

Razvoj kalibracijskih modela za aktivne farmaceutske tvari

Herceg, Tesa

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Chemical Engineering and Technology / Sveučilište u Zagrebu, Fakultet kemijskog inženjerstva i tehnologije**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:149:646971>

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

Download date / Datum preuzimanja: **2024-07-13**



Repository / Repozitorij:

[Repository of Faculty of Chemical Engineering and Technology University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET KEMIJSKOG INŽENJERSTVA I TEHNOLOGIJE
SVEUČILIŠNI DIPLOMSKI STUDIJ

Tesa Herceg

DIPLOMSKI RAD

Zagreb, siječanj 2023.

SVEUČILIŠTE U ZAGREBU
FAKULTET KEMIJSKOG INŽENJERSTVA I TEHNOLOGIJE
SVEUČILIŠNI DIPLOMSKI STUDIJ

Tesa Herceg

**Razvoj kalibracijskih modela za aktivne farmaceutske
tvari**

DIPLOMSKI RAD

Voditelj rada: doc. dr. sc. Željka Ujević Andrijić

Članovi komisije:

doc. dr. sc. Željka Ujević Andrijić

prof. dr. sc. Nenad Bolf

prof. dr. sc. Jasna Prlić Kardum

Zamjena: izv. prof. dr. sc. Krunoslav Žižek

Zagreb, siječanj 2023.

Ovaj rad izrađen je na Fakultetu kemijskog inženjerstva i tehnologije, Sveučilišta u Zagrebu, Zavodu za mjerenje i automatsko vođenje procesa, akademske godine 2021./2022.

Zahvaljujem se svojoj mentorici doc. dr. sc. Željki Ujević Andrijić i svim asistentima na Zavodu za mjerenje i automatsko vođenje procesa na pomoći, razumijevanju i uloženom vremenu i trudu pri izvedbi i izradi ovog rada. Hvala Filipu na pomoći i strpljenju oko Pythona, hvala Marku na pomoći oko pisanja. Od srca zahvaljujem svojoj obitelji i prijateljima na podršci i razumijevanju tijekom studija. Za kraj se zahvaljujem svojim kolegama, zbog kojih je studiranje bilo jedno predivno i nezaboravno iskustvo.

SAŽETAK

Proces kristalizacije vrlo je važan proces u farmaceutskoj industriji. Njegova složenost i ovisnost o mnogim čimbenicima, kao i strogi zahtjevi za visokom kvalitetom i čistoćom farmaceutskih aktivnih tvari, stvorili su potrebu za razvojem naprednih instrumenata i tehnologija za kontinuirano praćenje i vođenje procesa. Procesna analitička tehnologija (PAT) sve se više primjenjuje u procesu kristalizacije za kontinuirano praćenje neke od ključnih značajki kvalitete proizvoda. PAT- tehnologija podrazumijeva kontinuirano praćenje koncentracije naprednim analitičkim instrumentima kao što su optičke sonde te razvoj i primjenu kalibracijskoga modela koji povezuje izmjerene spektralne podatke s koncentracijom farmaceutske aktivne tvari. Kalibracijski model moguće je razviti nekom od metoda strojnoga učenja.

Cilj ovog rada je razviti i ispitati modele neuronskih mreža za modeliranje ovisnosti koncentracije farmaceutske aktivne tvari, ksilometazolin hidroklorida, u n-butanolu, o temperaturi i spektralnim podacima dobivenih mjerenjima ATR-FTIR spektrometrom.

U radu je provedena predobrada spektralnih podataka, analiziran je broj neurona u prvom i drugom skrivenom sloju, broj skrivenih slojeva, vrsta primijenjenoga algoritma učenja te utjecaj vrste prijenosne funkcije na kvalitetu razvijenih neuronskih mreža.

Ključne riječi: kristalizacija, procesna analitička tehnologija, ATR-FTIR spektrometrija, strojno učenje, ksilometazolin

ABSTRACT

The crystallization process is a very important process in the pharmaceutical industry. The complexity of the crystallization process and the dependence of this process on many factors, as well as the strict requirements for high quality and purity of active pharmaceutical ingredients, have led to the need to develop advanced instruments and technologies for continuous monitoring and process control. Process analytical technology (PAT) is increasingly being applied in the crystallization process to continuously monitor some of the key product quality characteristics. The PAT technology involves continuous monitoring of the concentration with advanced analytical instruments such as optical probes and the development of a calibration model that relates the measured spectral data to the concentration of the pharmaceutically active substance. The calibration model can be developed using machine learning methods.

The aim of this work was to develop and investigate neural network models for modeling the dependence of the concentration of the active pharmaceutical ingredient, xylometazoline hydrochloride in n-butanol on temperature and spectral data obtained by ATR-FTIR spectrometer measurements.

The effects of the preprocessing of the spectral data, the number of neurons in the first and second hidden layers, the number of hidden layers, and the type of optimizer and transfer function used on the quality of the evolved neural networks were analyzed.

Keywords: crystallization, process analytical technology, ATR-FTIR spectrometry, machine learning, xylometazoline

SADRŽAJ

1. UVOD.....	1
2. OPĆI DIO	2
2.1. Kristalizacija	2
2.1.1. Kristalizacija iz otopine	2
2.1.2. Topljivost	3
2.1.3. Metastabilna zona	4
2.1.4. Nukleacija i rast kristala.....	5
2.2. Procesne analitičke tehnologije.....	5
2.2.1. ATR-FTIR spektrometrija	7
2.3. Strojno učenje	8
2.3.1. Razvoj modela strojnog učenja.....	9
2.3.2. Prikupljanje i predobrada podataka	10
2.3.3. Podjela podataka	11
2.3.4. Odabir algoritma strojnog učenja.....	11
2.3.5. Odabir značajki iz predobrađenih podataka.....	12
2.4. Umjetne neuronske mreže.....	12
2.4.1. Klasifikacija neuronskih mreža.....	14
2.4.2. Struktura umjetne neuronske mreže.....	14
2.4.3. Učenje neuronske mreže	15
2.4.4. Koeficijent učenja	16
2.4.5. Prijenosna funkcija.....	16
2.4.6. Pretreniranost neuronske mreže	17
2.4.7. Ocjena izvedbe neuronske mreže.....	18
3. EKSPERIMENTALNI DIO	20
3.1. Materijali.....	20
3.2. Uređaji.....	20
3.3. Postupak rada	21
3.3.1. Određivanje krivulje topljivosti	21
3.3.2. Priprema otopina i prikupljanje temperatura	22
3.3.3. Prikupljanje i predobrada spektralnih podataka.....	23
3.3.4. Razvoj kalibracijskog modela.....	23
3.3.6. Optimizacija modela neuronske mreže	25
4. REZULTATI I RASPRAVA	27
4.1. Rezultati za skup podataka D1	27
4.1.1. Ispitivanje broja neurona za pojedine aktivacijske funkcije	27
4.1.2. Ispitivanje utjecaja vrste optimizatora	29

4.1.3. Ispitivanje utjecaja broja epoha i <i>batch size</i> -a	30
4.1.4. Najbolji model za skup podataka D1	31
4.2. Rezultati za skup podataka D2.....	32
4.2.1. Ispitivanje broja neurona za pojedine aktivacijske funkcije	33
4.2.2. Ispitivanje utjecaja vrste optimizatora	35
4.2.3. Ispitivanje utjecaja broja epoha i <i>batch size</i> -a	36
4.2.4. Najbolji model za skup podataka D2	37
4.3. Rezultati za skup podataka D3.....	38
4.3.1. Najbolji model za skup podataka D3	38
4.4. Rezultati za skup podataka D4.....	40
4.4.1. Najbolji model za skup podataka D4	40
4.5. Usporedba najboljih modela	41
5. ZAKLJUČAK	43
6. LITERATURA	44
7. POPIS SIMBOLA.....	46
8. ŽIVOTOPIS.....	47

1. UVOD

Kristalizacija je važan proces u mnogim industrijskim granama (farmaceutskoj, kemijskoj i prehrambenoj) [1]. U farmaceutskoj industriji proces kristalizacije od iznimnoga je značaja prilikom pripreme visokokvalitetnih i čistih proizvoda. Stoga je cilj farmaceutske industrije unaprijediti i automatizirati proces kristalizacije radi bržega, jeftinijeg i sigurnijeg dobivanja istih visokokvalitetnih i čistih proizvoda. Za navedena poboljšanja potrebne su nove tehnologije, a među njima se sve više koristi procesna analitička tehnologija (eng. *Process Analytical Tehnology*, PAT). Njezinom upotrebom moguće je projektiranje, praćenje i vođenje proizvodnih procesa kontinuiranim mjerenjem ključnih pokazatelja kvalitete i rada procesa. Implementacija procesne analitičke tehnologije zahtjeva specijalizirane alate što podrazumijeva analitičke instrumente, alate za kemometrijsko modeliranje, softver za razvoj PAT-metode, obradu podataka i integraciju u sustav te softver i algoritme za vođenje procesa [2].

Ključni pokazatelji kvalitete u procesu kristalizacije jesu temperatura i koncentracija otopine. Kontinuirano mjerenje koncentracije moguće je uz napredne analitičke instrumente koji pomoću svjetlosnoga spektra kvantitativno i/ili kvalitativno karakteriziraju određeno svojstvo. U takve metode spada i prigušena totalna refleksija *Fourier* transformacijske infracrvene spektroskopije (eng. *Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy*, ATR-FTIR), korištena u ovome radu [3].

Kako bi se ATR-FTIR spektrometrija mogla koristiti za mjerenje koncentracije u otopini, ključni je korak izrada kalibracijskoga (kemometrijskog) modela koji zahtjeva dizajn eksperimenata (DoE) za prikupljanje ATR-FTIR spektralnih podataka potrebnih za razvoj kalibracijskoga modela [4]. U zadnje je vrijeme popularan razvoj empirijskih modela tehnikama strojnog učenja (eng. *Machine Learning*) u koje se ubrajaju i umjetne neuronske mreže (eng. *Artificial Neural Networks*) korištene u ovome radu. Ovakav način izrade modela uvelike ovisi o broju i vrsti eksperimentalnih podataka koji se modeliraju.

U ovome je radu primjenom umjetnih neuronskih mreža razvijen kalibracijski model za kontinuirano praćenje koncentracije ksilometazolin hidroklorida u procesu kristalizacije.

2. OPĆI DIO

2.1. Kristalizacija

Kristalizacija je toplinski separacijski proces pri kojemu dolazi do izdvajanja nove, čvrste faze (kristala) iz otopine, taline ili iz parne faze. S obzirom na to da se radi o procesu prijenosa tvari, potrebna je pokretačka sila – prezasićenost. Pokretačka sila procesa jest razlika koncentracija, odnosno odstupanje od ravnoteže, što se ostvaruje hlađenjem otopine ili isparavanjem te kemijskom reakcijom, dodatkom aditiva uz isoljavanje ili taloženje. Da bi došlo do procesa kristalizacije, moraju postojati centri kristalizacije, odnosno nukleusi.

Proces kristalizacije iznimno je složen te ovisi o većemu broju čimbenika. Za opis takvoga procesa potrebno je poznavati ravnotežu zadanoga sustava, kao i kinetiku procesa kako bi se dobili kristale odgovarajuće čistoće i granulometrijskih svojstava (oblik, veličina, raspodjela veličina kristala). Kristalizacija je od iznimnoga značaja u mnogim industrijskim granama, od kojih su najznačajnije farmaceutska, prehrambena i kemijska industrija.

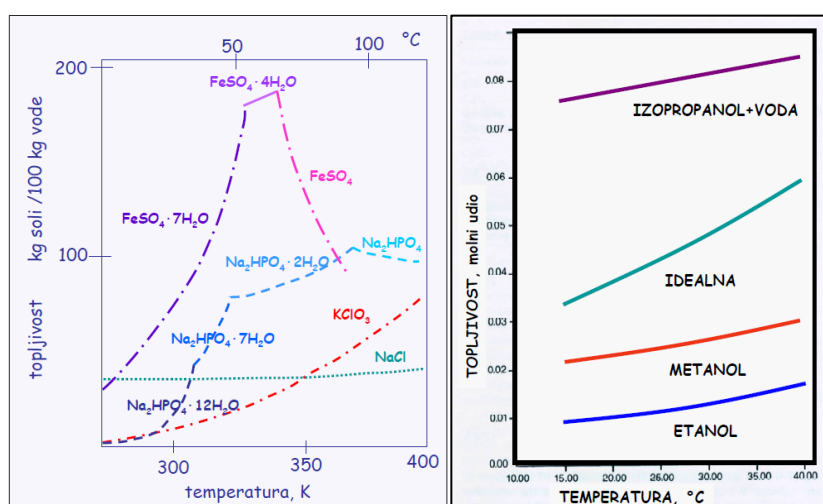
2.1.1. Kristalizacija iz otopine

Otopina je homogena smjesa dvaju ili više tvari koja se sastoji od otapala i otopljene tvari. Pri odabiru otapala nužno je obratiti pažnju na sljedeća svojstva: sposobnost otapala da otopi određenu tvar (tvar mora biti topljiva u odabranome otapalu), čistoća (nečistoće uvelike utječu na cjelokupni proces kristalizacije), kemijska reaktivnost (otapalo ne smije reagirati s otopljenom komponentom osim ako se radi s dodatkom aditiva) te rad s otapalima (djelotvornost, mogućnost regeneracije, toksičnost, zapaljivost).

Otopljeni komponenta prenosi se tijekom kristalizacije mehanizmom difuzije u smjeru manje koncentracije, dok se otapalo prenosi u obrnutome smjeru. Brzina prijenosa tvari difuzijom ovisi o temperaturi i koncentraciji otopine. Kako bi prijenos tvari bio djelotvoran, potrebno je osigurati dobro miješanje kako bi proces bio pod kontrolom ugradnje u kristalnu rešetku, a ne difuzijom. Naime, pri stvaranju kristala postoji više otpora. To su otpor prijenosu tvari difuzijom kroz otopinu do površine kristala te otpor prijenosu tvari na površini kristala pri ugradnji u kristalnu strukturu. Kada su hidrodinamički uvjeti povoljni (što se osigurava dobrim miješanjem), otpor prijenosu tvari difuzijom kroz otopinu je smanjen, a na brzinu ugradnje u kristalnu rešetku ne može se utjecati, stoga je to najsporiji stupanj koji ograničava brzinu rasta kristala.

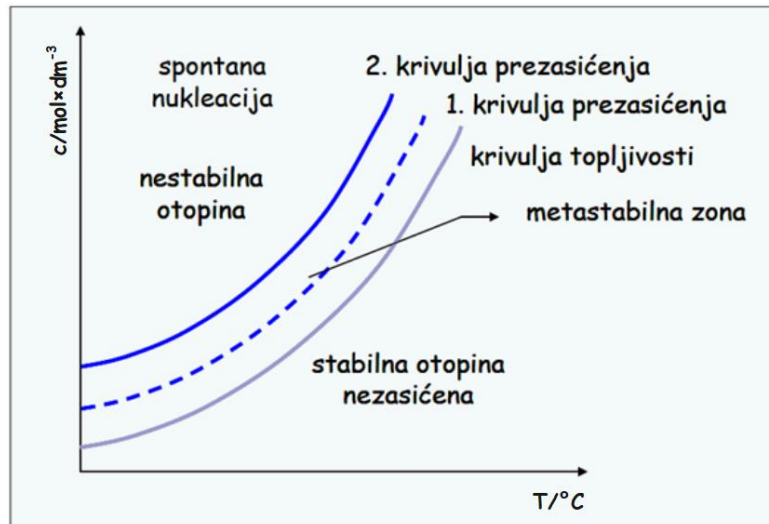
2.1.2. Topljivost

Pri proučavanju procesa kristalizacije, važno je poznavati topljivost pojedine komponente. Svaka pojedina tvar pri određenoj temperaturi ima određenu maksimalnu količinu koja se može otopiti u određenoj količini otapala. Stanje zasićenosti postiže se kada daljnjim dodatkom tvari ne dolazi do povećanja koncentracije u otopini. Kod većine krutih tvari topljivost raste s porastom temperature. Ukoliko tvar kristalizira tvoreći hidrate, na dijagramu topljivosti može se očitati u kojemu temperaturnom intervalu egzistira pojedini hidrat te prema tome odabrati radne uvjete. Osim temperature, na topljivost tvari utječu vrsta otapala, aditivi, pH otopine te prisutnost nečistoća. Primjeri utjecaja temperature i vrste otapala na topljivost soli prikazani su na *Slici 2.1*.



