matrica koja sadrži težine veza između skrivenog i izlaznog sloja.  $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$  se još naziva Moore–Penroseov pseudo-inverz matrice  $\mathbf{X}$ , te se ova metoda još zove metoda pseudo-inverza.

### 3.2 Određivanje veličine radijalne mreže

Problem određivanja veličine radijalne mreže odnosi se na problem određivanja broja čvorova u skrivenom sloju. Dok je broj čvorova u ulaznom i izlaznom sloju stalan za vrijeme cjelokupnog postupka treniranja, broj skrivenih čvorova je potrebno odrediti na početku treniranja.

Prema [15], izazov u određivanju veličine radijalne mreže predstavljaju problemi pretjerane i premale prilagođenosti podacima za treniranje. Pretjerana prilagođenost podacima za treniranje (engl. *overfitting*) se pojavljuje kod prevelikog broja čvorova u skrivenom sloju, te označava rezultat analize koji pretjerano odgovara danim podacima zbog čega postoji mogućnost smanjene sposobnosti generalizacije i pouzdanog predviđanja. S druge strane, premala prilagođenost

podacima za treniranje (engl, *underfitting*) se odnosi na premali broj čvorova u skrivenom sloju, što rezultira smanjenjem mogućnosti modeliranja skupa podataka za treniranje, s istim krajnjim učinkom kao pretjerana prilagođenost podacima za treniranje.

Jedna od metoda za određivanje veličine mreže je metoda pokušaja i pogreške, koja se dijeli na dva pristupa [15]:

1. Pristup unaprijed – uzima se manji broj čvorova u skrivenom sloju (uobičajeno dva), nakon čega se radi treniranje i validacija neuronske mreže korištenjem neke od mjera prilagođenosti mreže podacima za treniranje (primjerice MSE ili SSE). Nakon toga se slijedno povećava broj čvorova i ponavlja treniranje i validacija s ciljem poboljšanja rezultata.
2. Pristup unazad – izvršava se na isti način kao pristup unaprijed, osim što se počinje s većim brojem čvorova i slijedno smanjuje dok se rezultati ne pogoršaju.

Panchal et al. u [15] navode i druge metode, koje uključuju metodu općenitih smjernica, metodu dvije faze, i sekvencijalni ortogonalni pristup.



## 4. OSTVARENO PROGRAMSKO RJEŠENJE

Programsko rješenje obuhvaća implementaciju radijalne neuronske mreže za regresiju i korisničko sučelje u komandnom prozoru (engl. *Command Line Interface, CLI*). Ostvareno je koristeći programski jezik Python, verzije 3.6.2. Python je interpreterski jezik visoke razine, te za njega postoji bogata podrška za rad sa strojnim učenjem.

Za samu implementaciju radijalne mreže korištena je biblioteka *numpy*, koja omogućava podršku za rad s matricama. Dodatno, korišteni su *matplotlib* za iscrtavanje grafova, *pandas* za rad s skupovima podataka, te *sklearn* iz kojeg je korišten algoritam *k-means* i raspodjela skupa podataka na podskupove za treniranje i validaciju.

### 4.1 Način rada programskog rješenja

Prije svega je potrebno učitati datoteku koja sadrži odabrani skup podataka tako da se preda putanja do datoteke. Programsko rješenje radi s datotekama CSV (engl. *Comma Separated Values*) formata. Učitani skup podataka se naziva okvir podataka (engl. *Dataframe*) za koji je potrebno obaviti određene postupke predobrade, a koji uključuju uklanjanje duplikata, parsiranje broječnih vrijednosti i obradu zaglavlja (ako postoji) korištenjem *pandas* biblioteke. Nakon toga se okvir podataka dijeli na ulazne i izlazne podskupove, gdje ulazni podaci uključuju jedan ili više stupaca CSV datoteke, a izlazni jedan stupac. Oba podskupa su korisnički definirana i ne smiju uključivati vrijednosti koje nedostaju (primjerice "?") te podatke koji nisu brojevi (primjerice nazivi ili datumi). Potom se oba podskupa pretvaraju u matrični zapis korištenjem *numpy* biblioteke kako bi se s njima moglo dalje raditi.

Implementacija radijalne mreže za regresiju uključuje klasu čiji su članovi parametri koje koristi radijalna mreža: veličina mreže, središta i širine radijalnih funkcija, te matrica s vrijednostima težina veza između skrivenog i izlaznog sloja. Važnije metode uključuju:

- **rbf** – aktivacijska funkcija i implementacija Gaussove funkcije dane formulom (2-8) koja ulazni podatak koristi kao ulaz u Gaussovu funkciju i vraća izlaznu vrijednost
- **feedforward** – metoda koja za primljeni ulaz računa dvije matrice koje sadrže izlaze iz skrivenog i izlaznog sloja i vraća ih.
- **train** – metoda koja prima listu s ulazima i ciljanim izlazima i ažurira vrijednosti težina veza između skrivenog i izlaznog sloja. Metoda koja se koristi za treniranje je Moore-Penroseov pseudo-inverz dan formulom (3-8). U izračunu se ponekad pojavljuje singularna matrica koja nema inverz, te u tom slučaju nije moguće izvršiti treniranje. Za

takve slučajeve metoda vraća neistinu (engl. *false*) i preskače se u evaluaciji uspješnosti treniranja.

- **get\_MSE** – metoda koja za primljenu listu ulaza i ciljanih izlaza računa srednju kvadratnu pogrešku po formuli (2-11).
- **calculate\_centers** – metoda koja računa koordinate središta radijalnih funkcija. Prima broj središta, ulazni skup podataka te način određivanja. Način određivanja može biti korištenje nasumičnih točaka iz ulaznog skupa podataka ili korištenje algoritma *k-means*. U slučaju da je odabrano korištenje algoritma *k-means*, koristi se implementacija prema [16] iz biblioteke *sklearn*.
- **calculate\_std** – metoda koja računa širine radijalnih funkcija za primljenu listu središta radijalnih funkcija. Koristi implementaciju algoritma *p-najbližih susjeda* (*NearestNeighbors()*) prema [17] iz biblioteke *sklearn* kako bi našla jednog najbližeg susjeda. Ako se traži prvi najbliži susjed, metodi *NearestNeighbors()* je potrebno predati broj dva iz razlog što se u kalkulaciji svako središte radijalne funkcije smatra vlastitim susjedom. Širina se nakon dobivene udaljenosti računa po formuli (3-4).
- **analyze** – metoda koja prema primljenim parametrima (ulazi i odgovarajući očekivani izlazi, raspon broja čvorova u skrivenom sloju, faktor skaliranja širine radijalne funkcije, te način određivanja središta radijalnih funkcija) poziva metode *calculate\_centers()* i *calculate\_std()*, stvara instancu radijalne mreže, trenira ju, te računa srednju kvadratnu pogrešku korištenjem algoritma *k-rezova*.

Ulazne vrijednosti programa su skup podataka, te parametri mreže. Program radi tako da učitava skup podataka, te za dane parametre mreže poziva metodu *analyze()* za svaki član iz skupa faktora skaliranja širina  $q$  koja vraća listu koja sadrži MSE za sve zadane veličine radijalne mreže. Izlazna vrijednost programa je objekt s vrijednostima MSE za sve kombinacije veličine mreže i faktora skaliranja širine  $q$ , a koji se prikazuje korisniku na više načina, kako je objašnjeno u nastavku.

## 4.2 Prikaz i način uporabe programskog rješenja

Korisničko sučelje programskog rješenja je prikazano slikom 4.1. Prvo je potrebno učitati datoteku koja sadrži skup podataka, i to na način da korisnik unese apsolutnu ili relativnu putanju do željene datoteke. Također je potrebno specificirati nalazi li se na vrhu datoteke redak s imenima stupaca (engl. *header*). Nakon toga se skup podataka odvaja na ulazni i izlazni podskup i to tako da korisnik određuje stupac ili raspon stupaca za ulazni i stupac koji označava

izlazni podskup. Ako je uspješno učitavanje datoteke i konfiguracija skupa podataka prelazi se na konfiguraciju radijalne mreže.

```
Enter dataset filepath: data/real_estate_valuation_data.csv
Does your dataset have a header? (y/n): y

Enter input column range separated by a comma (first column is numbered 0), e.g. 0,6: 0,6
Enter the column number for output (first column is numbered 0): 6

-----
Dataset import complete.
-----

Use Default radial basis function network configuration?
- cluster number range: [2, 20]
- width factor q values: [1, 1.2, 1.4, 1.6, 1.8, 2]
- K-Means clustering enabled
If no, configuration prompts will be shown. (y/n): n

Enter cluster number range separated by a comma, e.g. 2,20 or press Return to use default: 2,20
Enter width factor q values separated by a comma, e.g. 1,1.2,1.4,1.6 or press Return to use default:
Use K-means clustering? (y/n): y

-----
Radial basis function network configuration finished.
-----

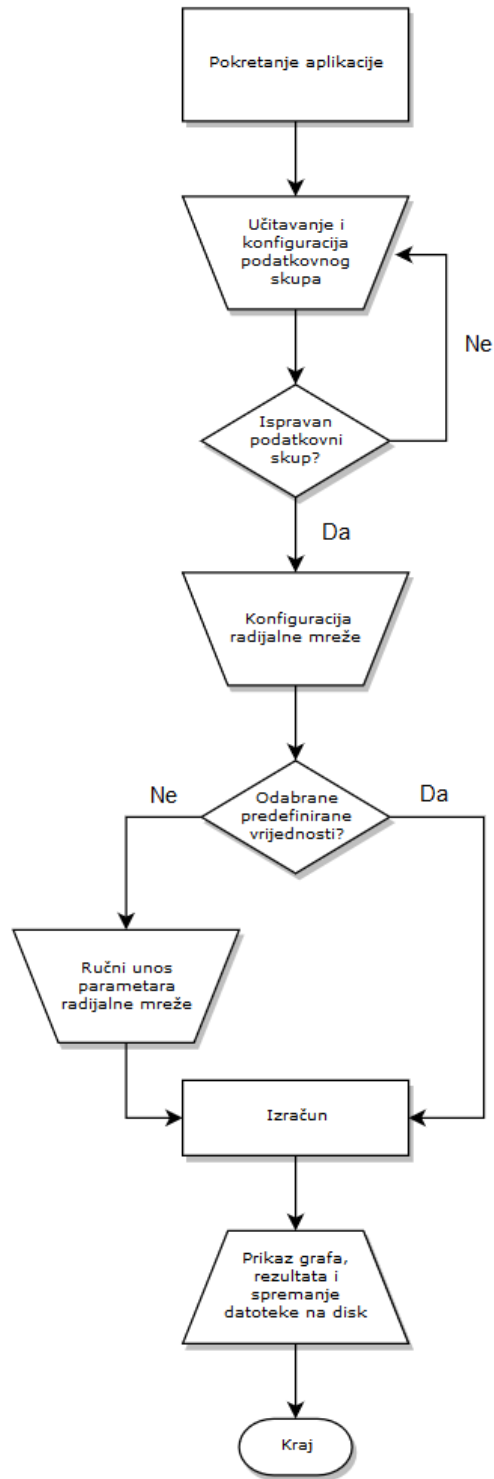
Enter path for data export: results

-----
Export configuration finished. Starting...
-----
```

Sl. 4.1. Izgled korisničkog sučelja programskog rješenja

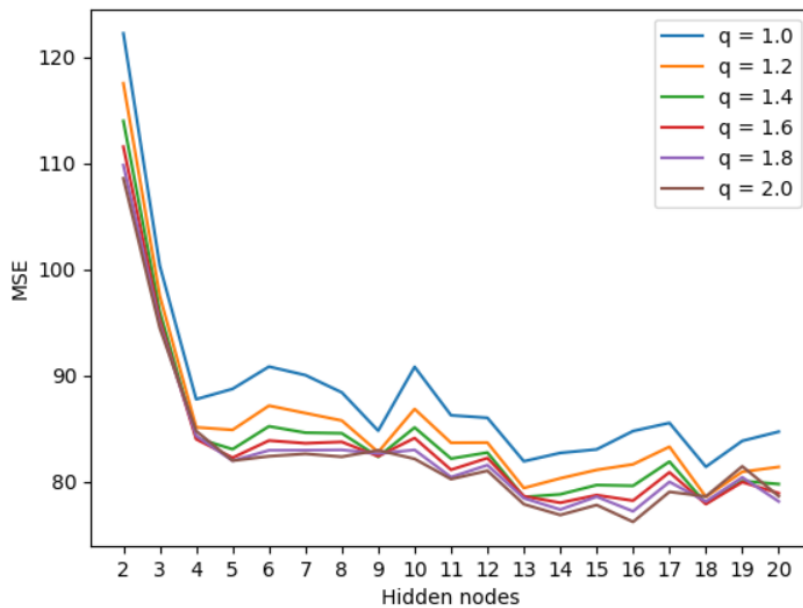
Radijalnu mrežu je moguće konfigurirati u jednom koraku koristeći predefinirane vrijednosti parametara, ili je moguće ručno unositi parametre ako je potrebno. Predefinirane vrijednosti su, kako je prikazano slikom 4.3., raspon broja čvorova u skrivenom sloju od 2 do 20, faktor skaliranja širine  $q$  je skup  $\{1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$ , te se koristi algoritam *k-means* za odabir središta radijalnih funkcija. Cjelokupni tijek rada programskog rješenja prikazan je slikom 4.2.

Nakon završetka analize se korištenjem *matplotlib* biblioteke korisniku na uvid daje graf rezultata, što je prikazano primjerom danim slikom 4.3. Na grafu je prikazan MSE u ovisnosti o veličini mreže i faktoru skaliranja širine  $q$ .



Sl. 4.2. Dijagram toka izvođenja programskog rješenja

Osim grafa, rezultati se spremaju na disk u datoteku XLSX formata kako bi bili dostupni za daljnju obradu. XLSX je format koji koristi program Microsoft Excel, i koji pruža bogatu podršku za prikaz i obradu podataka, a uz to omogućuje izvoz podataka u druge formate po potrebi.



Sl. 4.3. Primjer grafa rezultata pri završenoj analizi

Dodatno, programsko rješenje za vrijeme svog rada u konzolu ispisuje vrijednosti MSE pri svakom izračunu, kao na slici 4.4. Ovo korisniku daje na uvid ispravnost rada programa i rezultate za vrijeme izvođenja.

```

Hidden nodes: 2, q=1.0, MSE: 121.98996874162489
Hidden nodes: 3, q=1.0, MSE: 100.3575237987736
Hidden nodes: 4, q=1.0, MSE: 87.76321040053475
Hidden nodes: 5, q=1.0, MSE: 88.74460616400583
Hidden nodes: 6, q=1.0, MSE: 90.8440913594813
Hidden nodes: 7, q=1.0, MSE: 90.04064168754643
Hidden nodes: 8, q=1.0, MSE: 88.41140291385982
Hidden nodes: 9, q=1.0, MSE: 84.79738659526234
Hidden nodes: 10, q=1.0, MSE: 90.82178468727184
Hidden nodes: 11, q=1.0, MSE: 86.25458291530717
Hidden nodes: 12, q=1.0, MSE: 86.01452695185158
Hidden nodes: 13, q=1.0, MSE: 81.91520505120218
Hidden nodes: 14, q=1.0, MSE: 82.70470471789832
Hidden nodes: 15, q=1.0, MSE: 83.03787355544083
Hidden nodes: 16, q=1.0, MSE: 84.76769870205672
Hidden nodes: 17, q=1.0, MSE: 85.53507273940298
Hidden nodes: 18, q=1.0, MSE: 81.39579827011951
Hidden nodes: 19, q=1.0, MSE: 83.84081878853047
Hidden nodes: 20, q=1.0, MSE: 84.70655827217513

```

Sl. 4.4. Primjer ispisa rezultata na ekran za vrijeme izvođenja

## 5. EKSPERIMENTALNA ANALIZA

Cilj eksperimentalne analize je utvrđivanje razlike u kvaliteti dobivenih rezultata regresijske analize u ovisnosti o veličini radijalne neuronske mreže te načinu odabira središta radijalnih funkcija  $\mu$  i širina radijalnih funkcija  $\sigma$ . Kao mjera kvalitete dobivenih rezultata korištena je MSE (srednja kvadratna pogreška). Eksperimentalna analiza provedena je na pet skupova podataka. Skupovi podataka su "Airfoil Self-Noise", "Concrete Compressive Strength", "Energy efficiency", "Real Estate Valuation" te "Residential Building", koji su prikazani tablicom 5.1. Odabrani skupovi podataka preuzeti su s [18].

Tab. 5.1. Svojstva korištenih skupova podataka

Ime	Oznaka	Broj primjeraka	Broj korištenih ulaznih značajki	Broj korištenih izlaznih značajki	Raspon zavisne varijable
Airfoil Self-Noise	ASN	1500	6	1	103.38-140.987
Concrete Compressive Strength	CCS	1030	8	1	2.33-82.6
Energy efficiency	EE	768	8	1	10.9-48.3
Real Rstate Valuation	REV	414	6	1	7.6-117.5
Residential Building	RB	372	8	1	50-6800

Dodatno, slijedi sažeti opis navedenih skupova podataka:

- ASN je NASA-in skup podataka dobiven ispitivanjima aerodinamičkih i zvučnih svojstava dvodimenzionalnih i trodimenzionalnih sekcija poprečnih presjeka krakova propelera provedenih u zračnim tunelima bez odjeka. Ulazne značajke opisuju frekvenciju, napadni kut, duljina zamišljene linije aeroprofila koja spaja njegov napadni i izlazni rub, brzina, i debljina pomaka usisne strane, a izlazna opisuje razinu zvučnog pritiska.
- CCS je skup podataka skup koji opisuje čvrstinu betona u ovisnosti o sastojcima i starosti. Ulazne značajke opisuju udio sastojaka betona (cementa, troske visoke peći, letećeg pepela, superplastifikatora, grubog i finog agregata) i starosti. Izlazna značajka opisuje tlačnu čvrstoću betona.

- EE je skup podataka koji opisuje energetska učinkovitost zgrade kao funkciju njenih građevinskih svojstava koja uključuju relativnu kompaktnost, ukupnu površinu, površinu zida, površinu krova, visinu, orijentaciju, površinu ostakljenja i njenu raspodjelu. Izlazna značajka je opterećenje hlađenja.
- REV je skup podataka koji prikazuje cijene nekretnina u okrugu Sincan New Taipei Cityja u Tajvanu. Cijene nekretnina kao izlazna značajka su funkcija datuma transakcije, starosti, udaljenosti do najbliže tranzitne postaje, broja trgovina u blizini, i geografske duljine i širine.
- RB je skup podataka koji opisuje cijenu prodaje stanova u Teheranu kao funkciju lokacije, stambene površine, površine zemljišta, tri preliminarne procjene cijene gradnje na početku projekta, trajanja gradnje i cijene kvadratnog metra na početku gradnje.