Slika 2.1 Utjecaj temperature i vrste otapala na topljivost soli [1]

Dijagram topljivosti prikazuje ovisnost koncentracije otopljene tvari o temperaturi (*Slika 2.2*). Krivulja topljivosti i krivulja prezasićenja dijele dijagram na tri dijela. Ispod krivulje topljivosti nalazi se područje stabilne nezasićene otopine u kojemu ne dolazi do kristalizacije. Iznad krivulje prezasićenja nalazi se nestabilno područje u kojemu dolazi do spontane nukleacije jer je osigurana dovoljno velika pokretačka sila. Između tih dviju krivulja nalazi se metastabilno područje u kojemu i najmanja promjena koncentracije omogućuje kristalizaciju. Drugim riječima, u metastabilnoj zoni ne dolazi do spontane nukleacije već će se ona dogoditi samo uz dodatak cjepiva (male količina kristala) ili ulaskom nečistoće u sustav koja će izazvati nukleaciju. Dakle, do kristalizacije dolazi samo ako je koncentracija otopine veća od ravnotežne. Što je veći stupanj prezasićenosti, veća je pokretačka sila.



Slika 2.2. Krivulja topljivosti [1]

Prezasićenost otopine izražava se sljedećim izrazima:

$$\text{Prezasićenost: } \Delta c = c - c^* \quad (1)$$

$$\text{Omjer prezasićenosti: } S = \frac{c}{c^*} \quad (2)$$

$$\text{Relativna prezasićenost: } \sigma = \frac{\Delta c}{c^*} = S - 1 \quad (3)$$

2.1.3. Metastabilna zona

Metastabilna zona predstavlja područje u dijagramu topljivosti u kojemu je otopina u stanju labilne ravnoteže. Do spontane nukleacije dolazi nakon dovoljno dugoga zadržavanja u tome području. Širina metastabilne zone ovisi o vremenu, brzinama hlađenja ili isparavanja, intenzitetu miješanja te prisutnosti nečistoća. Porastom brzine hlađenja raste i širina metastabilne zone stoga je za dobivanje većeg broj sitnih kristala poželjna i veća brzina hlađenja. Širina metastabilne zone smanjuje se porastom brzine vrtnje miješala zbog jačih smičnih naprežanja kapljevine koja utječu na dovoljnu pokretačku silu za nukleaciju te zbog povoljnijih hidrodinamičkih uvjeta za prijenos topline i tvari, no samo do određene, optimalne brzine vrtnje pri kojoj su otpori prijenosu tvari minimalni. Daljnjim povećanjem intenziteta miješanja ne postiže se nikakvo poboljšanje hidrodinamičkih uvjeta. Za dobivanje kristala uže raspodjele veličina kristala, kristalizacija se provodi uz dodatak cjepiva, a prezasićenost se održava unutar metastabilne zone.

2.1.4. Nukleacija i rast kristala

Stabilni nukleusi, odnosno centri kristalizacije, nužni su za rast kristala. Nukleacija se dijeli na primarnu i sekundarnu, homogenu i heterogenu. Heterogena nukleacija potaknuta je nečistoćama ili hrapavim površinama u reaktoru te nastaje pri manjim stupnjevima prezasićenosti od primarne nukleacije. Sekundarni nukleusi nastaju cijepljenjem (dodatkom sitnih kristala (cjepiva) u prezasićenu otopinu), sudarima (kontaktna nukleacija) i miješanjem (zbog smičnih naprezanja fluida u prisustvu kristala većih od kritičnoga nukleusa). Sekundarna nukleacija poželjna je kod kontinuiranih kristalizatora u kojima se nastali kristali kontinuirano odvođe, međutim kod nje je česta pojava nastanka sitnijih kristala te višemodalna raspodjela veličina. To se događa jer je brzina rasta kristala manja kako se koncentracija (pokretačka sila) troši i na nastajanje novih nukleusa. Sekundarna nukleacija može se kontrolirati manjim brzinama vrtnje miješala ili otapanjem kristala manjih od definirane veličine. Primarna se nukleacija također ne može kontrolirati jer se ne javlja uvijek pri istome stupnju prezasićenosti.

Cilj je cijepljenja izbjeći primarnu te minimizirati sekundarnu nukleaciju i omogućiti „trošenje“ prezasićenosti samo na rast kristala. Ako se pokretačka sila održava stalnom i brzina rasta kristala bit će konstantna. Masa i veličina cjepiva, kao i način i mjesto uvođenja u kristalizator, utječu na dobivanje željene veličine kristala. Kako bi se spriječilo dobivanje bimodalne raspodjele veličina kristala, ukupna površina cjepiva mora biti dovoljno velika. Osim toga, kako bi se osigurala željena veličina kristala cjepivo mora biti dovoljno maloga promjera i ne smije se dodati u prevelikome broju. Prevelik promjer cjepiva rezultira manjom količinom kristala, dok prevelika količina cjepiva rezultira kristalima malih dimenzija jer se pokretačka sila troši na velik broj cjepiva [1].

2.2. Procesne analitičke tehnologije

Povijesno gledano, farmaceutska industrija oslanjala se na šaržne procese proizvodnje, nakon kojih su slijedila laboratorijska ispitivanja i analize uzorka kako bi se ispitala kvaliteta dobivenoga proizvoda. Današnje procesno inženjerstvo i moderna proizvodnja napuštaju taj princip rada te se okreću ka *on-line* praćenju i vođenju procesa te tako postižu visoku kvalitetu željenoga proizvoda. Taj način praćenja i vođenja procesa omogućuje bržu i kvalitetniju proizvodnju, što u konačnici omogućuje farmaceutskoj industriji brže plasiranje proizvoda na tržište te manju vjerojatnost pojave pogreške [4].

Američka agencija za hranu i lijekove (eng. *Food and Drug Administration*, FDA) definirala je procesnu analitičku tehnologiju (eng. *Process Analytical Technology*, PAT) kao alate koji „omogućuju razumijevanje procesa za znanstveni i farmaceutski razvoj, proizvodnju i osiguravanje kvalitete“ [5]. Ti alati omogućuju projektiranje, praćenje i vođenje proizvodnih procesa kontinuiranim mjerenjem ključnih pokazatelja kvalitete i rada procesa. Mjeriti se mogu svojstva sirovina, međuprodukata i produkata, ali najčešće se mjere ključne veličine koje utječu na djelotvornost procesa i kvalitetu produkta [6]. Cilj korištenja procesne analitičke tehnologije jest bolje razumijevanje procesa u ranoj laboratorijskoj fazi te praćenje tijekom razvoja i uvećanja. Procesne analitičke tehnologije iznimno su važne u industriji jer omogućuju identifikaciju parametara koji utječu na pojedini proces, praćenje procesa u realnome vremenu, dobivanje podataka koji se mogu obraditi multivarijantnim metodama analize te razvoj strategije vođenja procesa [2].

Nakon definiranja kritičnih značajki kvalitete, QbD pristupom (eng. *Quality by Design*), pronalazi se funkcijska veza s kritičnim procesnim parametrima (eng. *Critical Process Parameters*, CPP). Nizom eksperimenata razvija se viševeličinski model koji definira način na koji promjene važnih varijabli i parametara u procesu utječu na kritične značajke kvalitete. Također se određuju granice prostora dizajna (eng. *Process Design Space*) u kojemu se dobiva produkt zadovoljavajućih karakteristika. Prostor dizajna opisuje viševeličinski model koji definira radne granice procesa. Kad se definiraju funkcijske veze kritičnih procesnih parametara i kritičnih značajki kvalitete te prostor dizajna, primjenom PAT-mjerenja i vođenja procesa moguće je ostvariti traženu kvalitetu i povećati djelotvornost procesa.

Kako bi se pomoću PAT-a mogao kontinuirano pratiti proces i na temelju toga dobiti njegovo optimalno stanje, potrebni su senzori koji skupljaju nužne informacije o procesu koji se prati. Postoji više načina mjerenja procesnih veličina:

- *off-line* – analitički laboratorij izvan proizvodnoga pogona;
- *at-line* – analitički laboratorij unutar proizvodnoga pogona;
- *on-line* – mjerenje u preusmjerenome toku uzorka koji se može vratiti u proces;
- *in-line* – mjerenje u stvarnome vremenu integrirano u proces bez uzorkovanja, ali proces se pritom može omesti;
- *non-invasive* – integrirano mjerenje u stvarnome vremenu, pri čemu senzor nije u kontaktu s tvari kojom se mjeri i pri čemu se ne ometa procesni tok.

Primjena procesne analitičke tehnologije zahtjeva specijalizirane alate za primjenu u farmaceutskoj industriji.

Sustav za mjerenje i vođenje koji uključuje PAT-tehnike sastoji se od slijedećih dijelova:

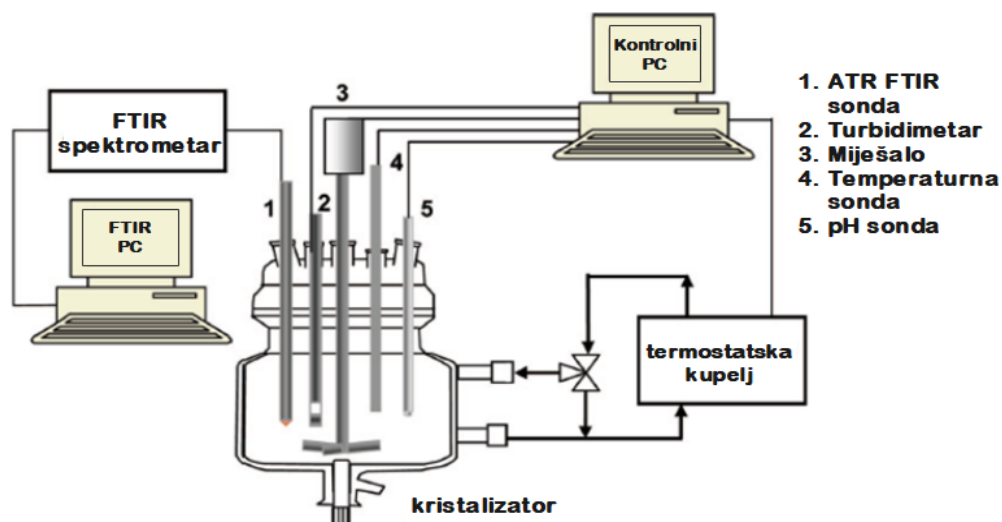
- analitički instrumenti,
- alati za kemometrijsko modeliranje,
- softveri za razvoj PAT-metode, obradu podataka i integraciju u sustav,
- softver i algoritmi za vođenje procesa.

Za mjerenje se obično koriste napredni analitički instrumenti koji rade na načelu da pomoću svjetlosnoga spektra kvantitativno i/ili kvalitativno karakteriziraju određeno svojstvo (npr. NIR, FTIR, Raman, masena i UV-Vis spektroskopija, FBRM) [6].

2.2.1. ATR-FTIR spektrometrija

Brzi razvoj spektrometrijske tehnologije pratio je i razvoj odgovarajućih instrumenta za *in-situ/online* mjerenja u svrhu praćenja procesa kristalizacije. Jedna od takvih spektrometrijskih tehnologija jest i prigušena totalna refleksija *Fourier* transformacijske infracrvene spektroskopije (eng. *Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy*, ATR-FTIR). FTIR je tehnika koja se koristi za dobivanje infracrvenoga spektra apsorpcije ili emisije krutine, tekućine ili plina. FTIR-spektrometar može prikupljati spektralne podatke visoke razlučivosti u širokome rasponu spektra. ATR jest jedan od dodataka na FTIR spektrometar, a koristi se za dobivanje infracrvenog spektra površine tvari te može dati valjane spektre uzoraka koji su veće debljine ili prejako apsorbiraju zračenje [3].

Kako bi se ATR-FTIR spektrometrija mogla koristiti za mjerenje koncentracije u otopini, ključni je korak izrada kalibracijskoga (kemometrijskog) modela. Prije eksperimentiranja poželjno je provesti dizajn eksperimenata (DoE) za prikupljanje ATR-FTIR spektralnih podataka za razvoj kalibracijskog modela [7].



Slika 2.3. Shema praćenja procesa kristalizacije primjenom procesne analitičke tehnologije

[8]

2.3. Strojno učenje

Umjetna inteligencija (eng. *Artificial Intelligence*) dio je znanosti o računalima koji se bavi projektiranjem inteligentnih računalnih sustava odnosno sustavima koji predočavaju karakteristike koje povezujemo s inteligencijom u ljudskome ponašanju [9]. Strojno učenje (eng. *Machine Learning*) grana je umjetne inteligencije koja omogućuje računalima da „uče“ iz podataka bez eksplicitnoga programiranja [10]. Algoritmi strojnoga učenja „uče“ informacije i odnose među njima izravno iz podataka, ne oslanjajući se na fundamentalne zakonitosti. Pri tome se modeli unaprjeđuju porastom broja uzoraka dostupnih za učenje. Dakle, glavna zadaća strojnoga učenja jest pronalazak prirodnih uzroka i poveznica u podacima te na temelju njih stjecanje uvida u proces i zatim odlučivanje i predviđanje.

Strojno učenje primjenjuje se prilikom rješavanja kompleksnih zadataka ili problema koji uključuju veliku količinu podataka i velik broj varijabli, pri čemu ne postoje unaprijed definirane formule ili jednadžbe (modeli). Ono je aktualna tema u istraživanju i industriji gdje ima važnu ulogu u prediktivnom održavanju uređaja i procesne opreme, procjeni i nadziranju kvalitete proizvoda, ali i kada je priroda podataka takva da se stalno mijenja što pak zahtijeva čestu prilagodbu računalnoga programa. Zbog velikih količina podataka koristi se i u ostalim područjima (proračun financija, računalna obrada slika, računalna biologija, obrada prirodnog jezika,...).

Razlikuju se dvije osnovne tehnike strojnoga učenja: nadzirano učenje (eng. *Supervised Learning*) i nenadzirano učenje (eng. *Unsupervised Learning*).

Kod nadziranoga učenja modeli se izvode na temelju skupa ulaznih i izlaznih podataka prikupljenih eksperimentom ili akvizicijom tako da mogu predviđati buduće izlaze na temelju dostupnih ulaznih podataka. Dakle, algoritmi rade s poznatim skupom ulaznih (ulazi modela) i izlaznih (izlaz modela) podataka te uvježbavaju model za predviđanje izlaznih podataka.

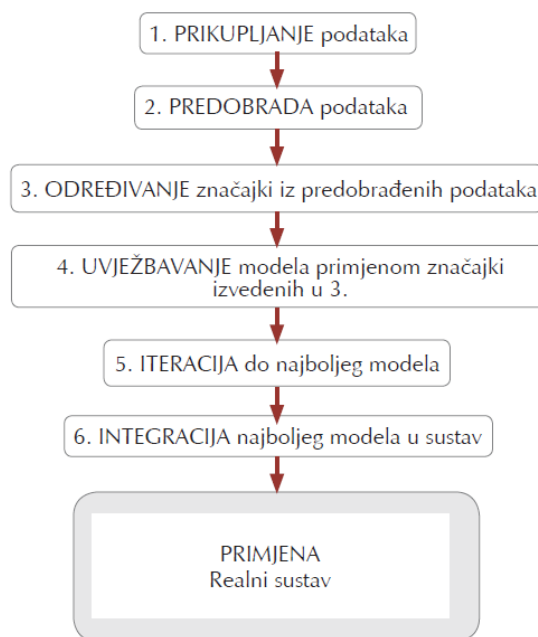
Nadzirano učenje primjenjuje se u sljedećim postupcima:

- Klasifikacija ili razvrstavanje – koristi se za predviđanje diskretnih odziva, pri čemu se ulazni podaci razvrstavaju u kategorije. Primjeri: razvrstavanje željene i neželjene e-pošte, u medicinskoj dijagnostici, npr. je li tumor dobroćudan ili zloćudan.
- Regresija – predviđa kontinuirane odzive. Ovim postupkom predviđaju se kontinuirane varijable poput promjene temperature ili koncentracije u reaktoru, sastav i kvaliteta proizvoda, fluktuacije potrošnje električne energije u elektroenergetskom sustavu i sl.

Nenadzirano učenje (eng. *Unsupervised Learning*) radi na principu pronalaska skrivenih uzoraka i inherentne strukture u ulaznim podacima bez poznavanja izlaza. Njime se uzorci grupiraju ili otkrivaju određene strukture uzoraka. Ono se primjenjuje za izvođenje zaključaka iz skupova podataka koje čine samo ulazni podaci, bez poznavanja odzivnih podataka (izlaza) [11].

2.3.1. Razvoj modela strojnog učenja

Jedan od izazova pri razvoju modela strojnoga učenja jest baratanje podacima. Datoteke s podacima javljaju se u različitim oblicima i veličinama pa i različitim vrstama (slike, tekst, signali sa senzora), stoga predobrada podataka može biti ključna u razvoju modela. Na *Slici 2.4.* prikazan je tijek razvoja modela strojnoga učenja.