## 5.1. Postavke eksperimenta

Veličina radijalne neuronske mreže  $n$  određena je slijedno, ispitivanjem rezultata za 2 do 20 čvorova u skrivenom sloju. Korišteni načini odabira središta radijalnih funkcija  $\mu$  su nasumičan odabir  $j$  točaka iz skupa za treniranje, gdje je  $j$  broj skrivenih čvorova mreže, te algoritam za grupiranje podataka *k-means*. Korištene postavke algoritma *k-means* su:

- zadani broj grupa odgovara broju čvorova u skrivenom sloju mreže  $j$
- odabir početnih središta grupa je nasumičan
- algoritam se izvršava 10 puta s različitim početnim središtima grupa te se uzima najbolji rezultat
- maksimalan broj iteracija je 300

Način izračuna širina radijalnih funkcija  $\sigma$  je korištenje algoritma p-najbližih susjeda prema formuli (3-3). Korištene postavke algoritma p-najbližih susjeda su:

- broj susjeda je jednak jedan
- korištena je euklidska udaljenost

Eksperimentalna analiza provedena je za dva scenarija, zadanih u ovisnosti o načinu određivanja središta radijalnih funkcija  $\mu$ . Zadani scenariji su:

1. Algoritam *k-means* korišten je za određivanje središta i algoritam p-najbližih susjeda za određivanje širina radijalnih funkcija
2. Korišten je nasumičan odabir središta i algoritam p-najbližih susjeda za određivanje širina radijalnih funkcija

Dodatno, za svaki scenarij sve dobivene širine  $\sigma$  pomnožene su faktorom skaliranja širine  $q$  prema formuli (3-4), gdje  $q$  poprima vrijednost svakog člana iz skupa  $\{1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$ .

## 5.2. Rezultati

Rezultati ostvareni provedenim eksperimentima dani u nastavku. Prikazane su vrijednosti srednje kvadratne pogreške tablično i grafički za sve skupove podataka i scenarije. Zelenom bojom u tablici je označen najbolji rezultat, a crvenom najlošiji za svaki od scenarija. U slučaju da su rezultati vrlo slični za različite veličine mreže, izvršena je dodatna obrada korištenjem funkcije dobrote (engl. *fitness function*) [19]:

$$f(v) = MSE + \lambda j, \quad (5-1)$$

gdje je vrijednost  $\lambda$  određena eksperimentalno, a  $j$  je broj čvorova u skrivenom sloju. Koeficijent  $\lambda$  se određuje posebno za svaki skup podataka. Zbog razlike u rasponima izlaznih vrijednosti, a prema tome u MSE, nije moguće odrediti jednostruki koeficijent  $\lambda$ . Dodatno, funkcija dobrote se koristi samo za scenarij u kojem su dobivene manje vrijednosti MSE.

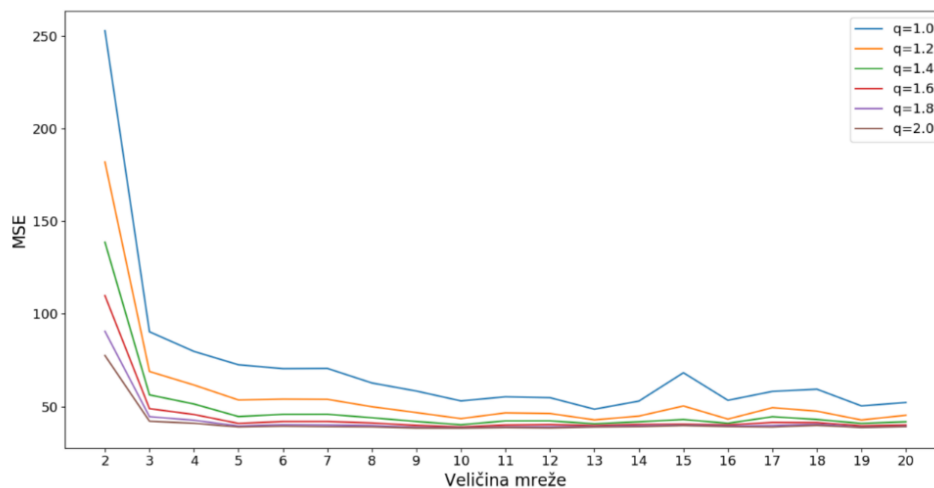
Iz tablice 5.2. i grafa 5.1. može se iščitati ovisnost MSE o broju čvorova u skrivenom sloju mreže i faktoru skaliranja širine  $q$  za scenarij 1 u analizi skupa ASN. Promatrajući tablicu 5.2. po stupcima može se primijetiti da se kvaliteta predviđanja povećava s porastom broja čvorova u skrivenom sloju do određene granice. Uzimajući u obzir retke, također se vidi smanjenje MSE s povećanjem faktora skaliranja širine  $q$ .

Pad MSE se događa do određene točke, nakon koje se dodavanjem čvorova u skriveni sloj samo povećavaju potrebni računalni resursi i vrijeme bez stvarnog povećanja kvalitete predviđanja. Na primjer, za  $q=1.8$  može se primijetiti da se za 5 i više skrivenih čvorova dobiva vrlo sličan rezultat, a dobivenim razlikama u MSE je uzrok nasumičan odabir primjera za rezove. Ovakve varijacije se ne mogu izbjeći ako se koristi algoritam k-rezova, ali su dovoljno male kako bi se mogle zanemariti. Isto vrijedi i za  $q=2$ , s marginalno boljim rezultatima. S ciljem određivanja najboljeg rezultata s obzirom na računalnu kompleksnost, na rezultate u tablici 5.2. primijenjena je funkcija dobrote prema formuli (5-1) za vrijednost  $\lambda = 0.5$ . Novi rezultati su dani tablicom 5.3.



Tab. 5.2. Ovisnost MSE o veličini mreže na skupu ASN i scenarij 1 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	252.73	181.89	138.60	109.77	90.52	77.47
3	90.30	68.87	56.24	48.78	44.46	42.02
4	79.69	61.56	51.31	45.64	42.56	40.91
5	72.48	53.48	44.53	40.84	39.44	38.91
6	70.40	53.95	45.74	41.88	40.13	39.32
7	70.51	53.84	45.77	41.89	39.99	39.06
8	62.65	49.83	43.88	41.08	39.68	38.93
9	58.30	46.67	41.86	39.77	38.80	38.36
10	52.98	43.42	40.12	39.00	38.55	38.31
11	55.21	46.52	42.19	39.96	38.94	38.60
12	54.74	46.18	42.22	40.22	39.11	38.41
13	48.50	42.75	40.63	39.66	39.16	38.87
14	52.89	44.78	41.73	40.29	39.52	39.06
15	68.18	50.24	42.93	40.43	39.87	39.72
16	53.31	43.17	40.91	40.02	39.46	39.07
17	58.12	49.31	44.44	41.36	39.66	38.88
18	59.30	47.44	42.97	41.27	40.39	39.80
19	50.29	42.68	40.84	39.55	38.84	38.62
20	52.14	45.24	41.72	39.94	39.27	39.06



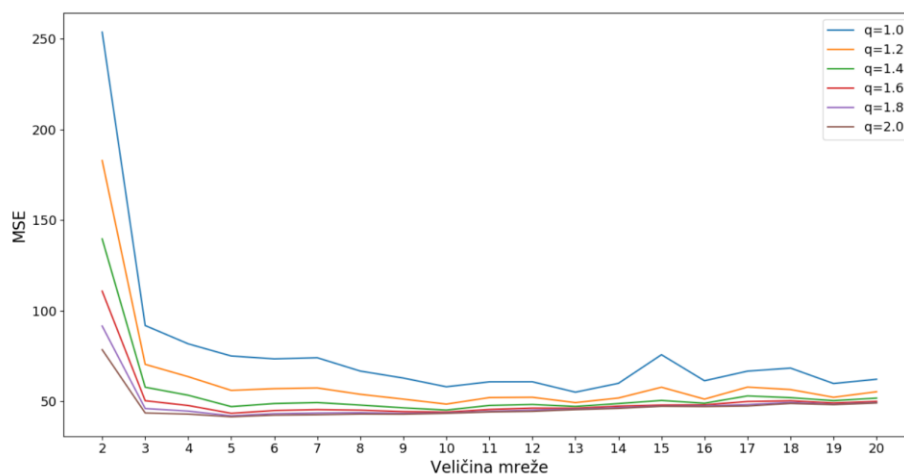
Sl. 5.1. Ovisnost MSE o veličini mreže na skupu ASN i scenarij 1 za različite faktore skaliranja širina  $q$

Iz tablice 5.3 i grafa na slici 5.2. može se iščitati porast MSE ako se povećava broj čvorova u skrivenom sloju jer funkcija dobrote kažnjava korištenje većih mreža. Nova najniža vrijednost je za pet čvorova u skrivenom sloju i faktor  $q=2$ . Vrijednost MSE za istu veličinu mreže i faktor

$q=1.8$  je vrlo slična, a u takvim slučajevima se uzima manja vrijednost, što znači da je vrijednost  $q=1.8$  najpogodnija.

Tab. 5.3. Ovisnost MSE o veličini mreže na skupu ASN i scenarij 1 za različite faktore skaliranja širina  $q$  nakon primjene funkcije dobrote

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	253.73	182.89	139.60	110.77	91.52	78.47
3	91.80	70.37	57.74	50.28	45.96	43.52
4	81.69	63.56	53.31	47.64	44.56	42.91
5	74.98	55.98	47.03	43.34	41.94	41.41
6	73.40	56.95	48.74	44.88	43.13	42.32
7	74.01	57.34	49.27	45.39	43.49	42.56
8	66.65	53.83	47.88	45.08	43.68	42.93
9	62.80	51.17	46.36	44.27	43.30	42.86
10	57.98	48.42	45.12	44.00	43.55	43.31
11	60.71	52.02	47.69	45.46	44.44	44.10
12	60.74	52.18	48.22	46.22	45.11	44.41
13	55.00	49.25	47.13	46.16	45.66	45.37
14	59.89	51.78	48.73	47.29	46.52	46.06
15	75.68	57.74	50.43	47.93	47.37	47.22
16	61.31	51.17	48.91	48.02	47.46	47.07
17	66.62	57.81	52.94	49.86	48.16	47.38
18	68.30	56.44	51.97	50.27	49.39	48.80
19	59.79	52.18	50.34	49.05	48.34	48.12
20	62.14	55.24	51.72	49.94	49.27	49.06