Slika 2.4. Koraci u razvoju modela strojnog učenja [11]

2.3.2. Prikupljanje i predobrada podataka

Odabir podataka iz baze podataka postrojenja ili laboratorija te njihova predobrada mogu oduzeti preko 50 % ukupnoga vremena razvoja modela. Sirovi podaci često nisu prikladni za direktnu primjenu u razvoju modela te ih je potrebno predobraditi i prilagoditi u oblik pogodan za proračun. Predobrada podataka podrazumijeva pronalaženje ekstremnih vrijednosti (eng. *outliers*), njihovo uklanjanje, zamjenu nedostajućih vrijednosti, filtriranje podataka, generiranje dodatnih vrijednosti varijabli, odabir utjecajnih varijabli, detrendiranje te skaliranje podataka.

Ekstremne vrijednosti značajno odstupaju od okolnih podataka te mogu negativno utjecati na razvoj modela. To su najčešće mjerne pogreške koje ne treba uzimati u obzir pri razvoju modela jer bi dale krivu predodžbu procesa.

Filtriranje je postupak kojim se nastoji smanjiti mjerni šum prisutan kod mjernih pretvornika. Svrha filtriranja je „izgladiti“ podatke te olakšati razvoj modela i ukloniti visokofrekventne i niskofrekventne smetnje jer one mogu dovesti do netočnoga i nestabilnog modela. Neke od metoda filtriranja i zaglađivanja podataka su: *Moving average*, *LOWESS* i *LOESS* filtri, *Savitzky-Golay* filter te *Multiplicative Scatter Correction* (MSC) filter.

Skaliranje podataka provodi se kako bi se smanjio utjecaj brojčano većih veličina nad manjim tijekom postupka razvoja modela. Podaci, ovisno o fizikalnim jedinicama i prirodi procesa,

mogu imati različite iznose, stoga je bitno svesti ih na isto mjerilo, tj. normalizirati. Razlikuju se dvije metode skaliranja podataka:

- *Min-max* normalizacija

$$x_{norm}^i = \frac{x^i - x_{min}}{x_{max} - x_{min}} \quad (4)$$

pri čemu je x neskaliрана varijabla, a x_{norm} skalirana varijabla.

- *Z-score* normalizacija (standardizacija) [15]

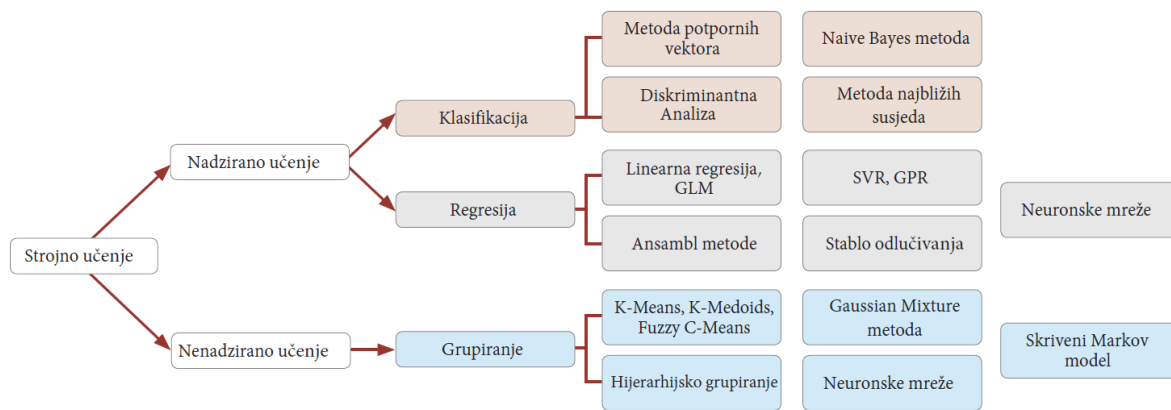
$$x_{std}^i = \frac{x^i - \mu_x}{\sigma_x} \quad (5)$$

2.3.3. Podjela podataka

Prije samoga procesa učenja kod razvoja modela primjenom strojnog učenja, podatke je potrebno podijeliti na skup za učenje modela te skup za provjeru modela. Skup podataka za učenje trebao bi sadržavati kompletan raspon mogućih vrijednosti podataka kako bi dobivenim modelom mogao predviđati njihovo vladanje. Skup za provjeru (validaciju, vrednovanje) modela sadrži podatke nepoznate dobivenome modelu te se njima provjerava kako se model vlada na nepoznatome skupu podataka. Uz njegovu se pomoć ocjenjuju generalizacijska svojstva modela. Omjer u kojemu se podaci dijele treba pažljivo odabrati jer se smanjenjem broja podataka u skupu za učenje mogu uskratiti informacije važne prilikom učenja, a smanjenjem broja podataka u skupu za provjeru procjena pogreške generalizacije postaje nepouzdanom. Skup za učenje treba biti uvijek veći od skupa za provjeru. U praksi se najčešće uzimaju omjeri 60:40, 70:30 ili 80:20, ovisno o količini podataka. Za jako velike količine podataka prihvatljivi i uobičajeni omjeri su 90:10 ili čak 99:1 [12]. Za konačnu provjeru modela može se koristiti i treći skup podataka – skup za testiranje [9].

2.3.4. Odabir algoritma strojnog učenja

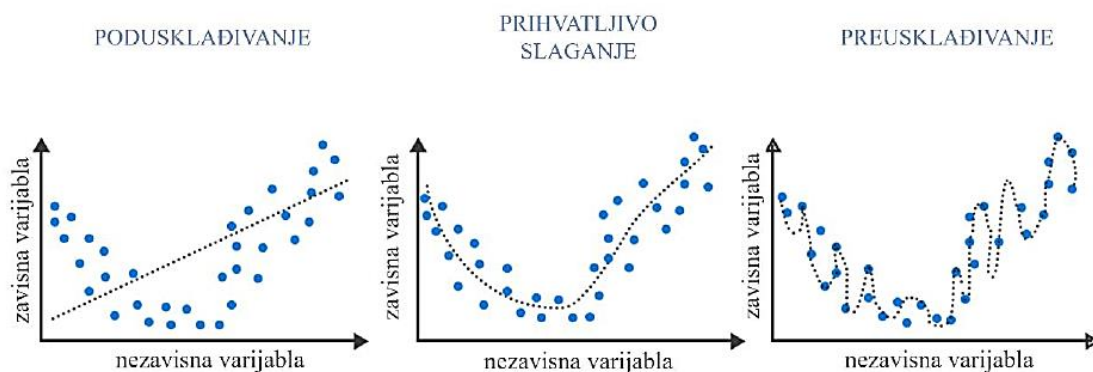
Odabir algoritma strojnog učenja ovisi o broju i vrsti podataka, svrsi razvoja modela te o tome gdje i kako će se rezultati primjenjivati. Postoji više desetaka vrsta algoritama, a na *Slici* 2.5 prikazane su neke od tehnika provedbe strojnog učenja [11].



Slika 2.5. Podjela metoda strojnog učenja [11]

2.3.5. Odabir značajki iz predobrađenih podataka

Ako vrijednost kriterija uspješnosti nije zadovoljavajuća na nekome od skupova podataka, može se zaključiti kako model ima problema s generalizacijom. Razlog tomu može biti nedovoljna kompleksnost modela, tzv. „podusklađivanje“ (eng. *Underfitting*) ili pak prevelika kompleksnost modela, tzv. „preusklađivanje“ (eng. *Overfitting*). Na Slici 2.5. prikazani su primjeri podusklađivanja, preusklađivanja te prihvatljivoga slaganja modela s eksperimentalnim podacima. Navedeni problemi mogu se riješiti povećavanjem podataka za učenje, penaliziranjem (kažnjavanjem) složenosti modela postupkom regularizacije, odabirom složenijeg/jednostavnijega modela s više/manje parametara ili smanjivanjem dimenzionalnosti podataka [12].

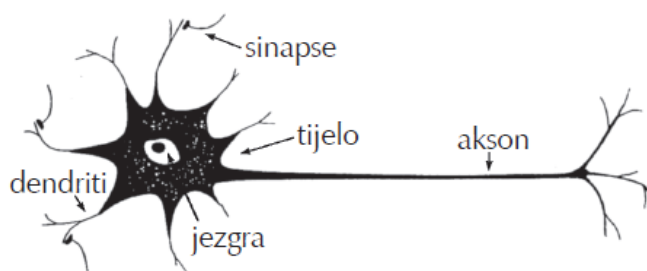


Slika 2.6. Podusklađivanje, prihvatljivo slaganje i preusklađivanje modela s eksperimentalnim podacima [12]

2.4. Umjetne neuronske mreže

Umjetna neuronska mreža skup je međusobno povezanih jednostavnih procesnih elemenata (neurona) čija se funkcionalnost temelji na biološkome neuronu te služi distribuiranoj

paralelnoj obradi podataka. Svaki biološki neuron, kao što je prikazano na *Slici 2.7.* sastoji se od sljedećih dijelova:

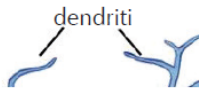
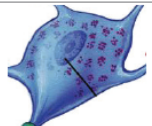
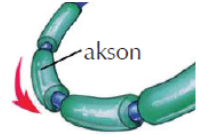



Slika 2.7. Prikaz biološkog neurona [13]

- tijela stanice – sadrži jezgru ili nukleus s informacijama o nasljednim značajkama
- dendrita – kraće niti oko stanice, uloga im je prenošenje signala s drugih neurona
- aksona – duge tanke niti koje prenose signale do drugih neurona pri čemu se granaju u vlakna
- sinapse – funkcionalne jedinice između završetka aksona prethodnog neurona i dendrita sljedećeg neurona, oslobađaju materijal potreban stanici za prijenos signala, neurotransmiter, pri čemu se odvija elektrokemijska reakcija.

Spoznaja o građi i načinu funkcioniranja ljudskog mozga i neurona potaknula je istraživanje i razvoj umjetnih neuronskih mreža. Topološka analogija umjetnih neuronskih mreža s biološkim prikazana je u *Tablici 2.1.* [13].

Tablica 2.1. Usporedba biološkog i umjetnog neurona [13]

	Biološki neuron	Umjetni neuron
	Prima ulazni signal putem dendrida (sinaptičke veze)	Prima ulaze (i) koji su određeni težinskim koeficijentima (w)
	Obrada signala u somi	Obrada ulaza, unutarnji prag – bias (b)
	Pretvara obrađeni ulaz u izlaz putem aksona	Pretvara ulaze u izlaz (prijenosna funkcija)
	Šalje informacije putem sinapsi do svih neurona s kojima je neuron povezan	Šalje informaciju prema izlazu i sljedećim neuronima

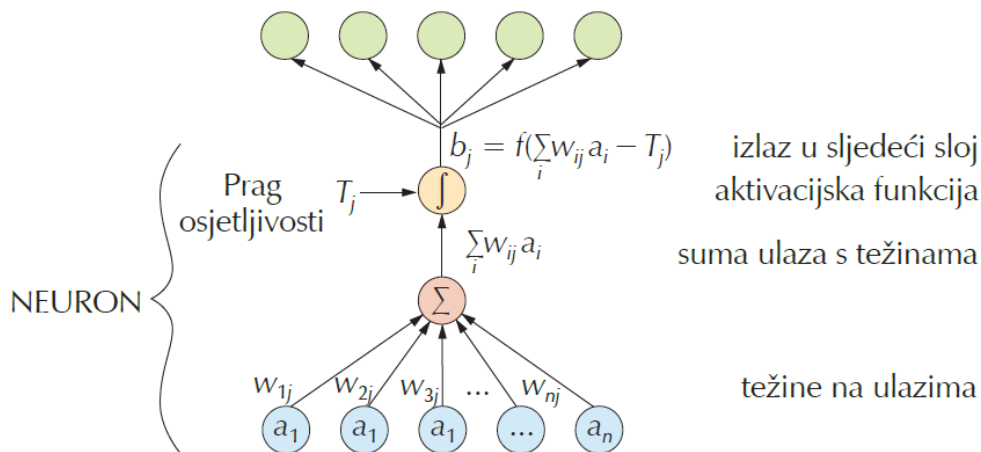
2.4.1. Klasifikacija neuronskih mreža

Sa strukturnoga stajališta, ovisno o modelu neurona od kojega su građene te po načinu prostiranja signala kroz mrežu, neuronske mreže mogu se podijeliti na statičke unaprijedne (eng. *Feedforward*) i dinamičke povratne (eng. *Recurrent*). U primjeni su najčešće višeslojne statičke neuronske mreže. Od dinamičkih neuronskih mreža najčešće se koriste višeslojne neuronske mreže s elementima zadržke (eng. *Time Delay Neural Networks*). Uz spomenute, postoje i neizrazite neuronske mreže (eng. *Fuzzy Neural Networks*). Neuronske mreže koje mogu imati po nekoliko stotina skrivenih slojeva nazivaju se dubokim učenjem (eng. *Deep Learning*).

Specificiranje umjetne neuronske mreže podrazumijeva definiranje strukture mreže (shema povezivanja neurona, broj skrivenih slojeva i broj neurona u slojevima), prijenosne funkcije i pravila učenja. Model neuronske mreže definiran je dvama vrstama parametara: parametrima modela koji se podešavaju tijekom učenja (težinski koeficijenti i pristranost u neuronskim mrežama) te hiperparametrima koji se biraju na temelju prethodno provedenih eksperimenata (prijenosna funkcija, broj neurona, broj skrivenih slojeva, algoritam učenja) [13].

2.4.2. Struktura umjetne neuronske mreže

Strukturu umjetne neuronske mreže čine ulazni, skriveni i izlazni sloj. Ulazni sloj (eng. *Input Layer*) prima informacije, ulazne signale, i prosljeđuje ih mreži. Svaki ulazni signal množi se težinskim koeficijentom w_i , koji su analogni dendritima biološkog neurona. Tako otežani ulazni signali se zbrajaju, a njihov se zbroj uspoređuje s pragom osjetljivosti neurona, T_j (eng. *Threshold*). Skriveni sloj (eng. *Hidden Layer*) prima informacije iz ulaznoga sloja i obrađuje ih zbrajajući otežane ulaze pomoću neke funkcije sumiranja i tako stvara vlastitu internu aktivaciju. Ako je zbroj otežanih signala veći od praga osjetljivosti neurona, nelinearna aktivacijska funkcija f generira izlazni signal neurona iznosa b_j . Prijenosna funkcija može biti diskontinuirana skokomična funkcija ili neka kontinuirana funkcija, kao npr. sigmoida ili tangens-hiperbolna funkcija. Izlazni sloj (eng. *Output Layer*) prima obrađene informacije i daje rezultate.



Slika 2.8. *Struktura umjetne neuronske mreže* [9]

2.4.3. Učenje neuronske mreže

Učenje neuronske mreže podrazumijeva iteracijski postupak podešavanja veza među neuronima. Težinski koeficijenti su „nositelji znanja“ u neuronskoj mreži i do njihovih vrijednosti dolazi se učenjem mreže na određenome skupu ulaznih podataka. Dakle, učenje neuronske mreže podrazumijeva podešavanje težinskih koeficijenata tako da se dobije minimalna vrijednost pogreške između izlaza dobivenih modelom i stvarne vrijednosti izlazne veličine. Jedno od pravila učenja jest pravilo širenja unatrag ili algoritam unatražne propagacije izlazne pogreške (eng. *Back Propagation*). Pravilo učenja širenja unatrag radi na principu da se na temelju pogreške između stvarne i predviđene vrijednosti ispravljaju vrijednosti težinskih faktora metodom unatražne propagacije. Postupak se ponavlja sve dok se ne zadovolji zadani kriterij uspješnosti. Razlika vrijednosti dobivene modelom i stvarnih izlaznih vrijednosti služi za korekciju težinskih koeficijenata pomoću tzv. Delta pravila, koje je prikazano sljedećim jednadžbama:

$$w_{ij}^{novi} = w_{ij}^{stari} + \Delta w_{ij} \quad (6)$$

pri čemu je w_{ij} neuron i povezan s neuronom j .

Korektivni faktor računa se prema:

$$\Delta w_{ij} = \eta y_i \varepsilon_j \quad (7)$$

pri čemu je η koeficijent učenja, y_i izlaz iz prethodnog neurona (neurona i) te ε_j pogreška sljedećeg neurona (neurona j).

Pogreška sljedećeg neurona jest:

ε_j = vrijednost koju daje neuronska mreža – stvarna vrijednost

Globalna pogreška mreže je stoga:

$$E = \frac{1}{2} \sum_k (d_k - y_k)^2 \quad (8)$$

pri čemu je k broj neurona u izlaznom sloju, d_k željena vrijednost, dobivena eksperimentom, a y_k vrijednost dobivena neuronskom mrežom [13].

2.4.4. Koeficijent učenja

Koeficijent učenja diktira brzinu promjene težinskih koeficijenata s obzirom na vrijednost pogreške ε . Taj parametar mora se zadati i o njemu ovisi brzina konvergencije. Važno je zadati optimalnu vrijednost koeficijenta učenja jer premala vrijednost može rezultirati presporim učenjem, dok prevelika vrijednost nikada ne može pronaći minimum funkcije pogreške [13].



Slika 2.9. Utjecaj vrijednosti koeficijenta brzine učenja na učenje mreže [13]

2.4.5. Prijenosna funkcija

Prijenosna funkcija obično je nelinearna funkcija koja uči nelinearne funkcijske ovisnosti. Najčešće korištene prijenosne funkcije su sigmoidna i tangens-hiperbolna. To su kontinuirane funkcije i aproksimacija su skokomične funkcije kakva se nalazi i u biološkom neuronu te ReLU funkcija. Sigmoidna funkcija uvodi nelinearnost u model pri čemu izlazne veličine poprimaju vrijednost od 0 do 1. Sigmoidna funkcija nije centrirana oko nule, stoga ažuriranje težinskih faktora može biti neefikasno. Tangens hiperbolna funkcija poprima vrijednosti u intervalu $[-1,1]$ i centrirana je oko nule. Problem se javlja kada su izlazne vrijednosti blizu vrijednostima -1 i 1 što uzrokuje nestanak gradijenta pogreške. ReLU (eng. *Rectified Linear Units*) je funkcija „ispravljene linearne jedinice“. Kod ReLU funkcije problem nestajućega

gradijenta ne postoji, a konvergencija je brža. Nedostatak ove funkcije jest činjenica da ne može opisati nelinearnost među podacima.

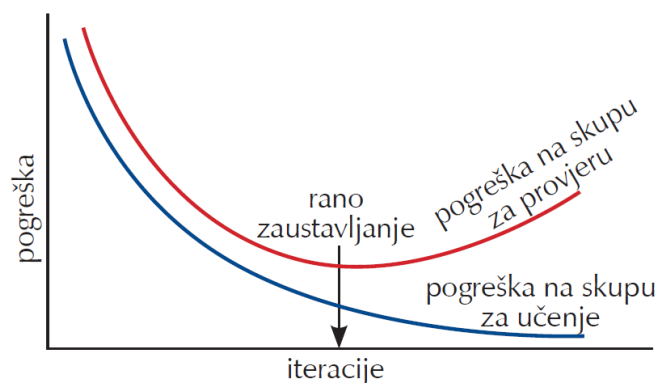
Ovisno o vrsti prijenosne funkcije koja se koristi, podatke je potrebno prethodno skalirati na vrijednosti s kojima funkcija može računati. U *Tablici 2.2.* prikazane su formule za spomenute prijenosne funkcije [13].

Tablica 2.2. Najčešće korištene prijenosne funkcije i načini skaliranja [14]

Prijenosna funkcija	Skaliranje podataka	Broj jednačbe	Skalirani podaci
<i>sigmoid</i>	$x_{skal}^i = \frac{x^i - x_{min}}{x_{max} - x_{min}}$	(9)	u rasponu od 0 do 1
<i>tanh</i>	$x_{skal}^i = 0,5 \cdot \tanh \left(0,01 \cdot \frac{x^i - \mu_x}{\sigma_x} \right)$	(10)	između -1 i 1
<i>ReLU</i>	$x_{skal}^i = \frac{x^i - \mu_x}{\sigma_x}$	(11)	centrirani oko 0

2.4.6. Pretreniranost neuronske mreže

Kao što je objašnjeno u poglavlju 2.3.3., podaci se dijele na tri skupa: skup za učenje, skup za provjeru (validaciju) te skup za testiranje. Skup za provjeru služi kako bi se provjerilo je li došlo do pretreniranosti (preusklađivanja) mreže. To se događa kada mreža daje odlične rezultate na podacima na kojima je učila, dok na novim podacima daje lošije rezultate. Kako bi se spriječilo pretreniranje mreže može se koristiti tehnika ranog zaustavljanja (eng. *Early Stopping*). Ona omogućuje zaustavljanje učenja onda kada pogreška na skupu za učenje i dalje opada, a pogreška na skupu za provjeru dosegne minimum.

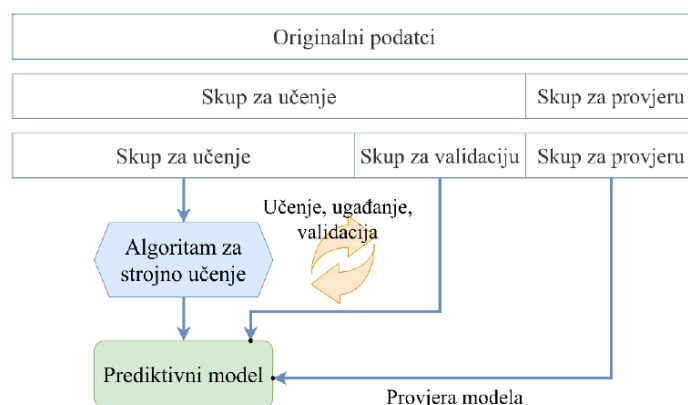


Slika 2.10. Učenje neuronske mreže tehnikom ranog zaustavljanja [13]

2.4.7. Ocjena izvedbe neuronske mreže

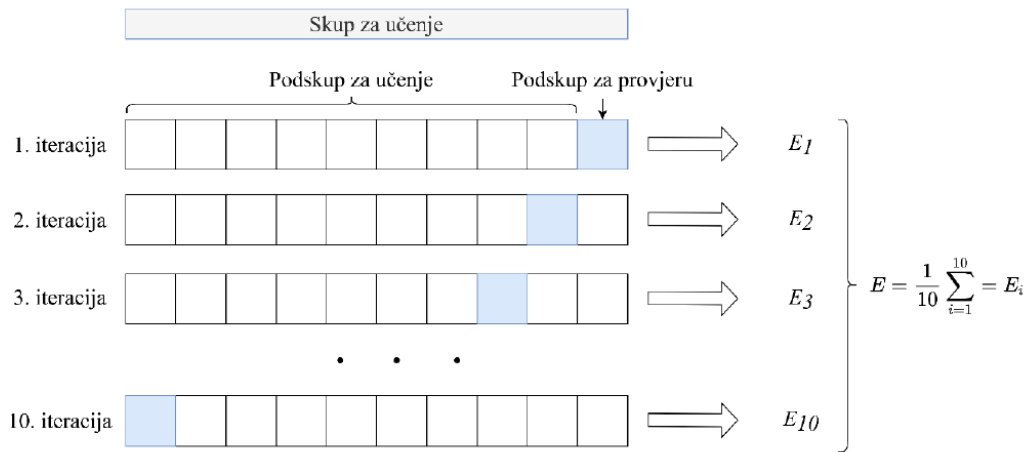
Konačno, model se ocjenjuje provjerom na skupu za testiranje, na potpuno novim, modelu do tad nepoznatim podacima. Najčešće se koristi kriterij pogreške koji govori o kvaliteti i generalizaciji mreže. Ako sposobnost predviđanja s tim skupom nije zadovoljavajuća, onda se struktura modela može izmijeniti promjenom broja skrivenih neurona, promjenom broja prošlih vrijednosti ulaza i/ili izlaza ili promjenom aktivacijske funkcije [9].

Kako bi se dobio optimalan model neuronske mreže, uobičajeno je optimirati parametre modela i hiperparametre križnom validacijom (eng. *Cross Validation*). Razlikuje se križna validacija uz zadržavanje (eng. *Holdout Cross-validation*) te križna validacija koja se ponavlja k puta (eng. *k-fold Cross-validation*). Križna validacija uz zadržavanje pogodna je za velike skupove podataka. Osjetljiva je na način na koji su podaci razdvojeni tako da ocjena modela može varirati za različite skupove podataka.



Slika 2.11. Shema postupka križne validacije uz zadržavanje [13]

Križna validacija koja se ponavlja k puta nasumično dijeli podatke u k skupina te se svaki podatak jednom nađe u validacijskome skupu podataka (skupu za provjeru) što rezultira manjom varijabilnosti pogreške u ocjeni izvedbe modela [13].

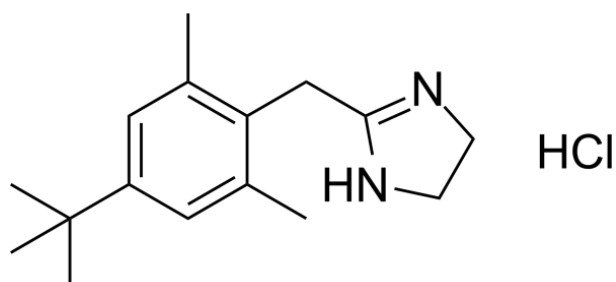


Slika 2.12. Shema postupka križne validacije koja se ponavlja k puta [13]

3. EKSPERIMENTALNI DIO

3.1. Materijali

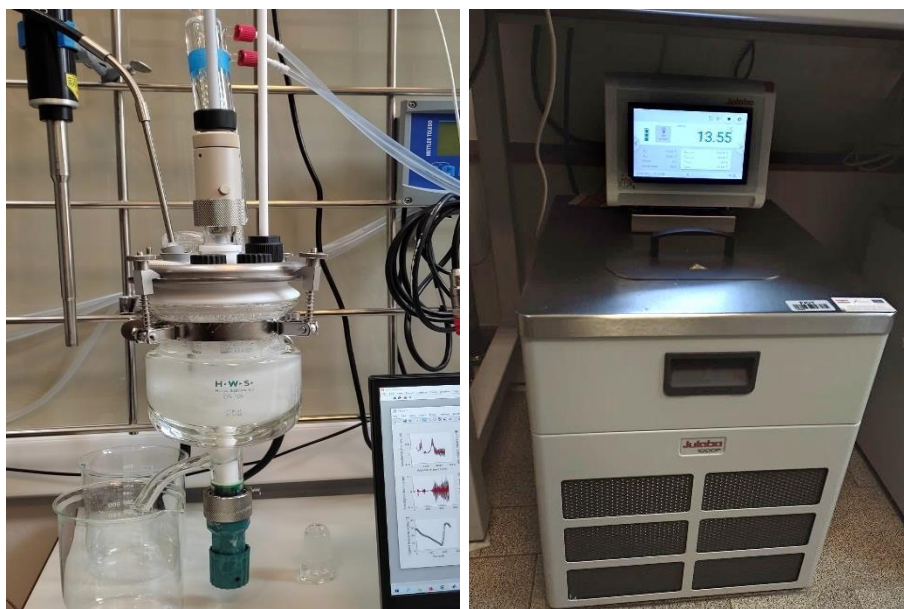
Za izradu kalibracijskoga modela iz prikupljenih spektralnih podataka korištena je djelatna tvar ksilometazolin hidroklorid (molekulska formula: $C_{16}H_{25}ClN_2$, molekulska masa: 280,8 g/mol) otopljena u n-butanolu (molekulska formula: C_4H_9OH , molekulska masa: 74,123 g/mol). Na *Slici 3.1.* prikazana je struktura molekule ksilometazolin HCl.



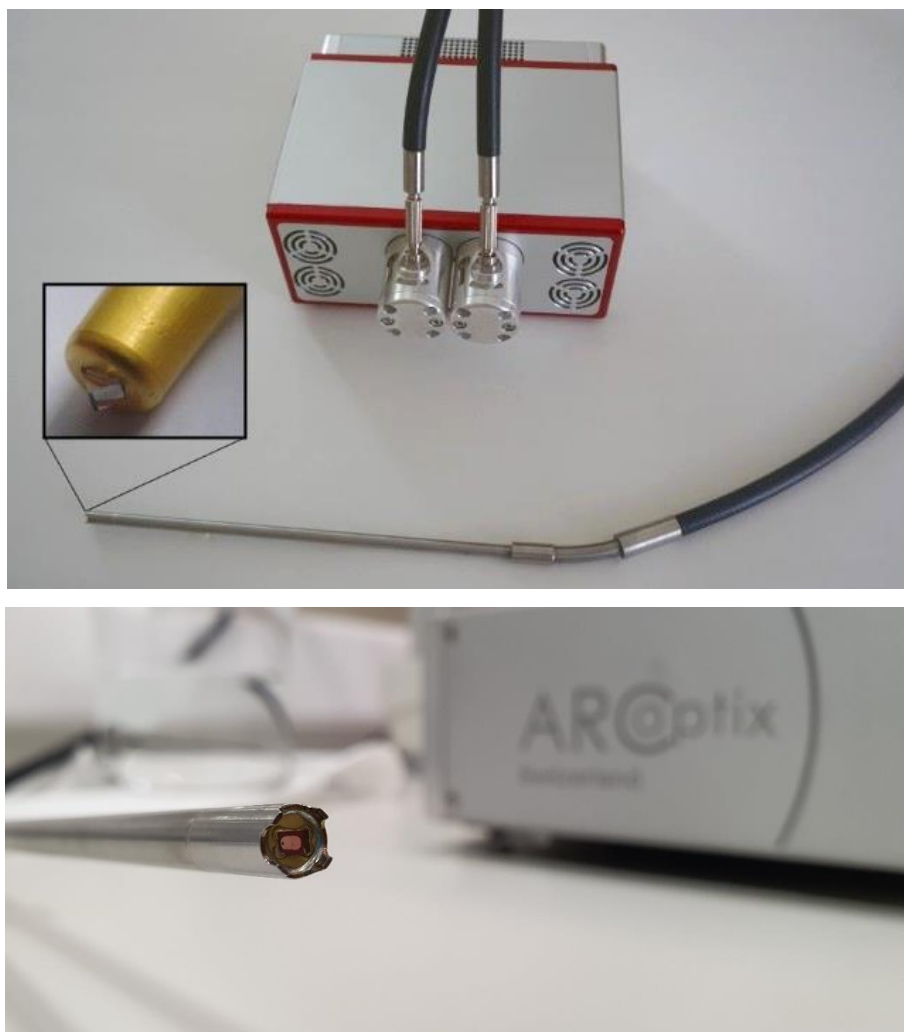
Slika 3.1. *Struktura molekule ksilometazolin hidroklorida* [16]

3.2. Uređaji

Aparatura za provedbu eksperimenta prikazana je na *Slici 3.2* te se sastoji od 500 ml staklenoga kristalizatora s plaštem, temperaturnoga osjetila Pt-100 i termostata za regulaciju temperature (Julabo, Magio MS-1000F), stupne miješalice s PTFE miješalom.. Za prikupljanje spektralnih podataka korišten je ATR-FTIR spektrofotometar (ArcOptix, FC-120-4TE), čije je sonda (*Slika 3.3.*) direktno uronjena u otopinu.



Slika 3.2. *Aparatura: kristalizator s plaštem, miješalom i ATR FTIR sondom (lijevo), termostat (desno)*

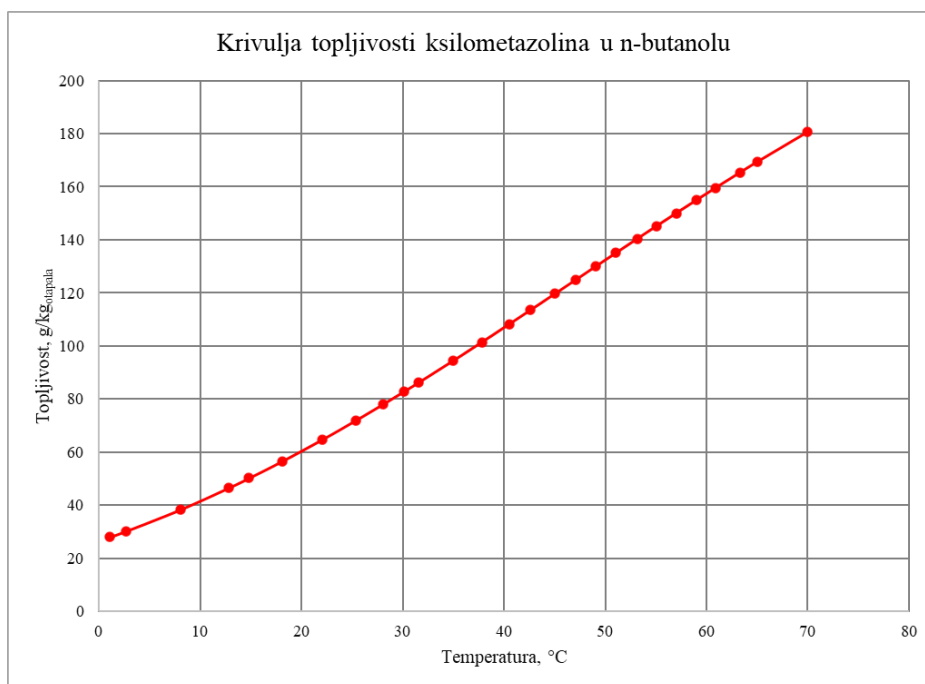


Slika 3.3. ATR FTIR spektrometar

3.3. Postupak rada

3.3.1. Određivanje krivulje topljivosti

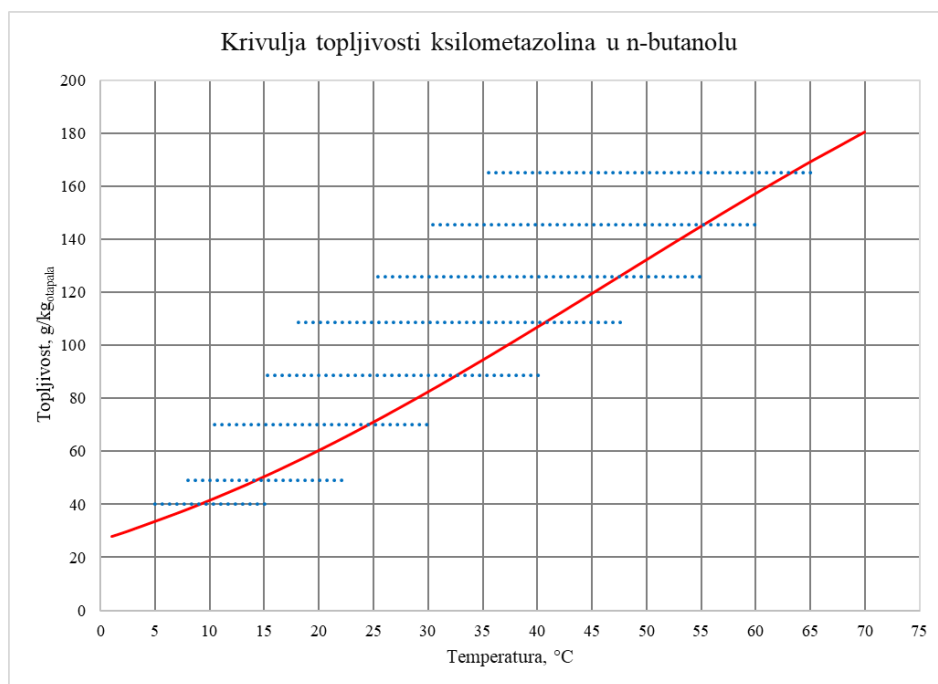
Prvi korak u provođenju eksperimenta bio je određivanje krivulje topljivosti ksilometazolin HCl u n-butanolu. N-butanol u reaktoru ohlađen je na 0 °C te su dodavane odvage ksilometazolina HCl. Nakon svake dodane odvage, po potrebi je povišena temperatura u reaktoru do potpunoga otapanja kristala. Temperature potpunoga otapanja zapisane su u tablicu te je dobiven grafički prikaz krivulje topljivosti prikazan na *Slici 3.4*.