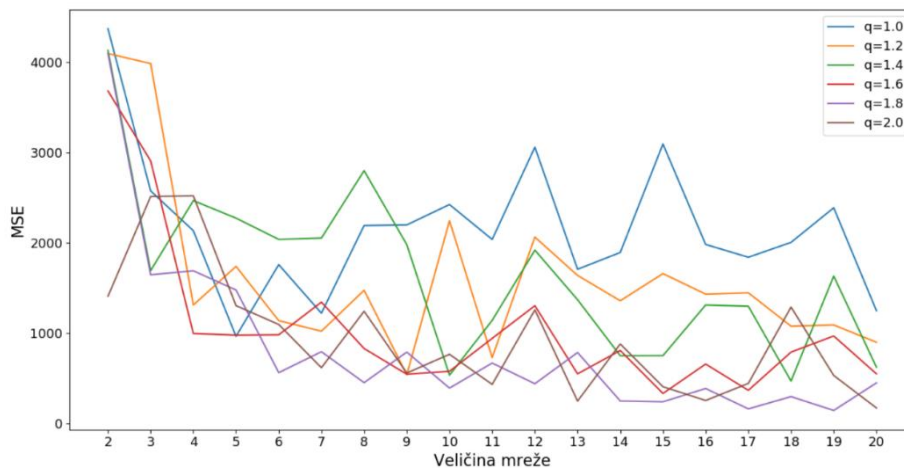
Sl. 5.2. Ovisnost MSE o veličini mreže na skupu ASN i scenarij 1 za različite faktore skaliranja širina  $q$  nakon primjene funkcije dobrote

Iz tablice 5.4. i slike 5.3. može se iščitati ovisnost MSE o broju čvorova u skrivenom sloju mreže i faktoru skaliranja širine  $q$  za scenarij 2 u analizi skupa ASN. Najveće su greške za dva skrivena

čvora. Vidljiv je trend pada MSE od dva do osam skrivenih čvorova, a za više od osam se može primijetiti da postoje velike varijacije, to jest nagli rast i pad vrijednosti MSE. Ovo je uzrokovano nasumičnim odabirom središta radijalnih funkcija. Manje širine radijalnih funkcija daju lošije rezultate, dok su najbolji rezultati dobiveni za  $q=1.6$  i više, slično kao kod scenarija 1. Središta radijalnih funkcija  $\mu$  su određena nasumično iz skupa za treniranje, i to se odražava na kvalitetu predviđanja - MSE je za sve kombinacije veličine mreže i faktora skaliranja širine  $q$  mnogo veći nego za scenarij 1. Izostanak algoritma za grupiranje ne može se nadoknaditi povećanjem broja skrivenih čvorova do 20 u slučaju skupa ASN, niti povećanjem širina radijalnih funkcija do udvostručenja vrijednosti određene korištenjem algoritma p-najbližih susjeda. Iz ovog razloga nije korištena funkcija dobrote za ovaj scenarij.

Tab. 5.4. Ovisnost MSE o veličini mreže na skupu ASN i scenarij 2 za različite faktore skaliranja širina  $q$

Veličina mreže	q					
	1.0	1.2	1.4	1.6	1.8	2.0
2	4372.70	4097.62	4133.04	3683.56	4098.85	1408.17
3	2576.09	3986.79	1693.34	2908.34	1647.32	2513.85
4	2135.10	1312.21	2468.78	996.41	1691.40	2521.68
5	965.32	1739.78	2274.41	978.69	1478.30	1305.16
6	1759.47	1139.04	2039.14	982.16	562.93	1096.53
7	1221.06	1022.03	2053.77	1344.31	794.64	617.25
8	2192.54	1475.56	2800.18	830.10	452.01	1242.79
9	2199.58	539.90	1980.84	545.37	790.65	558.76
10	2425.25	2245.43	535.13	578.16	391.70	766.21
11	2038.40	729.61	1141.87	942.66	669.21	432.78
12	3060.45	2064.54	1920.06	1304.50	440.22	1257.81
13	1708.32	1640.01	1370.55	549.59	785.91	248.20
14	1893.96	1358.99	750.02	808.73	249.14	880.43
15	3095.36	1660.32	751.30	332.29	240.26	406.80
16	1982.85	1432.74	1312.24	658.11	387.04	254.67
17	1841.28	1447.38	1298.78	366.99	161.62	443.48
18	2005.73	1077.54	469.71	790.32	297.38	1289.70
19	2389.17	1091.00	1632.69	968.84	144.20	531.56
20	1249.46	899.36	623.50	549.82	449.14	172.01

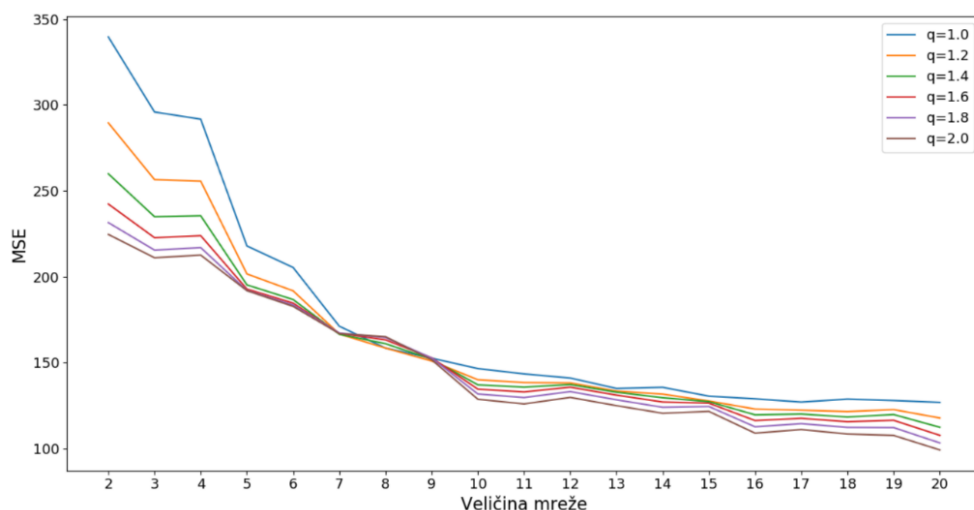


Sl. 5.3. Ovisnost MSE o veličini mreže na skupu ASN i scenarij 2 za različite faktore skaliranja širina  $q$

Iz tablice 5.5. i grafa na slici 5.4. može se iščitati ovisnost MSE o broju čvorova u skrivenom sloju mreže i faktoru skaliranja širine  $q$  za scenarij 1 u analizi skupa CCS. Promatrajući tablicu 5.5. po stupcima može se primijetiti da se kvaliteta predviđanja povećava s porastom broja čvorova u skrivenom sloju do određene granice. Uzimajući u obzir retke, također se vidi smanjenje MSE s povećanjem faktora skaliranja širine  $q$  kao i slučaju skupa ASN.

Tab. 5.5. Ovisnost MSE o veličini mreže na skupu CCS i scenarij 1 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	339.62	289.59	259.95	242.35	231.55	224.67
3	295.91	256.57	234.98	222.75	215.53	211.09
4	291.73	255.67	235.52	223.93	216.98	212.63
5	217.95	201.68	195.29	192.85	192.05	191.95
6	205.43	191.83	186.77	184.69	183.61	182.79
7	171.24	166.70	166.59	167.11	167.28	166.95
8	158.53	158.53	161.08	163.31	164.65	165.16
9	152.65	151.00	152.05	152.85	152.74	151.70
10	146.58	140.07	137.12	134.60	131.81	128.74
11	143.49	138.46	135.78	133.01	129.72	126.02
12	141.08	138.24	137.29	135.76	133.20	129.81
13	135.02	133.49	132.83	131.15	128.42	125.02
14	135.68	131.73	129.63	127.14	124.05	120.58
15	130.54	127.72	127.35	126.47	124.53	121.67
16	128.99	123.03	119.67	116.34	112.72	109.01
17	127.11	122.38	120.13	117.63	114.55	111.11
18	128.83	121.55	118.38	115.67	112.37	108.51
19	128.01	122.72	119.85	116.49	112.26	107.70
20	126.91	117.83	112.44	107.74	103.33	99.25



Sl. 5.4. Ovisnost MSE o veličini mreže na skupu CCS i scenarij 1 za različite faktore skaliranja širina  $q$

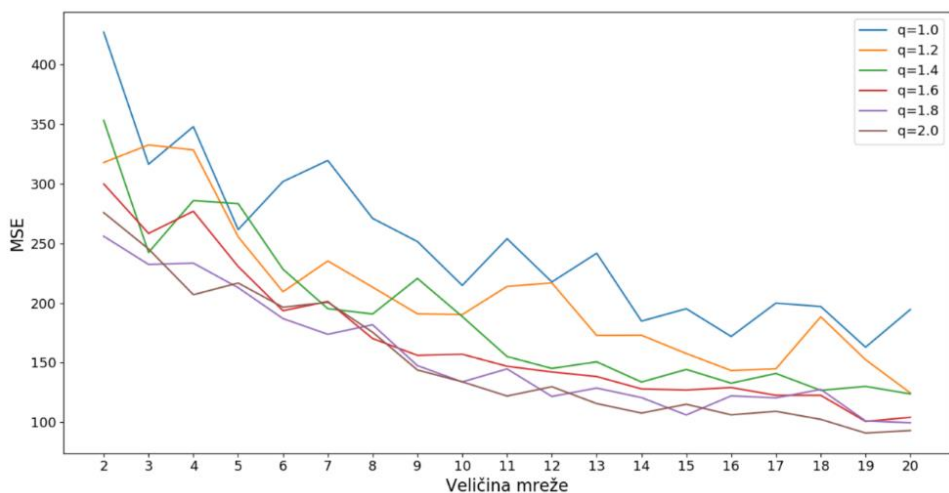
Najbolji rezultat je dobiven za najveću širinu radijalnih funkcija s najviše čvorova u skrivenom sloju. Za razliku od ASN skupa, razlike nisu zanemarive, to jest povećanjem veličine mreže i faktora  $q$  MSE opada. Iz ovog razloga nije korištena funkcija dobrote jer bi se korištenjem manjeg broja čvorova u skrivenom sloju znatno povećala pogreška.

Promatrajući tablicu 5.6. i graf na slici 5.5, može se zamijetiti da su za skup CCS rezultati bolji za veće širine radijalnih funkcija u slučaju da se ne koristi algoritam *k-means* za grupiranje. Razlike nisu velike, ali su zamjetne, pa se može zaključiti da nema potrebe za korištenjem grupiranja jer povećava složenost izračuna bez poboljšanja rezultata. Grupiranje korištenjem algoritma *k-means* može dati neoptimalne rezultate u slučaju da su podaci jednoliko raspoređeni ili su raspoređeni previše složeno za grupiranje prema prostornoj gustoći.

Tab. 5.6. Ovisnost MSE o veličini mreže na skupu CCS i scenarij 2 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	427.03	317.73	353.12	299.81	255.94	275.72
3	316.30	332.49	242.34	258.33	232.29	245.28
4	347.80	328.32	285.84	276.90	233.39	206.92
5	261.51	255.46	283.28	230.55	212.82	216.74
6	301.81	209.41	228.28	193.49	186.91	196.42
7	319.47	235.20	195.22	201.25	173.74	200.60
8	270.91	213.25	190.76	170.19	181.75	175.46
9	251.51	190.91	220.58	156.09	147.48	143.82
10	214.74	190.28	188.63	157.02	133.82	133.93
11	253.88	213.86	155.09	146.94	144.73	121.97

12	217.62	216.83	145.12	142.15	121.63	129.79
13	241.67	172.78	150.66	138.32	128.66	115.65
14	184.81	172.93	133.71	127.94	120.65	107.70
15	195.17	157.60	144.27	127.03	106.11	115.07
16	171.91	143.39	132.75	129.10	122.11	106.21
17	199.89	144.83	140.89	122.56	120.44	109.17
18	197.02	188.43	126.72	122.56	127.52	102.41
19	162.93	152.43	130.07	100.69	101.12	90.97
20	194.55	124.53	123.73	104.12	99.57	93.04



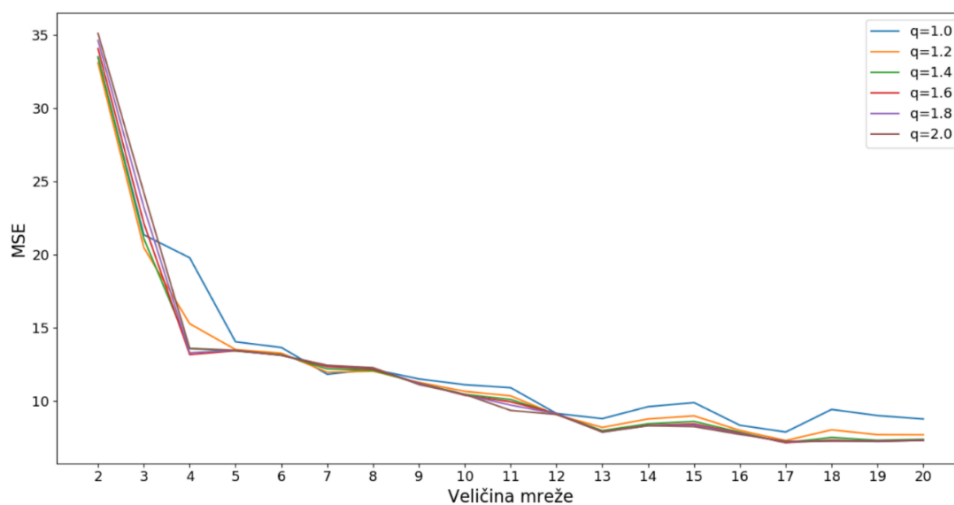
Sl. 5.5. Ovisnost MSE o veličini mreže na skupu CCS i scenarij 2 za različite faktore skaliranja širina  $q$

Tablicom 5.7 i grafom na slici 5.6 su dani rezultati analize za skup EE za scenarij 1. Pokazano je da, kao za skupove ASN i CCS, opada vrijednost MSE povećanjem broja čvorova u skrivenom sloju. Najniže vrijednosti se dobivaju za 13 i više skrivenih čvorova, s malim razlikama ako se povećava veličina mreže. Što se tiče širina radijalnih funkcija, promatrajući retke tablice može se primijetiti da povećanje širina radijalnih funkcija utječe vrlo malo, ili uopće ne utječe na vrijednost MSE.

Tab. 5.7. Ovisnost MSE o veličini mreže na skupu EE i scenarij 1 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	33.50	33.09	33.50	34.07	34.62	35.11
3	21.35	20.47	21.09	22.14	23.23	24.22
4	19.78	15.28	13.58	13.17	13.29	13.60
5	14.04	13.52	13.41	13.43	13.46	13.47
6	13.65	13.26	13.17	13.13	13.12	13.12
7	11.83	11.97	12.19	12.34	12.42	12.44

<b>8</b>	12.17	12.04	12.09	12.18	12.25	12.27
<b>9</b>	11.51	11.27	11.22	11.21	11.19	11.13
<b>10</b>	11.11	10.66	10.46	10.39	10.42	10.46
<b>11</b>	10.91	10.35	10.10	9.95	9.72	9.36
<b>12</b>	9.15	9.06	9.11	9.14	9.13	9.09
<b>13</b>	8.80	8.20	7.98	7.90	7.87	7.89
<b>14</b>	9.62	8.78	8.45	8.34	8.32	8.32
<b>15</b>	9.89	8.99	8.61	8.44	8.35	8.25
<b>16</b>	8.34	7.99	7.87	7.81	7.77	7.72
<b>17</b>	7.88	7.30	7.15	7.15	7.20	7.25
<b>18</b>	9.42	8.03	7.52	7.33	7.28	7.27
<b>19</b>	9.01	7.70	7.31	7.24	7.25	7.28
<b>20</b>	8.77	7.69	7.39	7.32	7.32	7.33



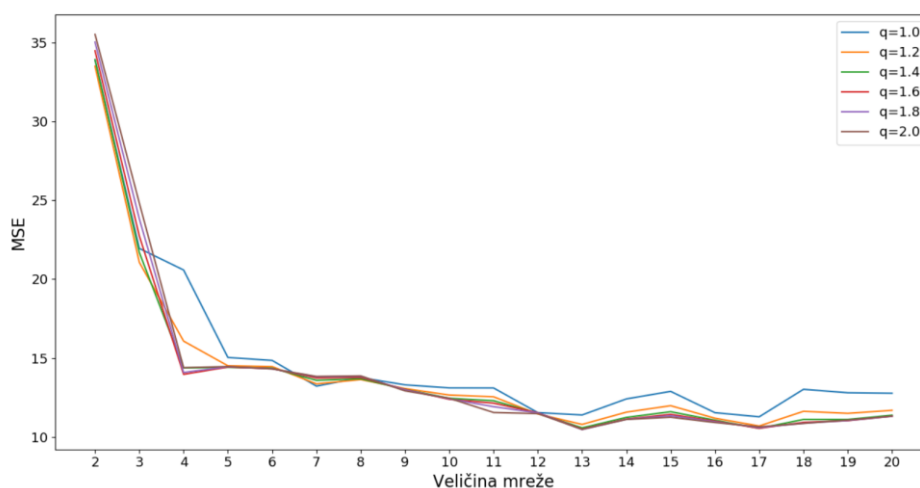
Sl. 5.6. Ovisnost MSE o veličini mreže na skupu EE i scenarij 1 za različite faktore skaliranja širina  $q$

Najlošiji rezultat je ponovno dobiven korištenjem vrlo malog broja skrivenih čvorova, ali za najveću širinu radijalnih funkcija. Ovakva nedosljednost se može objasniti načinom vrednovanja mreže, to jest korištenjem algoritma k-rezova, gdje su male varijacije ovog tipa česte. Najniže vrijednosti se dobivaju za 13 i više skrivenih čvorova, s malim ili nikakvim poboljšanjem dobivenim daljnjim povećavanjem mreže. Kako bi se odredio najbolji rezultat s obzirom na potrebne računalne resurse, rezultati za ovaj scenarij su dodatno obrađeni funkcijom dobrote (formula (5-1)) s varijablom  $\lambda = 0.2$ , što je dano tablicom 5.8 i grafom na slici 5.7.

Nakon primjene funkcije dobrote vidi se da je najniža vrijednost MSE ona dobivena korištenjem 13 čvorova u skrivenom sloju. Pošto faktor  $q$  nema stvaran utjecaj na kvalitetu rezultata u ovom slučaju, odabire se onaj s najmanjom vrijednosti, što je u ovom slučaju 1.2.