Slika 3.4. Grafički prikaz krivulje topljivosti ksilometazolin HCl-a u n-butanolu

3.3.2. Priprema otopina i prikupljanje temperatura

Drugi korak u provođenju eksperimenta bio je priprema otopina poznatih koncentracija. Na temelju prethodno određene krivulje topljivosti ksilometazolin HCl-a u n-butanolu, odabrano je 8 koncentracija koje se protežu kroz cijeli raspon temperatura. Otopine koncentracija 40,13; 55,39; 71,14; 88,51; 107,00; 125,10; 144,32; 164,77 g/kg_{otapala} pripremljene su otapanjem točno određenih masa u n-butanolu. Otopine su zatim zagrijavane do temperature više od krivulje topljivosti kako bi se osiguralo potpuno otapanje djelatne tvari. Otopine su nakon toga hlađene konstantnom brzinom hlađenja. Tijekom procesa hlađenja, prikupljene su temperature i spektralni podaci.



Slika 3.5. Krivulja topljivosti i snimano područje metastabilne zone obuhvaćeno s osam koncentracija

3.3.3. Prikupljanje i predobrada spektralnih podataka

Prikupljanje spektralnih podataka iz podzasićene zone i metastabilne zone ksilometazolin HCl-a u n-butanolu provedeno je ATR-FTIR spektroskopijom. ATR-sonda direktno je uronjena u otopinu te uređaj kontinuirano prikuplja spektralne podatke koji se spremaju na računalo. Za očitane temperature u trenutku snimanja pri određenoj koncentraciji dobiva se spektar s rasponom valnoga broja od 840 do 5000 cm^{-1} . Eksperimentalno dobiveni podaci spremaju se u *Excel* datoteku. Prije samoga razvoja modela, snimljeni spektralni podaci predobrađeni su tehnikom MSC (eng. *Multiplicative Scatter Correction*). Svrha ove tehnike predobrade podataka jest smanjivanje raspršenja u spektralnim podacima kao i eliminacija odstupanja (eng. *outliers*). Skaliranje podataka provedeno je prema jednadžbama za *Min-max* normalizaciju (4) te *Z-score* normalizaciju odnosno standardizaciju (5).

3.3.4. Razvoj kalibracijskog modela

Za razvoj modela neuronske mreže korišteni su snimljeni spektralni podaci u rasponu valnoga broja 840-1800 cm^{-1} s pripadajućim temperaturama. Snimljeni spektralni podaci iznad 1800 cm^{-1} sadržavali su šum te nisu korišteni za razvoj modela.

3.3.4.1. Softversko razvojno okruženje

U ovome je radu za razvoj i ispitivanje modela neuronske mreže korišten programski jezik *Python*, odnosno njegovo programsko okruženje *Spyder* (eng. *Scientific Python Development Environment*). *Python* je programski jezik općenite namjene razvijen kasnih 1980-ih kao nadogradnja programskoga jezika ABC. Koristi se za razvoj softvera, izradu web stranica, te se vrlo često koristi za primjenu metoda znanosti o podacima (eng. *data science*). [17] *Spyder* je platforma za programiranje unutar programskoga jezika *Python*. U sebi sadrži pakete poput *NumPy*, *SciPy*, *Matplotlib*, *Pandas*, *Keras*, *Sklearn*, *TensorFlow*. Biblioteka *Pandas* učitava podatke u obliku tablice direktno iz *Excel* datoteke. *Keras* je programsko sučelje (eng. *Application Programming Interface, API*) razvijeno za potrebe dubokoga učenja koje u svojoj pozadini koristi platformu *TensorFlow*, *CNTK* ili *Theano*. Razvijen je kako bi se omogućio brz i jednostavan razvoj modela dubokoga učenja. Ostale učitane biblioteke potrebne su za matematičke operacije i manipulaciju višedimenzionalnim matricama (*NumPy*), izradu neuronske mreže (*TensorFlow*), dijeljenje podataka na podskupove za učenje i validaciju, računanje kriterija uspješnosti (*Sklearn*) i njihovo pohranjivanje (*Json*) [18].

3.3.4.2. Razvoj modela neuronske mreže u Pythonu

Struktura neuronske mreže u *Pythonu* zadaje se pomoću funkcije *model* koja predstavlja organizaciju slojeva u mreži. U ovome radu korišten je sekvencijalni model, *model=Sequential()*. Takav tip modela predstavlja linearni stog slojeva i pogodan je za uporabu kada svaki sloj ima po jedan tenzor na ulazu i izlazu [19]. Slojevi se dodaju naredbom *model.add()*, a unutar zagrade upisuje se tip sloja. U ovome radu korišten je sloj *Dense* koji predstavlja standardni sloj umjetne neuronske mreže, duboko povezan s prethodnim i budućim neuronima u susjednim slojevima. Za svaki sloj potrebno je definirati koliko će neurona imati, koja je dimenzija ulaznoga sloja te koja je aktivacijska funkcija naredbom *activation*. Izlazni sloj bit će onaj koji je zadnji zapisan u nizu *.add()*, funkcija. U ovome radu definirana je jedna izlazna varijabla koju želimo naučiti, koncentracija, stoga izlazni sloj ima samo jedan neuron te mu je aktivacijska funkcija *linear*. To znači da se izlazna varijabla iz zadnjega sloja neće dodatno nelinearizirati.

Nakon što je zadana struktura neuronske mreže, definira se proces učenja naredbom *model.compile()* pri čemu se definira algoritam optimiranja, tzv. optimizator te funkcija pogreške. Optimizator je algoritam učenja ili metoda optimiranja koja se koristi za procjenu parametara neuronske mreže kao što su težinski koeficijenti i koeficijent učenja kako bi se minimizirala funkcija pogreške [20]. U ovome radu korišteni su neki od najpoznatijih

optimizatora koji se temelje na algoritmu gradijentnog spusta kao što su *Adam*, *Nadam*, *Adadelta* i *RMSprop*.

Naredba *model.fit()* služi za učenje tj. treniranje modela na podacima namijenjenim za učenje. U okviru te naredbe potrebno je zadati broj epoha (broj prolaska svih podataka kroz mrežu) te *batch size* (broj uzoraka sljedova podataka koji se obrađuju kao cjelina).

Koristeći prethodno naučeni model, naredba *model.predict()* računa vrijednosti izlazne varijable prema danim ulaznim varijablama [20].

3.3.6. Optimizacija modela neuronske mreže

Za ocjenu uspješnosti razvijenoga modela neuronske mreže korišteni su kriteriji srednje kvadratne pogreške (eng. *Mean Squared Error*, *MSE*) (12) i koeficijent determinacije (eng. *Coefficient of Determination*, R^2) (13). Pri tome su se navedeni kriteriji pratili na dva seta podataka, od kojih se jedan set odnosi na učenje modela, a drugi na validaciju modela.

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (12)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (13)$$

y_i predstavlja izmjerenu vrijednost, \hat{y}_i modelom procijenjenu vrijednost, \bar{y} srednju vrijednost mjerene veličine, a n broj podataka.

Kako bi se dobio što točniji model, odnosno kako bi se dobili što manji kriteriji srednje kvadratne pogreške te koeficijent determinacije što bliže jedinici, korišteni su skupovi podataka prikazani u *Tablici 3.1*. U svrhu dobivanja što točnijega modela, metodom pokušaja i pogreške provjeravao se utjecaj broja skrivenih slojeva, broja neurona u prvome i drugom skrivenom sloju, vrsta aktivacijske funkcije, vrsta algoritma učenja (optimizatora), *batch size* te broj epoha. Ispitivane prijenosne funkcije bile su: sigmoidna, tangens-hiperbolna i ReLU funkcija. Načini skaliranja podataka za spomenute funkcije prikazani su u *Tablici 2.2*. u teorijskome dijelu ovoga rada.

Za skupove podataka D1 i D2 prikazanih u *Tablici 3.1*. najprije je ispitan utjecaj broja neurona u samo jednome skrivenom sloju za pojedine funkcije. Zatim se pratio utjecaj vrste optimizatora na iznose *MSE* i R^2 za odabrani broj neurona te za pojedinu funkciju. Za

najbolju kombinaciju hiperparametra modela iz prethodna dva ispitivanja provjeren je utjecaj broja epoha te *batch size*.

Za najbolje modele iz skupova D1 i D2 ispitan je utjecaj dodavanja drugoga skrivenog sloja te broja neurona u oba sloja. Tako je dobiven skup D3 koji čine hiperparametri najboljeg modela skupa D1, ali s dva skrivena sloja, odnosno D4 koji čine hiperparametri najboljega modela skupa D2 s dva skrivena sloja. U sljedećem poglavlju prikazani su i komentirani dobiveni rezultati.

Tablica 3.1. *Skupovi podataka za optimizaciju modela neuronske mreže*

Skup podataka	Predobrada	Broj skrivenih slojeva	Ispitani broj neurona	Prijenosne funkcije	Ispitani optimizatori
D1	/	1	2-112, korak 10	<i>sigmoid</i>	Nadam, Adam, Adadelta, RMSprop
				<i>tanh</i>	
				<i>ReLU</i>	
D2	MSC	1	2-112, korak 10	<i>sigmoid</i>	Nadam, Adam, Adadelta, RMSprop
				<i>tanh</i>	
				<i>ReLU</i>	
D3	/	2	2-112, korak 10 u prvom i drugom skrivenom sloju	<i>tanh</i>	Adam
D4	MSC	2	2-112, korak 10 u prvom i drugom skrivenom sloju	<i>ReLU</i>	Nadam

4. REZULTATI I RASPRAVA

Cilj ovoga rada bio je razviti i ispitati modele neuronskih mreža za modeliranje ovisnosti koncentracije djelatne tvari, ksilometazolin hidroklorida, u n-butanolu, o temperaturi i spektralnim podacima dobivenih mjerenjima ATR-FTIR spektrometrom.

U radu je ispitan utjecaj predobrade podataka, broja skrivenih slojeva, broja neurona u prvome i drugom skrivenom sloju, vrste primijenjenoga optimizatora, kao i prijenosne funkcije na kvalitetu razvijenih neuronskih mreža.

U ovom su poglavlju grafički i tablično prikazani i protumačeni rezultati za svaki ispitivani skup podataka i to s obzirom na broj skrivenih slojeva, broj neurona u prvome i drugom skrivenom sloju, vrstu aktivacijske funkcije, vrstu korištenoga optimizatora te broj epoha i *batch size*.

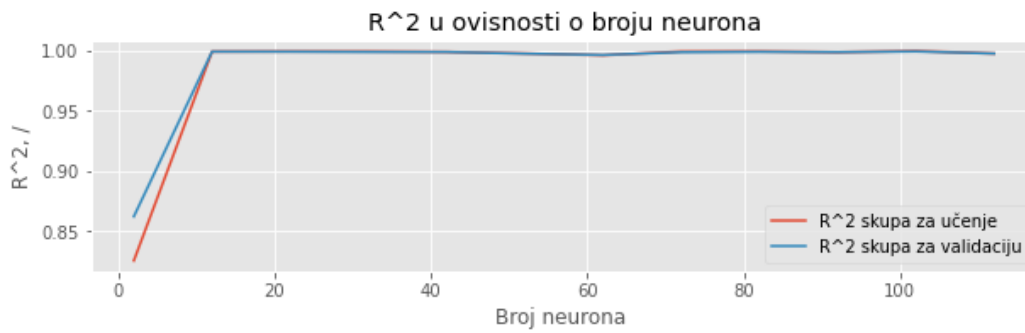
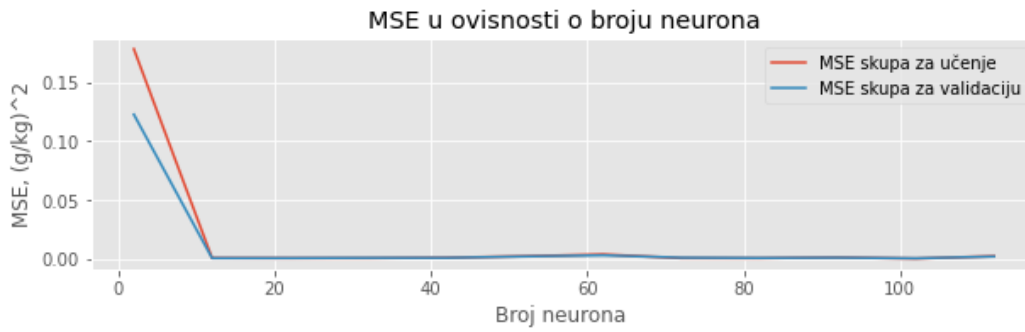
4.1. Rezultati za skup podataka D1

Za skup podataka D1 nije provedena predobrada podataka, već su se koristili sirovi podaci, s valnim brojem u rasponu 840-1800 cm^{-1} .