Tab. 5.8. Ovisnost MSE o veličini mreže na skupu EE i scenarij 1 za različite faktore skaliranja širina  $q$  nakon primjene funkcije dobrote

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	33.90	33.49	33.90	34.47	35.02	35.51
3	21.95	21.07	21.69	22.74	23.83	24.82
4	20.58	16.08	14.38	13.97	14.09	14.40
5	15.04	14.52	14.41	14.43	14.46	14.47
6	14.85	14.46	14.37	14.33	14.32	14.32
7	13.23	13.37	13.59	13.74	13.82	13.84
8	13.77	13.64	13.69	13.78	13.85	13.87
9	13.31	13.07	13.02	13.01	12.99	12.93
10	13.11	12.66	12.46	12.39	12.42	12.46
11	13.11	12.55	12.30	12.15	11.92	11.56
12	11.55	11.46	11.51	11.54	11.53	11.49
13	11.40	10.80	10.58	10.50	10.47	10.49
14	12.42	11.58	11.25	11.14	11.12	11.12
15	12.89	11.99	11.61	11.44	11.35	11.25
16	11.54	11.19	11.07	11.01	10.97	10.92
17	11.28	10.70	10.55	10.55	10.60	10.65
18	13.02	11.63	11.12	10.93	10.88	10.87
19	12.81	11.50	11.11	11.04	11.05	11.08
20	12.77	11.69	11.39	11.32	11.32	11.33



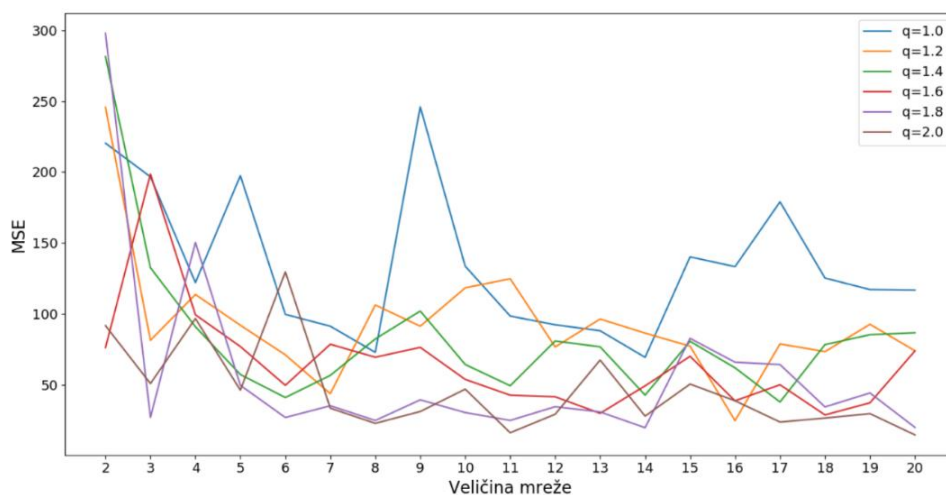
Sl. 5.7. Ovisnost MSE o veličini mreže na skupu EE i scenarij 1 za različite faktore skaliranja širina  $q$  nakon primjene funkcije dobrote

Kao i u slučaju skupa ASN, vide se velike varijacije za niske vrijednosti faktora  $q$  i broja skrivenih čvorova. MSE je nešto manji ako se uzme veća širina radijalnih funkcija, no prema tablici 5.9 i grafu na slici 5.7, najbolji rezultat je dvostruko lošiji od onog dobivenog u scenariju 1, i to korištenjem više skrivenih čvorova.



Tab. 5.9. Ovisnost MSE o veličini mreže na skupu EE i scenarij 2 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	220.35	245.67	281.32	76.33	297.74	91.87
3	196.64	81.55	132.65	198.61	27.24	51.08
4	122.19	113.80	91.10	99.52	150.46	96.88
5	197.49	92.13	57.43	77.14	49.82	46.43
6	99.80	71.31	41.24	49.85	27.18	129.72
7	91.45	43.94	56.85	78.76	35.45	33.74
8	73.05	106.42	82.49	69.62	25.07	23.08
9	245.90	91.45	102.13	76.57	39.61	31.40
10	133.71	118.47	64.50	54.00	30.67	47.14
11	98.62	124.89	49.55	42.91	25.13	16.44
12	92.48	76.86	81.06	41.77	34.79	29.47
13	88.27	96.53	76.95	30.13	31.09	67.62
14	69.56	86.60	42.96	49.47	20.00	28.16
15	140.27	77.40	80.83	70.27	82.91	50.72
16	133.41	24.96	62.15	38.93	66.10	38.97
17	179.03	78.97	38.09	50.24	64.45	23.99
18	125.38	73.58	78.50	29.00	34.63	26.77
19	117.26	92.81	85.48	37.52	44.46	29.90
20	116.88	74.17	86.85	74.04	20.24	14.94



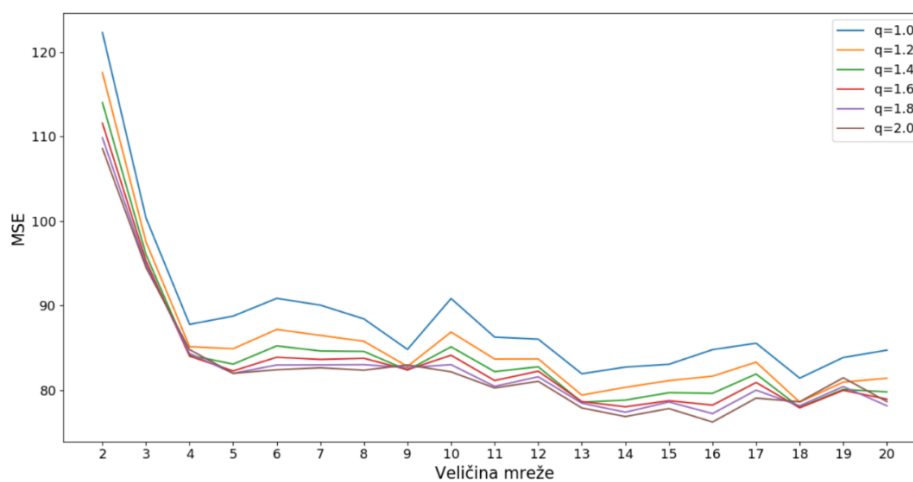
Sl. 5.8. Ovisnost MSE o veličini mreže na skupu EE i scenarij 2 za različite faktore skaliranja širina  $q$

Skup REV u scenariju 1 (tablica 5.10 i slika 5.9) daje slične rezultate kao u prijašnja tri slučaja. MSE opada s povećanjem broja čvorova u skrivenom sloju, a dodatna poboljšanja se dobivaju povećanjem širine radialnih funkcija, no za vrijednost  $q=1.4$  i više su poboljšanja mala. Pošto su vrijednosti MSE vrlo slične za 13 i više skrivenih čvorova, najbolji rezultat određen je

korištenjem funkcije dobrote (formula (5-1)) sa koeficijentom  $\lambda = 0.55$ , čiji su rezultati dani tablicom 5.11 i grafom na slici 5.10.

Tab. 5.10. Ovisnost MSE o veličini mreže na skupu REV i scenarij 1 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	122.32	117.57	114.02	111.58	109.85	108.60
3	100.36	97.52	96.03	95.27	94.83	94.49
4	87.76	85.12	84.11	84.00	84.32	84.77
5	88.74	84.88	83.05	82.25	81.96	81.97
6	90.84	87.16	85.21	83.87	82.95	82.39
7	90.04	86.46	84.61	83.61	82.97	82.62
8	88.41	85.76	84.56	83.74	82.99	82.34
9	84.80	82.78	82.34	82.39	82.64	82.96
10	90.82	86.85	85.10	84.11	82.99	82.14
11	86.25	83.66	82.17	81.11	80.42	80.23
12	86.01	83.67	82.74	82.22	81.56	81.02
13	81.92	79.39	78.55	78.62	78.43	77.87
14	82.70	80.31	78.80	78.00	77.36	76.85
15	83.04	81.10	79.68	78.73	78.58	77.80
16	84.77	81.63	79.60	78.20	77.20	76.20
17	85.54	83.29	81.89	80.88	79.98	79.03
18	81.40	78.56	77.94	77.88	78.12	78.60
19	83.84	80.92	80.06	79.95	80.40	81.44
20	84.71	81.38	79.77	78.91	78.13	78.63

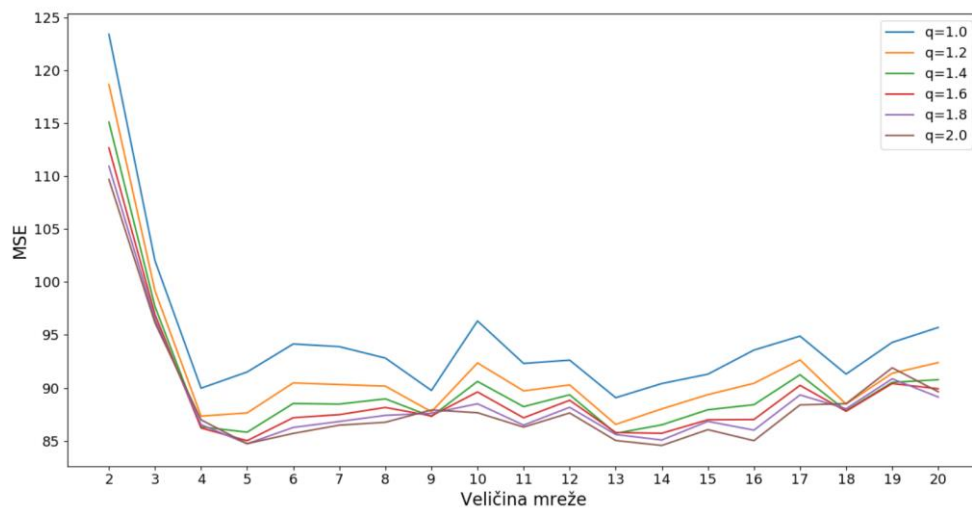


Sl. 5.9. Ovisnost MSE o veličini mreže na skupu REV i scenarij 1 za različite faktore skaliranja širina  $q$

Nakon korištenja funkcije dobrote iz tablice 5.11 i slike 5.10 može se iščitati da je najniža vrijednost MSE za 14 skrivenih čvorova i  $q=2$  jer se kažnjava korištenje veće mreže.

Tab. 5.11. Ovisnost MSE o veličini mreže na skupu REV i scenarij 1 za različite faktore skaliranja širina  $q$  nakon primjene funkcije dobrote

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	123.42	118.67	115.12	112.68	110.95	109.70
3	102.01	99.17	97.68	96.92	96.48	96.14
4	89.96	87.32	86.31	86.20	86.52	86.97
5	91.49	87.63	85.80	85.00	84.71	84.72
6	94.14	90.46	88.51	87.17	86.25	85.69
7	93.89	90.31	88.46	87.46	86.82	86.47
8	92.81	90.16	88.96	88.14	87.39	86.74
9	89.75	87.73	87.29	87.34	87.59	87.91
10	96.32	92.35	90.60	89.61	88.49	87.64
11	92.30	89.71	88.22	87.16	86.47	86.28
12	92.61	90.27	89.34	88.82	88.16	87.62
13	89.07	86.54	85.70	85.77	85.58	85.02
14	90.40	88.01	86.50	85.70	85.06	84.55
15	91.29	89.35	87.93	86.98	86.83	86.05
16	93.57	90.43	88.40	87.00	86.00	85.00
17	94.89	92.64	91.24	90.23	89.33	88.38
18	91.30	88.46	87.84	87.78	88.02	88.50
19	94.29	91.37	90.51	90.40	90.85	91.89
20	95.71	92.38	90.77	89.91	89.13	89.63

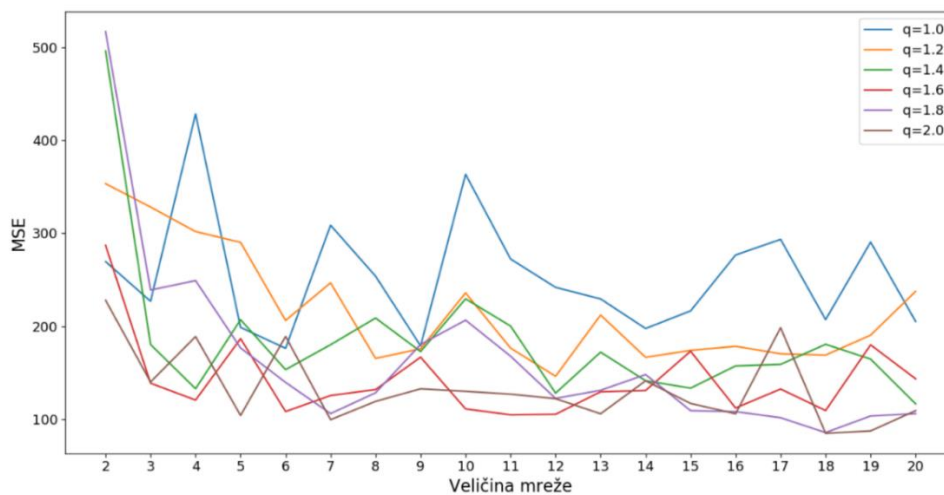


Sl. 5.10. Ovisnost MSE o veličini mreže na skupu REV i scenarij 1 za različite faktore skaliranja širina  $q$  nakon primjene funkcije dobrote

Tablica 5.12 i graf na slici 5.11 pokazuju vrijednost MSE za skup REV u scenariju 2. Varijacije su velike zbog nasumičnog određivanja središta radijalnih funkcija, no općenito bolji su rezultati s više skrivenih čvorova i većim koeficijentom  $q$ . Ipak, rezultati su lošiji nego kad se koristi algoritam za grupiranje, što se poklapa s analizom skupova ASN i EE.

Tab. 5.12. Ovisnost MSE o veličini mreže na skupu REV i scenarij 2 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	269.56	353.49	496.38	287.11	517.20	227.96
3	226.83	328.29	180.25	138.58	238.91	139.90
4	428.56	301.89	132.64	120.43	249.10	188.83
5	198.72	290.26	206.94	186.59	176.58	103.85
6	176.16	206.18	153.21	108.10	139.46	188.95
7	308.72	246.72	180.03	125.32	105.88	99.30
8	253.83	165.26	208.77	131.77	128.26	119.02
9	178.78	175.68	172.69	166.61	180.32	132.54
10	363.54	235.87	229.43	110.91	206.50	129.84
11	272.33	176.40	200.13	104.54	167.66	126.80
12	241.82	146.04	127.93	105.14	122.21	121.94
13	229.36	212.01	172.02	129.35	130.97	105.67
14	197.51	166.48	141.26	130.72	147.89	141.19
15	216.46	173.92	133.27	173.02	109.03	116.79
16	276.67	178.31	157.22	111.91	107.99	105.70
17	293.48	170.27	158.87	132.31	101.42	198.35
18	207.08	168.86	180.47	109.14	85.42	84.66
19	290.55	190.33	164.54	179.90	103.29	87.18
20	205.11	237.50	116.38	143.33	105.79	109.02



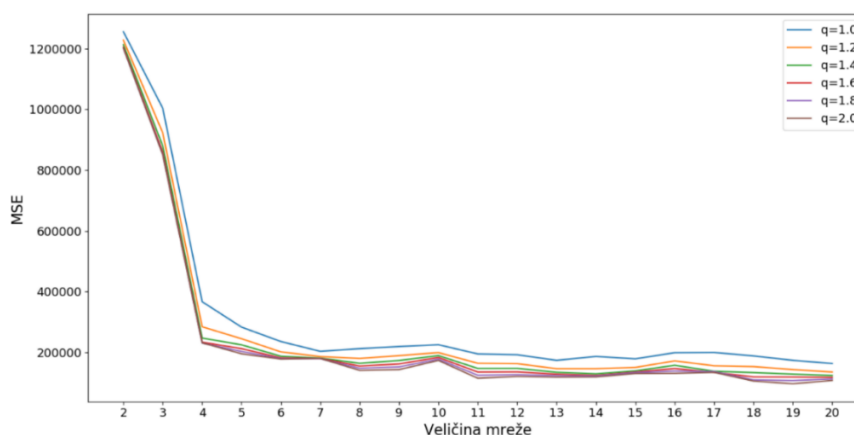
Sl. 5.11. Ovisnost MSE o veličini mreže na skupu REV i scenarij 2 za različite faktore skaliranja širina  $q$

Tablica 5.13 i graf na slici 5.12 pokazuju rezultate analize za skup RB u scenariju 1. Kao i ostali skupovi, općenito veća mreža znači manji MSE. Također, kao i kod skupova ASN, EE, i REV, povećanjem faktora  $q$  je moguće dodatno smanjiti vrijednost MSE. Najmanja vrijednost se dobiva za 19 skrivenih čvorova i  $q=2$  a najveća za jedan skriveni čvor i  $q=1$ . Funkcija dobrote

nije korištena za ovaj skup jer veliki “skok“ sa 18 na 19 skrivenih čvorova, to jest velika razlika u MSE, pa bi korištenje manje mreže značilo dosta veću pogrešku. Općenito se pokazalo za sve skupove podataka da se MSE smanjuje dodavanjem čvorova u skriveni sloj. Što se tiče širina radijalnih funkcija, na korištenim skupovima podataka se pokazala veća mogućnost generalizacije ako se poveća faktor  $q$ , no postoji gornja granica gdje daljnje povećavanje ne daje bolje rezultate, kako se pokazalo kod skupa EE.

Tab. 5.13. Ovisnost MSE o veličini mreže na skupu RB i scenarij 1 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	1255629	1227880	1212973	1204825	1201559	1202181
3	1003522	922574.1	879494.6	859767.4	852350.8	850084.9
4	367025.1	284502	247407.8	233849.5	231035.1	231685.5
5	283394.1	245176.3	225064.1	212532.2	203338.3	195630.9
6	235926.9	201817.5	187528.7	182039.9	179772.2	178130.7
7	203675.1	186505.1	181317.1	180435.2	179545.9	180321.9
8	212577.4	180178.5	164349.2	155118.9	147939.6	141013.3
9	219455	189541.3	173449.9	162162.5	152448.3	143541.7
10	225493.6	199463.2	189705.9	183559.9	177983.3	173752.2
11	195122.2	164401.2	147354.7	135510.4	124874.9	115061.7
12	192493.9	163163.4	147524.6	135922.5	126922.8	120945.2
13	173996	146436.5	134619.3	128492.6	123562.6	118738.7
14	186918.6	146533	129677.7	124197.7	122315.4	119375.7
15	178953.8	150655	139250.4	135562.7	133988.1	130749.7
16	199215.4	172242.2	157782	147009.6	138711.9	131246.8
17	199932	156098.8	138032	134564.3	137132.9	134529.1
18	188639.8	153510.1	133697.7	119411.9	109757.2	105910.8
19	174037.6	143537.3	128746	118992.8	107606.8	96896.46
20	163592.5	135508	124061	117772.4	112843.5	108100.9

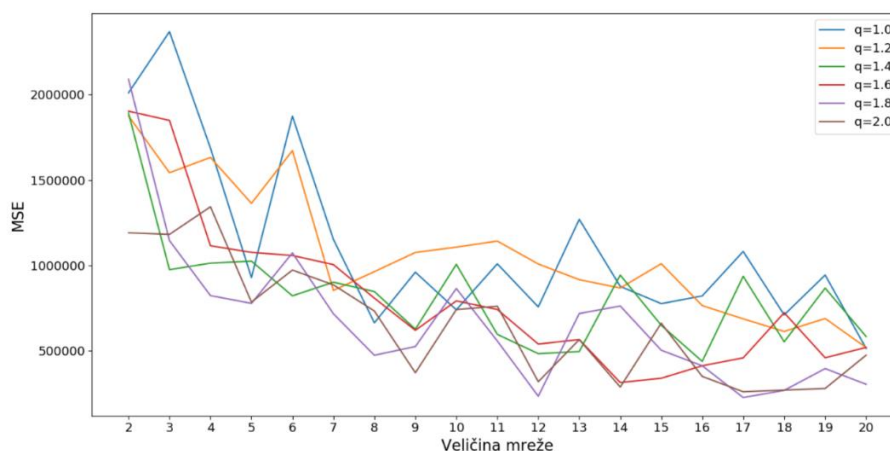


Sl. 5.12. Ovisnost MSE o veličini mreže na skupu RB i scenarij 1 za različite faktore skaliranja širina  $q$

Tablica 5.14 i graf na slici 5.13 prikazuju rezultate analize za skup RB u scenariju 2. Ponovno se, kao za ostale skupove javljaju velike varijacije, te je pogreška puno veća ako se ne koristi algoritam za grupiranje. Za sve skupove podataka (osim CCS) su dobiveni rezultati bolji ako se koristi *k-means*. Kako je već napomenuto, *k-means* može dati neoptimalne rezultate ako su podaci uniformno raspoređeni, no u slučaju kompleksnih oblika rasporeda podataka moguće je korištenje drugih algoritama za grupiranje koji nisu pokriveni ovim radom.