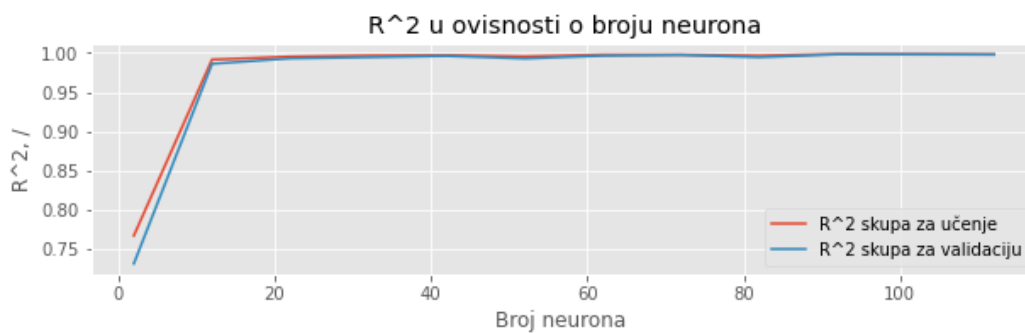
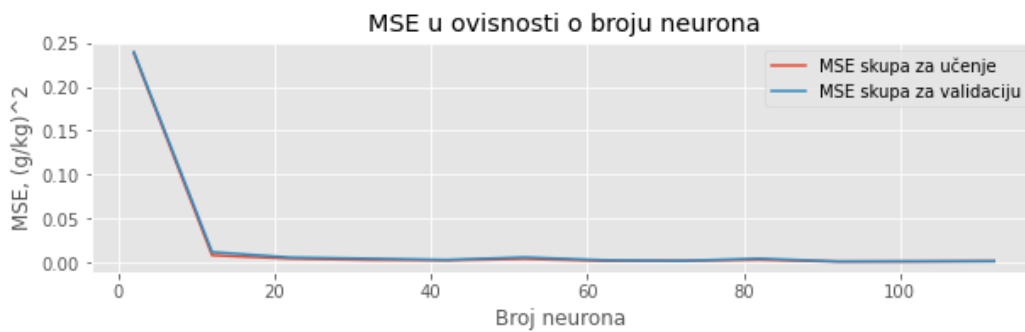
4.1.1. Ispitivanje broja neurona za pojedine aktivacijske funkcije

U postupku razvijanja i optimizacije modela neuronske mreže najprije se ispitao utjecaj broja neurona u samo jednome skrivenom sloju na iznose MSE te R^2 za pojedine aktivacijske funkcije. Rezultati su grafički prikazani na *Slikama 4.1., 4.2., i 4.3.*

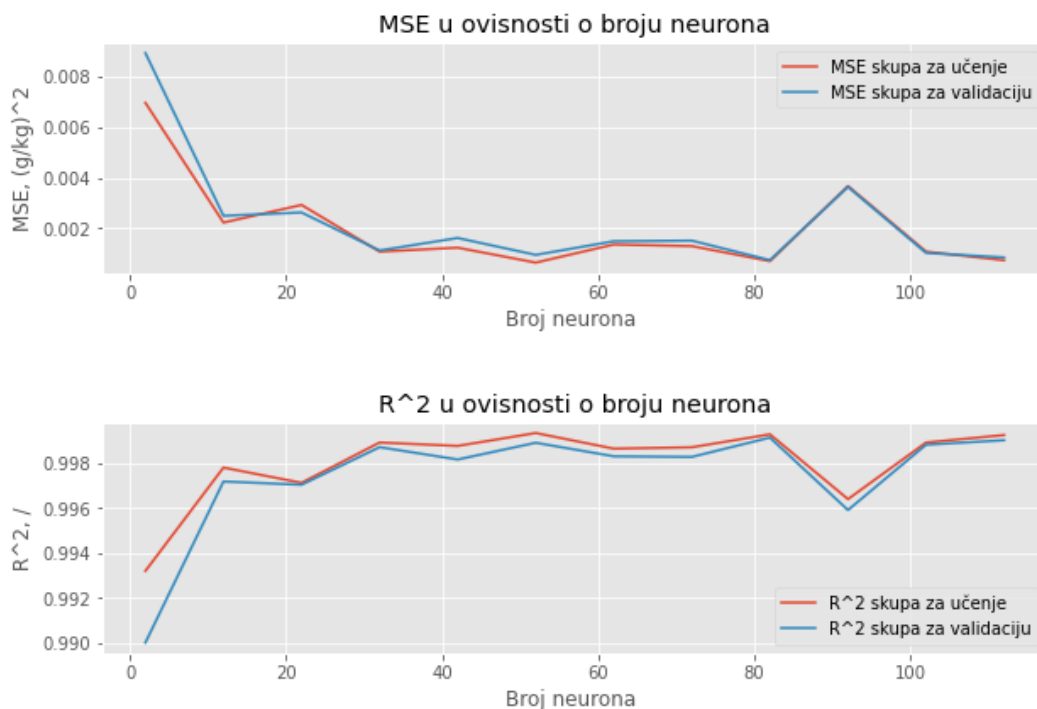
Iz dobivenih grafičkih prikaza može se vidjeti kako za aktivacijsku funkciju ReLU nema potrebe za brojem neurona većim od 10 jer se MSE ne smanjuje, odnosno R^2 ne povećava s daljnjim povećanjem broja neurona. Sličan je trend primijećen i kod sigmoidne aktivacijske funkcije za broj neurona veći od 20. Kod aktivacijske funkcije *tanh* vide se oscilacije u vrijednostima MSE i R^2 kako se broj neurona povećava. Za daljnja ispitivanja za *tanh* funkciju kao optimalan broj neurona uzet je broj 30 jer je za njega dobivena vrlo mala pogreška MSE , odnosno veliki R^2 , no raspon vrijednosti 10-20 također je zadovoljavajući.



Slika 4.1. MSE i R^2 u ovisnosti o broju neurona za $ReLU$ aktivacijsku funkciju



Slika 4.2. MSE i R^2 u ovisnosti o broju neurona za $sigmoid$ aktivacijsku funkciju



Slika 4.3. MSE i R^2 u ovisnosti o broju neurona za tanh aktivacijsku funkciju

4.1.2. Ispitivanje utjecaja vrste optimizatora

U Tablici 4.1. prikazani su iznosi MSE i R^2 na skupu za učenje i skupu za validaciju za ispitane optimizatore te aktivacijske funkcije. Iz tablice je vidljivo kako su za sve tri aktivacijske funkcije najniže vrijednosti MSE skupa za učenje i validaciju, a analogno najviše vrijednosti R^2 , postižu uporabom ADAM, NADAM te RMSprop optimizatorima, dok ADADELTA daje vrlo loše rezultate. Za funkciju ReLU najbolje rezultate daje NADAM, za sigmoidnu RMSprop, a za tanh ADAM. Ako se usporede sve funkcije i optimizatori, najbolji rezultati, odnosno najniže vrijednosti MSE, a vrijednosti R^2 najbliže jedinici za oba skupa, dobivene su kombinacijom tanh aktivacijske funkcije s 30 neurona te NADAM optimizatorom.

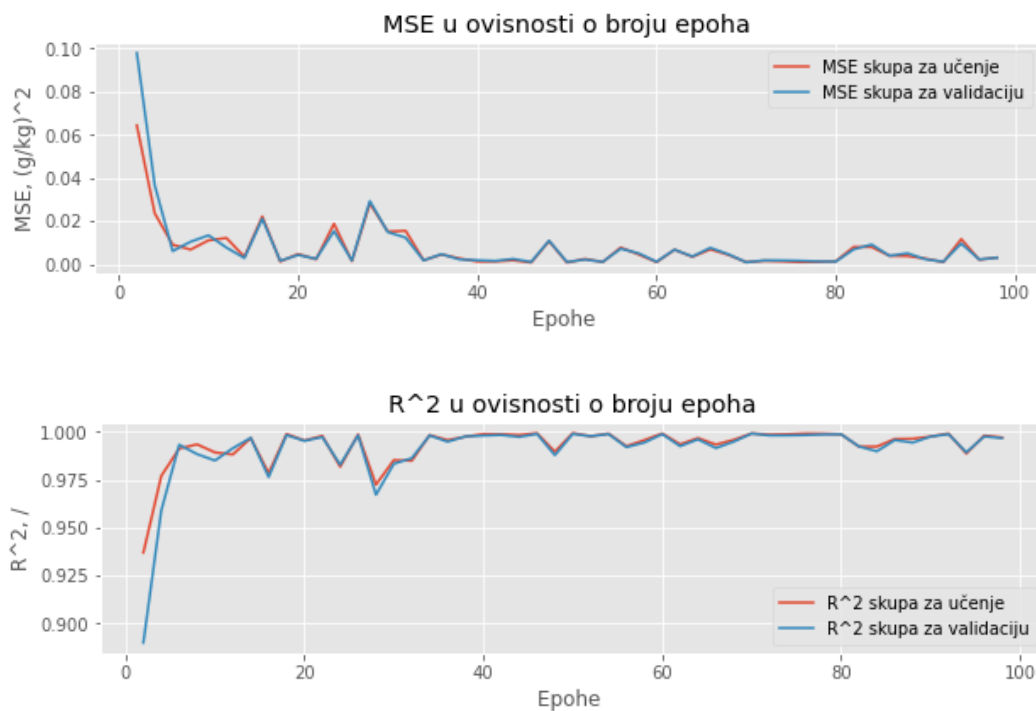
Tablica 4.1. MSE i R^2 u ovisnosti o pojedinoj optimizatoru i aktivacijskoj funkciji

	Skup za učenje		Skup za validaciju	
	MSE	R^2	MSE	R^2
<i>ReLU</i>				
ADAM	0,0049	0,9952	0,0061	0,9932
NADAM	0,0031	0,9970	0,0023	0,9974
ADADELTA	1,2952	-0,2666	1,0230	-0,1474
RMSprop	0,0102	0,9901	0,0137	0,9846
<i>sigmoid</i>				
ADAM	0,0079	0,9923	0,0118	0,9868
NADAM	0,0084	0,9918	0,0096	0,9893

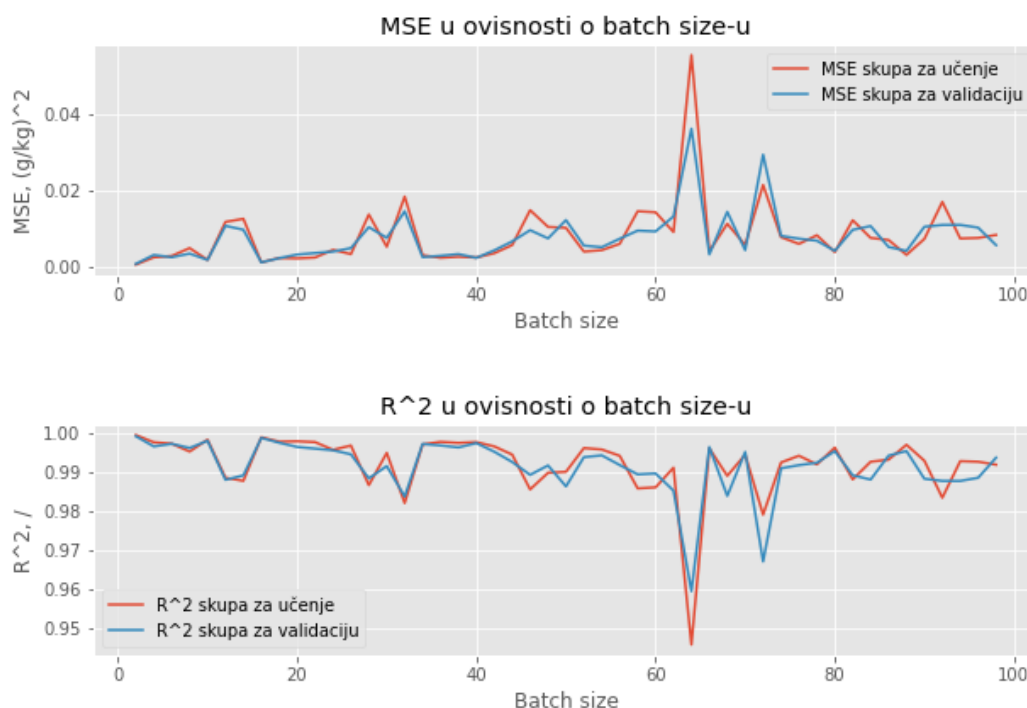
ADADELTA	1,6522	-0,6157	1,7861	-1,0034
RMSprop	0,0058	0,9943	0,0057	0,9936
<i>tanh</i>				
ADAM	0,0025	0,9975	0,0035	0,9961
NADAM	0,0011	0,9989	0,0014	0,9983
ADADELTA	1,144	-0,119	1,0101	-0,1329
RMSprop	0,0166	0,9837	0,0199	0,9776

4.1.3. Ispitivanje utjecaja broja epoha i *batch size*-a

Nakon što je pronađena najbolja kombinacija aktivacijske funkcije, broja neurona u skrivenom sloju te vrsta optimizatora, ispitao se utjecaj broja epoha i *batch size*. Iz prethodnih ispitivanja, za skup podataka D1, to je model koji sadrži tangens hiperbolnu funkciju u skrivenome sloju, 30 neurona u skrivenom sloju te NADAM kao optimizator. Na Slikama 4.4. te 4.5. prikazan je utjecaj broja epoha i *batch size*-a na iznose kriterije uspješnosti, *MSE* i R^2 . Iznos srednje kvadratne pogreške najprije naglo pada, uz poneka odstupanja između 5. i 40. epohe. Daljnjim povećanjem broja epoha iznad 40 nema većega utjecaja na kriterije uspješnosti. Trend je sličan i za *batch size*, ali uz povećanje srednje kvadratne pogreške s povećanjem *batch size*-a iznad 40. Za izradu modela odabrana je kombinacija od 38 epoha i *batch size* 6.



Slika 4.4. *MSE* i R^2 u ovisnosti o broju epoha



Slika 4.5. MSE i R^2 u ovisnosti o batch veličini

4.1.4. Najbolji model za skup podataka D1

Na temelju svih prethodnih ispitivanja dobiven je najbolji model neuronske mreže čiji su hiperparametri prikazani u *Tablici 4.2.*, a kriteriji uspješnosti u *Tablici 4.3.*

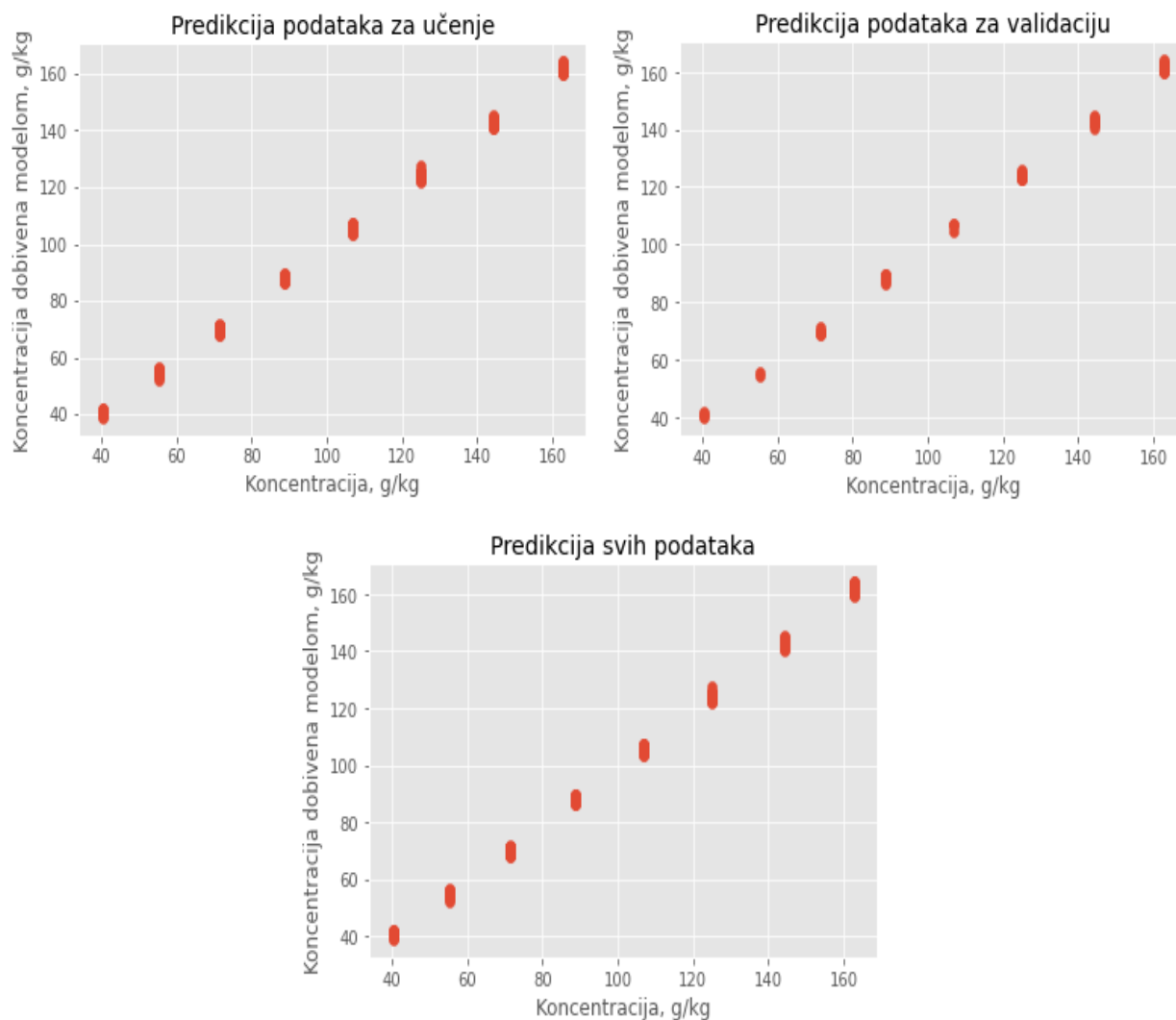
Tablica 4.2. Najbolji model neuronske mreže za skup podataka D1

Broj neurona u skrivenom sloju	Aktivacijska funkcija – skriveni sloj	Aktivacijska funkcija – izlazni sloj	Loss funkcija	Algoritam učenja	Broj epoha	Batch size
30	<i>tanh</i>	<i>LINEAR</i>	<i>MSE</i>	NADAM	38	6

Tablica 4.3. Kriterij uspješnosti za dobiveni model neuronske mreže za skup podataka D1