Tab. 5.14. Ovisnost MSE o veličini mreže na skupu RB i scenarij 2 za različite faktore skaliranja širina  $q$

Veličina mreže	$q$					
	1.0	1.2	1.4	1.6	1.8	2.0
2	2010328	1875959	1892147	1903330	2090829	1191525
3	2368896	1543205	975656.9	1849408	1145779	1182153
4	1687441	1632550	1013772	1115631	823789.5	1343638
5	929405.3	1363895	1025640	1076860	778406.5	785648.5
6	1874178	1672517	822436.3	1058480	1073749	973240.4
7	1153029	854187	902112.6	1005124	717421.1	886645.4
8	664358.7	964003.6	847365.9	808925.2	474534.2	733500.3
9	960767.9	1076275	627250.7	620656.6	525558.8	372030.5
10	740649.2	1107416	1006795	792991.5	865225.6	741992.7
11	1009642	1143047	596783.5	742609.2	558124.4	761954.4
12	757509	1008638	483896.9	540113.3	234601.5	319545
13	1270352	917078.6	496246.2	566591.6	718820.4	566376.4
14	877630.6	868110.3	943653.4	314940.9	762924.4	288388.7
15	776747	1010415	650232.1	339957.5	504037.5	662019.5
16	821890.9	765488.9	438814.4	412908	413020.6	350791.1
17	1082125	687656.5	936578.4	459350.3	227265.7	260967.2
18	710912.7	613792.8	553354.1	723517.9	268897.2	270757
19	944032.9	689695.9	867673.7	459931.6	396473.5	280080.6
20	514525.4	519546.9	584068.8	520444.7	304690.9	475137.7



Sl. 5.13. Ovisnost MSE o veličini mreže na skupu RB i scenarij 2 za različite faktore skaliranja širina  $q$

## 6. ZAKLJUČAK

U diplomskog radu opisan je problem regresije i neke moguće metode za rješavanje istog. Dan je opis koncepta i arhitekture radijalnih mreža i njihove primjene u regresijskoj analizi, te procesa izgradnje takve mreže. Opisano je nekoliko načina određivanja svakog od parametara korištenjem nadziranih i nenadziranih metoda učenja i postupak određivanja veličine mreže te vrednovanje kvalitete predviđanja. S ciljem izvršavanja regresijske analize napravljeno je programsko rješenje koje korisniku omogućuje učitavanje skupa podataka po izboru i konfiguraciju parametara radijalne mreže. Nadalje, po završetku analize, korisnik dobiva na uvid rezultate putem grafa, ispisa u konzolu i datoteke na disku. Korištenjem ostvarenog programskog rješenja izvršena je regresijska analiza na pet stvarnih skupova podataka, gdje su rezultati komentirani i obrazloženi. Dobiveni rezultati pokazuju uspješnost predviđanja u ovisnosti o odabranim parametrima i konfiguraciji mreže. U analizi korištenih skupova podataka, pokazala se potreba za korištenjem algoritma za grupiranje i skaliranja širina radijalnih funkcija. MSE se općenito smanjuje povećanjem faktora skaliranja širina, i u četiri od pet slučajeva je dobivena manja pogreška u slučaju da se koristi algoritam *k-means*. Daljnji rad može uključivati obradu većeg broja skupova podataka. Uz to, može se uvesti podržanost višestruke regresije, što se odnosi na korištenje više od jedne zavisne varijable. Dodatno, moguća je implementacija dodatnih mogućnosti konfiguracije mreže (primjerice drugačijih algoritama za grupiranje i određivanja širina radijalnih funkcija), te optimizacija procesa treninga u pogledu potrebnih računalnih resursa i vremena, iz razloga što to može biti problem na većim skupovima podataka. Nadalje, moguća je implementacija grafičkog korisničkog sučelja umjesto komandnog prozora kako bi se olakšalo rukovanje programskim rješenjem.

## Literatura

- [1] Y. Xin, G.S. Xiao, *Linear Regression Analysis: Theory and Computing*, World Scientific Publishing Co. Pte. Ltd, Singapur, 2009.
- [2] S. Chatterjee, A. S. Hadi, *Regression Analysis By Example, Fifth Edition*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2012.
- [3] A. Ruckstuhl, *Introduction to Nonlinear Regression*, IDP Institut für Datenanalyse und Prozessdesign, Švicarska, 2010.
- [4] D. S. Broomhead and D. Lowe, *Multivariable functional interpolation and adaptive networks*, Complex Systems, br. 3, sv. 2, str. 321-355, 1988.
- [5] A. P. Engelbrecht, *Computational Intelligence: An Introduction, Second Edition*, Wiley Publishing, UK, 2007.
- [6] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Inc., Birmingham, UK, 1995.
- [7] M. J. L. Orr, *Introduction to radial basis function networks*, tech. rep., University of Edinburgh, Centre for Cognitive Science, Edinburgh, UK, 1996.
- [8] Dokumentacija za scikit-learn, sklearn.model\_selection.Kfold, URL: [https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.KFold.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html), 03.08.2019
- [9] A. K. Jain, *Data clustering: 50 years beyond k-means*, Pattern Recognition Letters, sv. 31, str. 651-666, lipanj 2010.
- [10] S. Haykin, *Neural Networks A Comprehensive Foundation, Second Edition*, Prentice-Hall Inc, Indija, 1999.
- [11] J. Moody and C. J. Darken, *Learning with localized receptive fields*, Research Report, Yale Computer Science, New Haven, Connecticut, 1988.
- [12] A. Saha and J.D. Keeler, *Algorithms for Better Representation and Faster Learning in Radial Basis Function Networks*, Neural Information Processing Systems, Advances in Neural Information Processing Systems 2, str. 482–489, Denver, Colorado, 1989.
- [13] N. Benoudjit, C. Archambeau, A. Lendasse, J. Lee, M. Verleysen, *Width optimization of the Gaussian kernels in Radial Basis Function Networks*, ESANN 2002 proceedings – European Symposium on Artificial Neural Networks, str. 425-432. Bruges, Belgija, travanj, 2002.



- [14] R. Kruse, C. Borgelt, F. Klawonn, C. Moewes, M. Steinbrecher, and P. Held, *Computational Intelligence: A Methodological Introduction*. Springer-Verlag, London, UK, 2013.
- [15] F. S. Panchal et al, *Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network*, International Journal of Computer Science and Mobile Computing, sv. 3, br. 11, str. 455-464, studeni 2014.
- [16] Dokumentacija za scikit-learn, sklearn.cluster.Kmeans, URL: <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>, 04.08.2019.
- [17] Dokumentacija za scikit-learn, sklearn.neighbors.NearestNeighbors, URL: <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.NearestNeighbors.html#sklearn.neighbors.NearestNeighbors>, 03.08.2019.
- [18] UCI Machine Learning Repository, URL: <https://archive.ics.uci.edu/ml/datasets.php>, 04.08.2019.
- [19] D. Bajer, B. Zorić, and G. Martinović, *Automatic design of radial basis function networks through enhanced differential evolution*, Proceedings of the 10th International Conference on Hybrid Artificial Intelligent Systems (HAIS), str. 244-256, lipanj 2015.

## Sažetak

U diplomskom radu opisan je problem regresije, te pristup njegovom rješavanju korištenjem radijalnih mreža. Objašnjen je koncept i arhitektura radijalnih mreža i objašnjene su sličnosti i razlike s višeslojnim perceptronom te tijekom izgradnje koji se sastoji od određivanja parametara koristeći nadzirane i nenadzirane metode učenja. Navedeni su različiti načini određivanja svakog od parametara. Razvijeno je programsko rješenje koje omogućuje izvršavanje regresijske analize na danom skupu podataka tako da se učitava skup podataka, konfigurira radijalna mreža, te koje pruža na uvid rezultate. Izvršena je eksperimentalna analiza korištenjem razvijenog programskog rješenja za regresijsku analizu pet stvarnih skupa podataka čiji su rezultati prikazani i komentirani. Dobiveni rezultati pokazuju važnost korištenja algoritma za grupiranje i skaliranja širina radijalnih funkcija pri određivanju parametara radijalne mreže.

**Ključne riječi:** grupiranje podataka, parametri radijalne mreže, radijalna funkcija, radijalna neuronska mreža, regresija

## Abstract

### **Radial basis function network modelling for regression analysis**

Graduation thesis provides a description of regression analysis as well as an approach to regression analysis using radial basis function networks. The general concept and architecture of radial basis function networks are explained, and also the process of modelling such a network which is done by determining its parameters using supervised and unsupervised learning methods. Different approaches to determining the parameters are presented. A CLI application was developed to perform regression analysis, which enables loading of a dataset, network parameter configuration, and which provides results for inspection to the user. Experimental analysis was performed using the application for regression analysis of five real-world datasets, whose results were shown and commented. The results show the importance of clustering algorithm usage and radial basis function width scaling when determining network parameters.

**Keywords:** clustering, network parameters, radial basis function, radial basis function network, radial basis function, regression

## **Životopis**

Matej Đakovac rođen je 19. lipnja 1994.g. u Našicama. Živi u Osijeku te se u istom mjestu obrazuje. Pohađao je Osnovnu školu Ante Starčevića u Viljevu od 2001. do 2009. Zatim nastavlja svoje obrazovanje u Srednjoj školi Donji Miholjac smjera Opća gimnazija od 2009. do 2013. Završetkom srednje škole, uspješno upisuje preddiplomski studij na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija u Osijeku 2013. godine, gdje nakon tri godine završava preddiplomski sveučilišni studij računarstva.

Potpis:

---

## **Prilozi**

Na CD-u:

- Diplomski rad u .doc, .docx i PDF formatu
- Projekt programskog rješenja
- Korišteni skupovi podataka

GitHub repozitorij programskog rješenja: <https://github.com/mdakovac/rbfnn>