R^2 skupa za učenje	R^2 skupa za validaciju	R^2 svi podaci	MSE skupa za učenje	MSE skupa za validaciju
0,9985	0,9981	0,9985	0,0015	0,0017

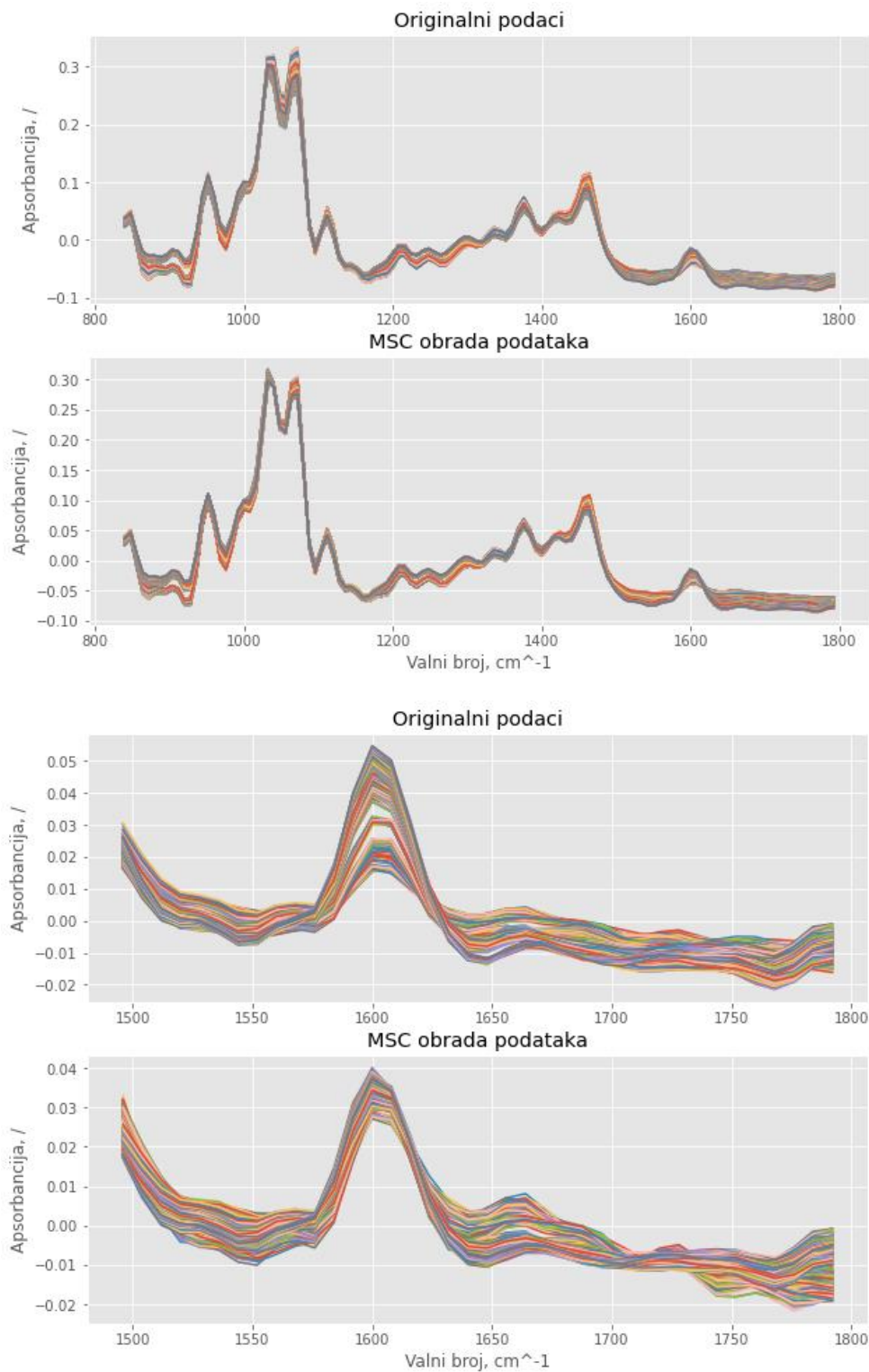
Na *Slici 4.6.* nalaze se grafički prikazi usporedbe stvarnih koncentracija te onih dobivenih modelom neuronske mreže. Linearni trend i malo rasipanje točaka oko pravca potvrđuju prikladnost ovoga modela za skup D1. Prikladnost modela također dokazuju visoki iznosi koeficijenta determinacije, koji su vrlo blizu jedinici te jako male vrijednosti srednje kvadratne pogreške koje su vrlo blizu nuli (*Tablica 4.3.*).



Slika 4.6. Usporedba stvarnih koncentracija i onih dobivenih modelom za skup podataka D1 na skupu podataka za učenje, validaciju i za cjelokupni skup podataka

4.2. Rezultati za skup podataka D2

Skup podataka D2 predobrađen je tehnikom MSC kako bi se smanjila raspršenja u spektralnim podacima. Ovim postupkom ispituje se utjecaj predobrade podataka na uspješnost modela neuronske mreže. Na *Slici 4.7.* prikazana je usporedba spektra bez MSC obrade i s njom.

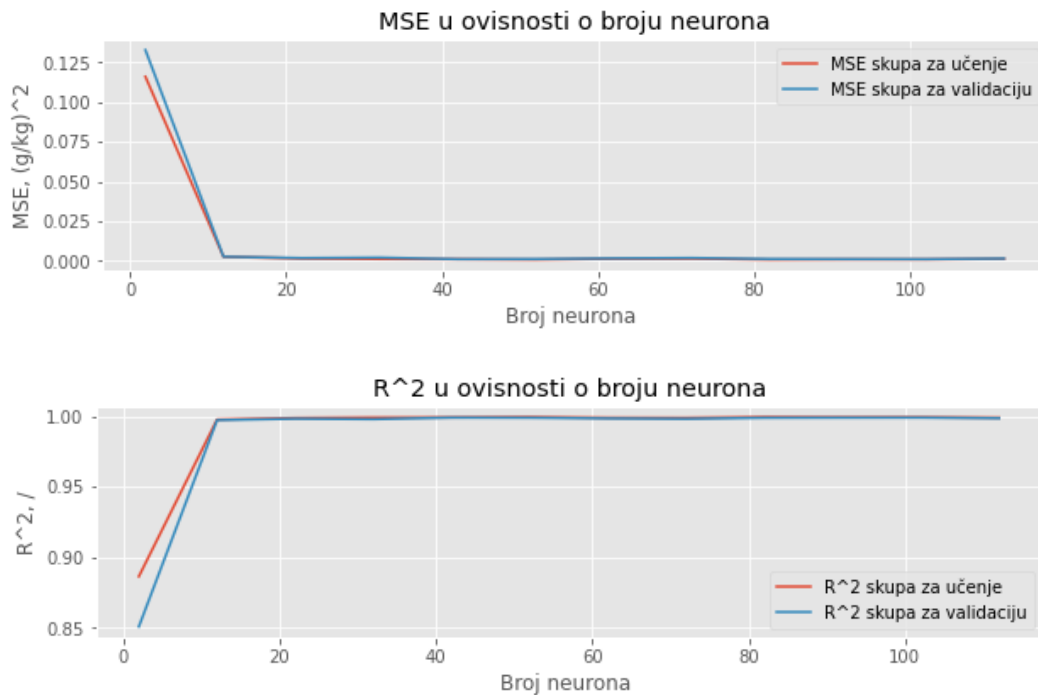


4.7. Usporedba spektra s originalnim podacima i podacima obrađenim MSC-tehnikom s rasponom valnih brojeva $840\text{-}1800\text{ cm}^{-1}$ (slika gore) i uvećan prikaz $1500\text{-}1800\text{ cm}^{-1}$ (slika dolje)

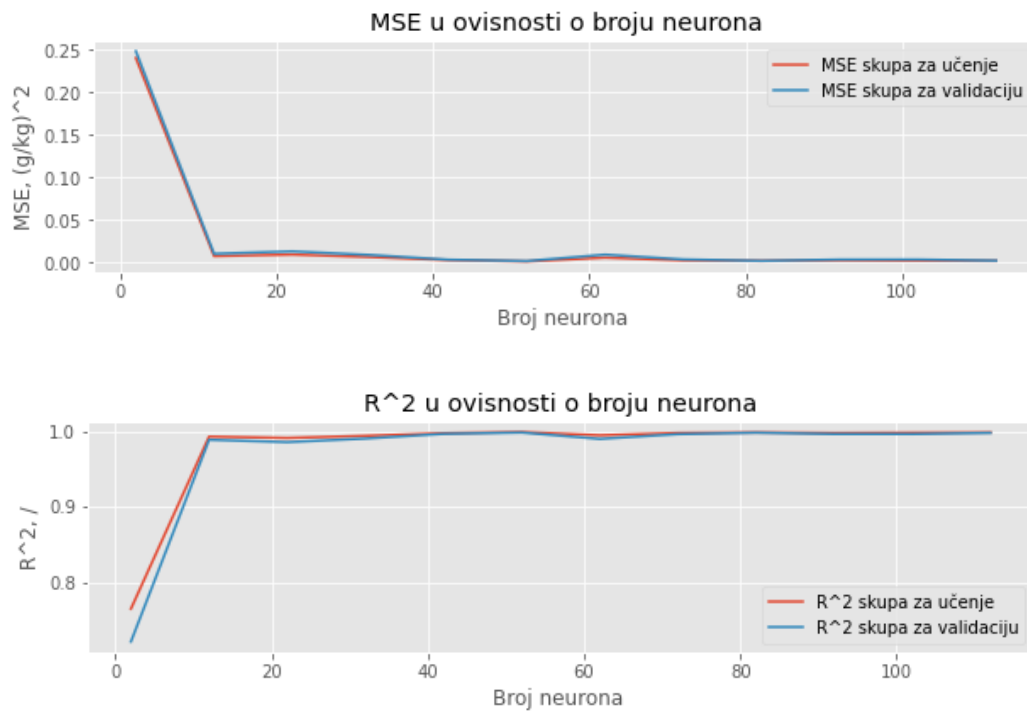
4.2.1. Ispitivanje broja neurona za pojedine aktivacijske funkcije

Na Slikama 4.8., 4.9., i 4.10. za skup podataka D2 dani su grafički prikazi ovisnosti MSE i R^2 o broju neurona u jednome skrivenom sloju za aktivacijske funkcije $ReLU$, $sigmoid$ i $tanh$. Za

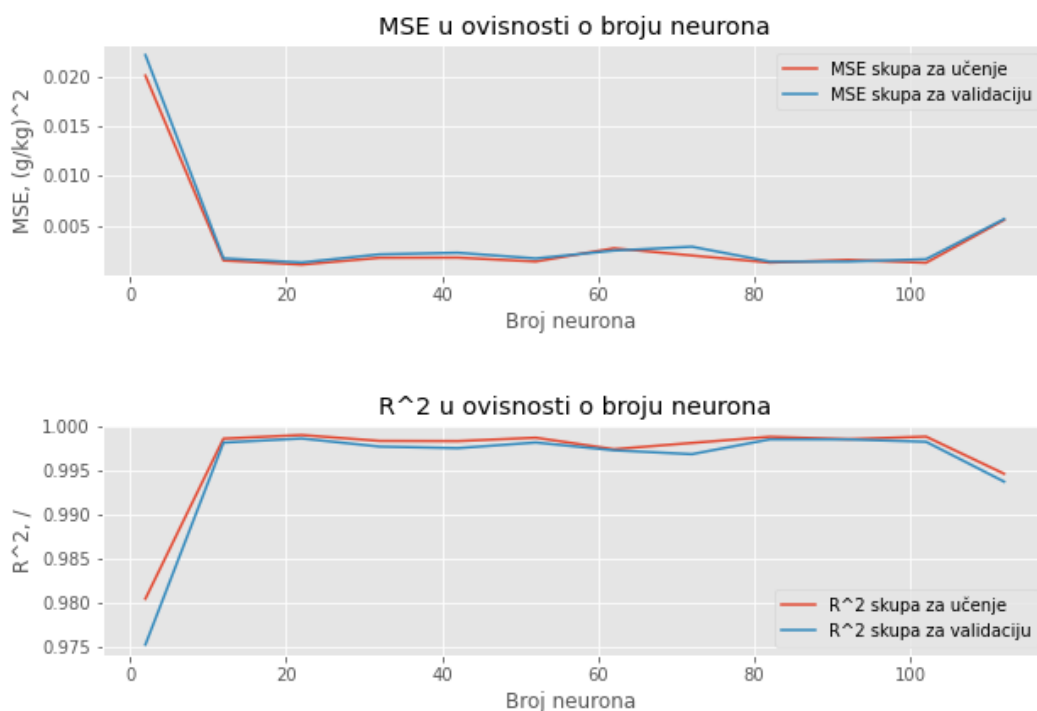
aktivacijske funkcije *ReLU* i *sigmoid* optimalni broj neurona je 10, kako se sa *Slika 4.8. i 4.9.* može vidjeti da daljnjim povećanjem broja neurona nema promjena u iznosima *MSE* i R^2 . Za funkciju *tanh* uzet je broj od 20 neurona u skrivenom sloju jer daljnjim povećanjem njihova broja dolazi do povećanja iznosa *MSE* te smanjivanja iznosa koeficijenta determinacije.



Slika 4.8. *MSE* i R^2 u ovisnosti o broju neurona za aktivacijsku funkciju *ReLU*



Slika 4.9. *MSE* i R^2 u ovisnosti o broju neurona za sigmoidnu aktivacijsku funkciju



Slika 4.10. MSE i R^2 u ovisnosti o broju neurona za tanh aktivacijsku funkciju

4.2.2. Ispitivanje utjecaja vrste optimizatora

U Tablici 4.4. prikazani su dobiveni rezultati pojedinih aktivacijskih funkcija s određenim brojem neurona s obzirom na vrstu primijenjenog optimizatora. Za aktivacijsku funkciju *ReLU* dobiven je ukupni najbolji rezultat s NADAM optimizatorom, što je vidljivo iz najnižih vrijednosti *MSE* te najviših vrijednosti koeficijenta determinacije R^2 . Za sigmoidnu i *tanh* aktivacijsku funkciju dobiveni su vrlo slični rezultati i gotovo jednako dobri uporabom ADAM optimizatora.

Tablica 4.4. MSE i R^2 u ovisnosti o pojedinom optimizatoru i aktivacijskoj funkciji

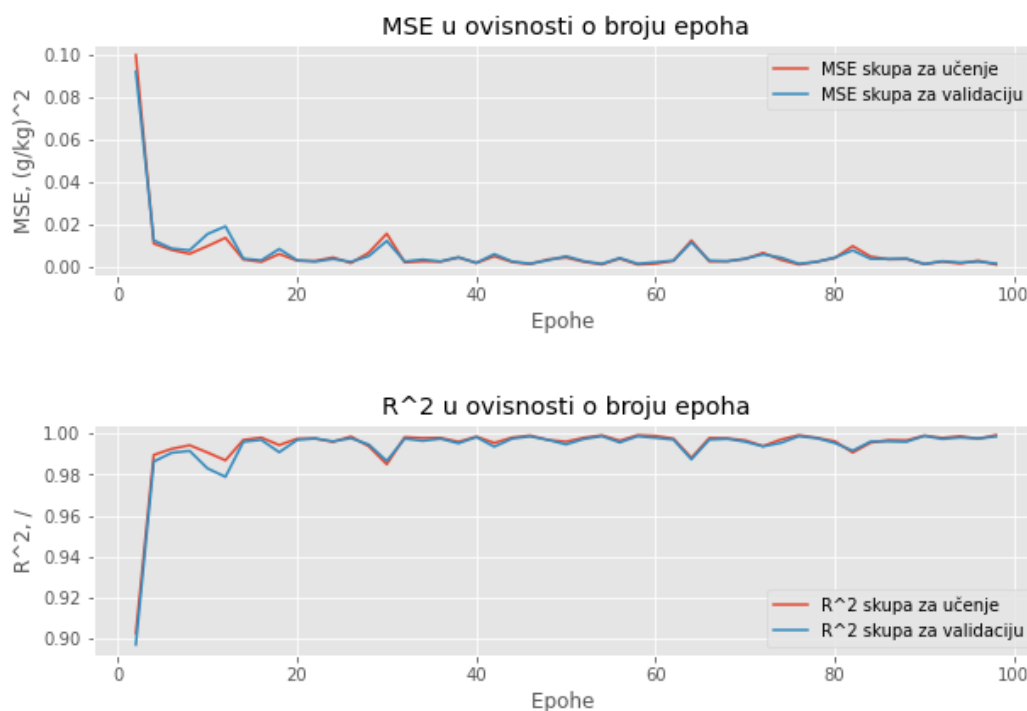
	Skup za učenje		Skup za validaciju	
	<i>MSE</i>	R^2	<i>MSE</i>	R^2
<i>ReLU</i>				
ADAM	0,0062	0,9940	0,0079	0,9912
NADAM	0,0011	0,9989	0,0012	0,9987
ADADELTA	2,3285	-1,2771	2,0722	-1,3242
RMSprop	0,0040	0,9961	0,0038	0,9958
<i>sigmoid</i>				
ADAM	0,0064	0,9937	0,0078	0,9912
NADAM	0,0305	0,9702	0,0421	0,9527
ADADELTA	0,8675	0,1517	0,7957	0,1075
RMSprop	0,0085	0,9917	0,0069	0,9922
<i>tanh</i>				

ADAM	0,0013	0,9987	0,0017	0,9981
NADAM	0,0033	0,9968	0,0037	0,9959
ADADELTA	0,8152	0,2028	0,8647	0,0301
RMSprop	0,0023	0,9978	0,0027	0,9970

4.2.3. Ispitivanje utjecaja broja epoha i *batch size*-a

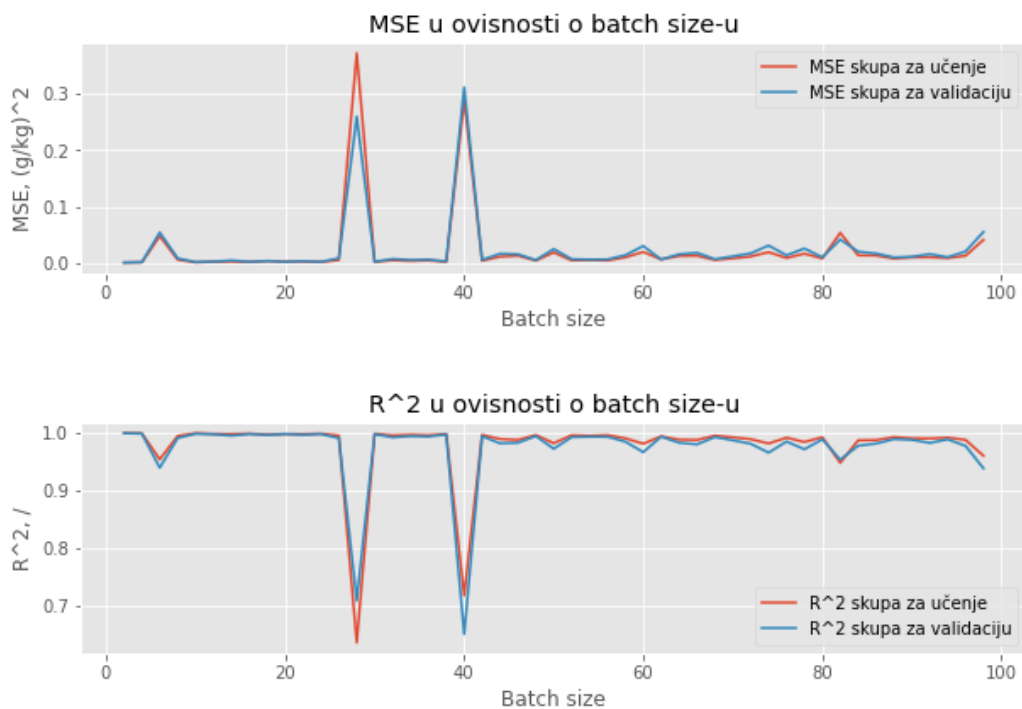
Nakon što je pronađena najbolja kombinacija aktivacijske funkcije, broja neurona u skrivenom sloju te vrste optimizatora, ispitao se utjecaj broja epoha te *batch size*. Rezultati za model s *ReLU* funkcijom s 10 neurona u skrivenome sloju te NADAM optimizatorom, prikazani su na Slikama 4.11. te 4.12.

Na Slici 4.11. prikazan je utjecaj povećanja broja epoha na kriterije uspješnosti *MSE* i R^2 . Na slici je vidljivo kako povećanjem broja epoha od 0 do 20 naglo opada iznos *MSE*, no daljnjim povećanjem broja epoha dolazi samo do manjih promjena u iznosu *MSE* i R^2 . Za izradu modela uzet je broj od 40 epoha jer je za njega dobivena najmanja *MSE*.



Slika 4.11. *MSE* i R^2 u ovisnosti o broju epoha

Na Slici 4.12. dan je grafički prikaz ovisnosti smanjenja *MSE* te povećanja R^2 o povećanju *batch size*-a za skup D2. Može se vidjeti kako povećanjem *batch size*-a u rasponu od 0 do 20 pada i iznos *MSE*, a daljnjim povećanjem *batch size*-a dolazi do oscilacija u iznosima *MSE* i R^2 . Za izradu modela odabran je *batch size* 10 zbog najvećega dobivenog R^2 te najmanje *MSE*.



Slika 4.12. MSE i R^2 u ovisnosti o batch veličini

4.2.4. Najbolji model za skup podataka D2

Nakon provedenih ispitivanja dobiven je model za skup podataka D2 s hiperparametrima prikazanim u *Tablici 4.5.* te kriterijima uspješnosti prikazanim u *Tablici 4.6.* Koeficijent determinacije na sva tri skupa vrlo je visok i blizu jedinici, što potvrđuje odlično slaganje modela i stvarnih podataka, kao i vrlo niske vrijednosti srednje kvadratne pogreške.

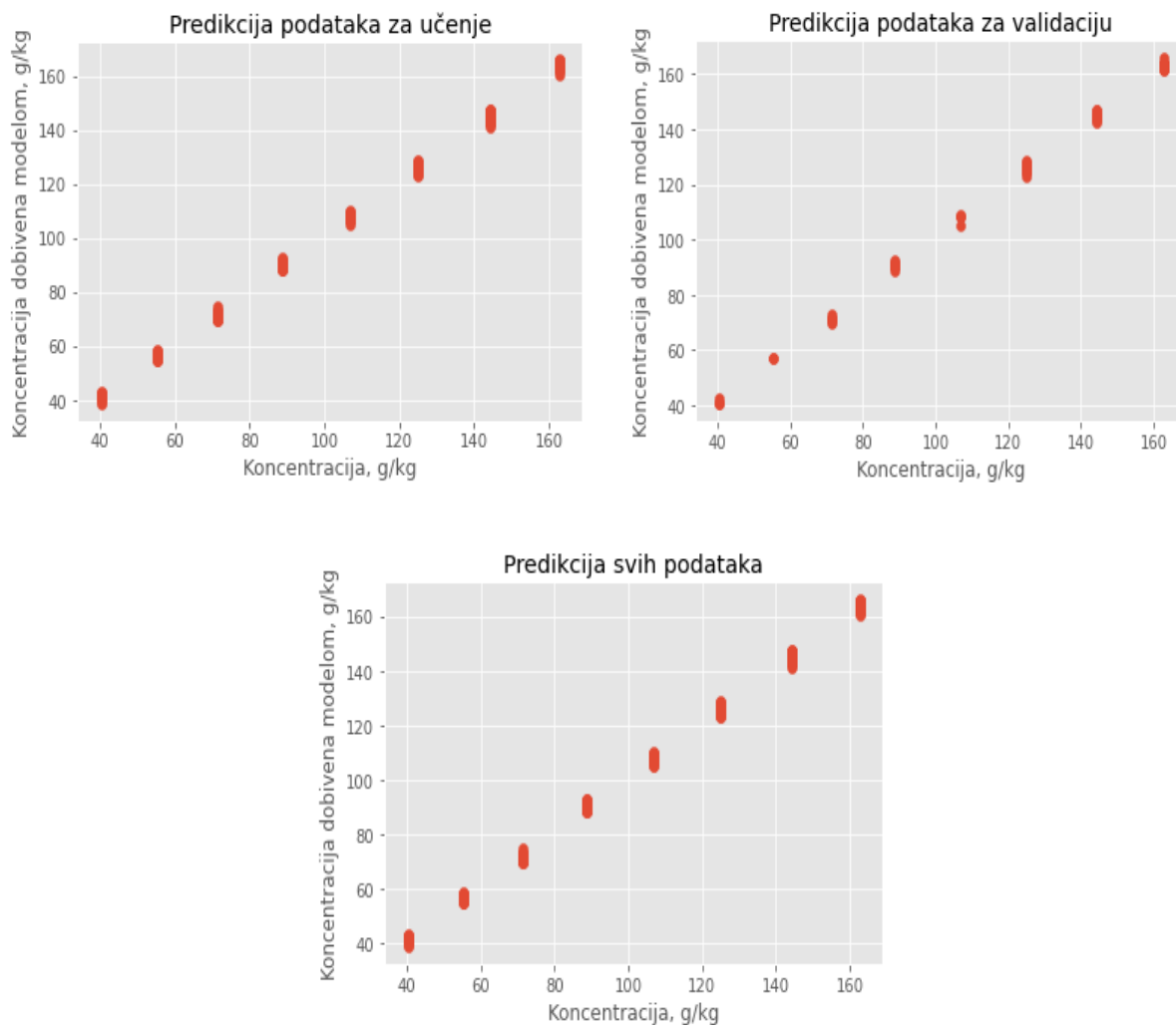
Tablica 4.5. Najbolji model neuronske mreže za skup podataka D2

Broj neurona u skrivenom sloju	Aktivacijska funkcija – skriveni sloj	Aktivacijska funkcija – izlazni sloj	Loss funkcija	Algoritam učenja	Broj epoha	Batch size
10	<i>ReLU</i>	<i>LINEAR</i>	<i>MSE</i>	NADAM	40	10

Tablica 4.6. Kriterij uspješnosti za dobiveni model neuronske mreže za skup podataka D2

R^2 skupa za učenje	R^2 skupa za validaciju	R^2 svi podaci	MSE skupa za učenje	MSE skupa za validaciju
0,9980	0,9979	0,9980	0,0020	0,0019

Usporedbe stvarnih koncentracija i onih previđenih modelom prikazane su na *Slici 4.13.* za sva tri skupa podataka (učenje, validaciju i sve podatke). Ponovno se vidi linearni trend što uz kriterije pogreške i visoke koeficijente determinacije, potvrđuje dobro slaganje modela i stvarnih koncentracija.



Slika 4.13. Usporedba stvarnih koncentracija i onih dobivenih modelom za skup podataka D2 na setu podataka za učenje, validaciju i za cjelokupni set podataka

4.3. Rezultati za skup podataka D3

Kao i kod skupa D1, za skup podataka D3 nije napravljena predobrada podataka, već su korišteni sirovi podaci u rasponu valnih brojeva 840-1800 cm^{-1} . D3 je ustvari skup D1, ali s dodanim skrivenim slojem. Sadrži hiperparametre najboljega modela skupa D1 te je provedeno ispitivanje broja neurona u prvome i drugom skrivenom sloju. Rezultati su prikazani u *tablicama 4.7. i 4.8.* te na *Slici 4.14.*

4.3.1. Najbolji model za skup podataka D3

Nakon provedenih ispitivanja dobiven je model za skup podataka D3 s hiperparametrima prikazanim u *Tablici 4.7.* te kriterijima uspješnosti prikazanim u *Tablici 4.8.* Pomoću *for* petlje određen je optimalan broj neurona u prvome i drugom skrivenom sloju, a ostali hiperparametri određeni su ispitivanjima na skupu podatka D1. Koeficijent determinacije i

vrlo niske vrijednosti srednje kvadratne pogreške upućuju na dobro slaganje modela i eksperimentalnih podataka.

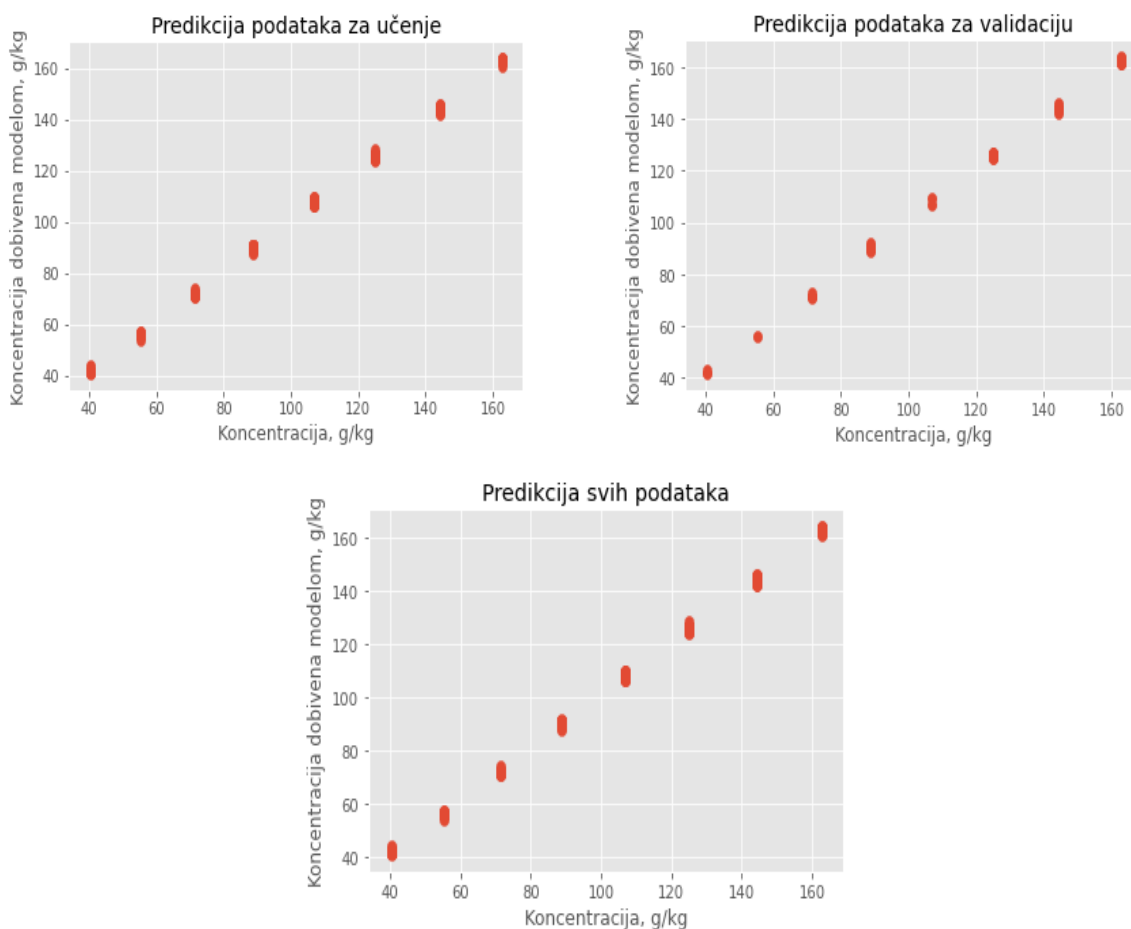
Tablica 4.7. Najbolji model neuronske mreže za skup podataka D3

Broj neurona u 1. i 2. skrivenom sloju	Aktivacijska funkcija – 1. skriveni sloj	Aktivacijska funkcija – 2. skriveni sloj	Aktivacijska funkcija – izlazni sloj	Loss funkcija	Algoritam učenja	Broj epocha	Batch size
20-20	<i>tanh</i>	<i>tanh</i>	<i>Linear</i>	<i>MSE</i>	NADAM	38	6

Tablica 4.8. Kriterij uspješnosti za dobiveni model neuronske mreže za skup podataka D3

R^2 skupa za učenje	R^2 skupa za validaciju	R^2 svi podaci	MSE skupa za učenje	MSE skupa za validaciju
0,9989	0,9985	0,9989	0,0011	0,0014

Na Slici 4.4. prikazane su usporedbe stvarnih koncentracija i onih predviđenim modelom za sva tri seta podataka. Linearni trend, uz kriterije uspješnosti, potvrđuje dobro slaganje modela i stvarnih koncentracija.



Slika 4.14. Usporedba stvarnih koncentracija i onih dobivenih modelom za skup podataka D3 na setu podataka za učenje, validaciju i za cjelokupni set podataka

4.4. Rezultati za skup podataka D4

Na skupu D4 provedena je predobrada podataka MSC te su dobiveni obrađeni spektri kao na *Slici 4.7.* kod skupa D2. Skup D4 dobiven je tako da su uzeti hiperparametri najboljega modela iz D2 skupa, dodan je drugi skriveni sloj te je pomoću *for* petlje nađena najbolja kombinacija broja neurona u prvome i drugom skrivenom sloju.

4.4.1. Najbolji model za skup podataka D4

Nakon određivanja broja neurona u prvom i drugom skrivenom sloju uz prethodno određene hiperparametre na skupu D2, dobiven je model za skup podataka D4 prikazan u *Tablici 4.9.* Prvi skriveni sloj ima 18, dok drugi skriveni sloj ima 20 neurona. U *Tablici 4.10.* prikazani su kriteriji uspješnosti. Koeficijent determinacije na sva tri seta jest kao i kod prethodnih triju skupova podataka vrlo visok, a srednja kvadratna pogreška vrlo niska.

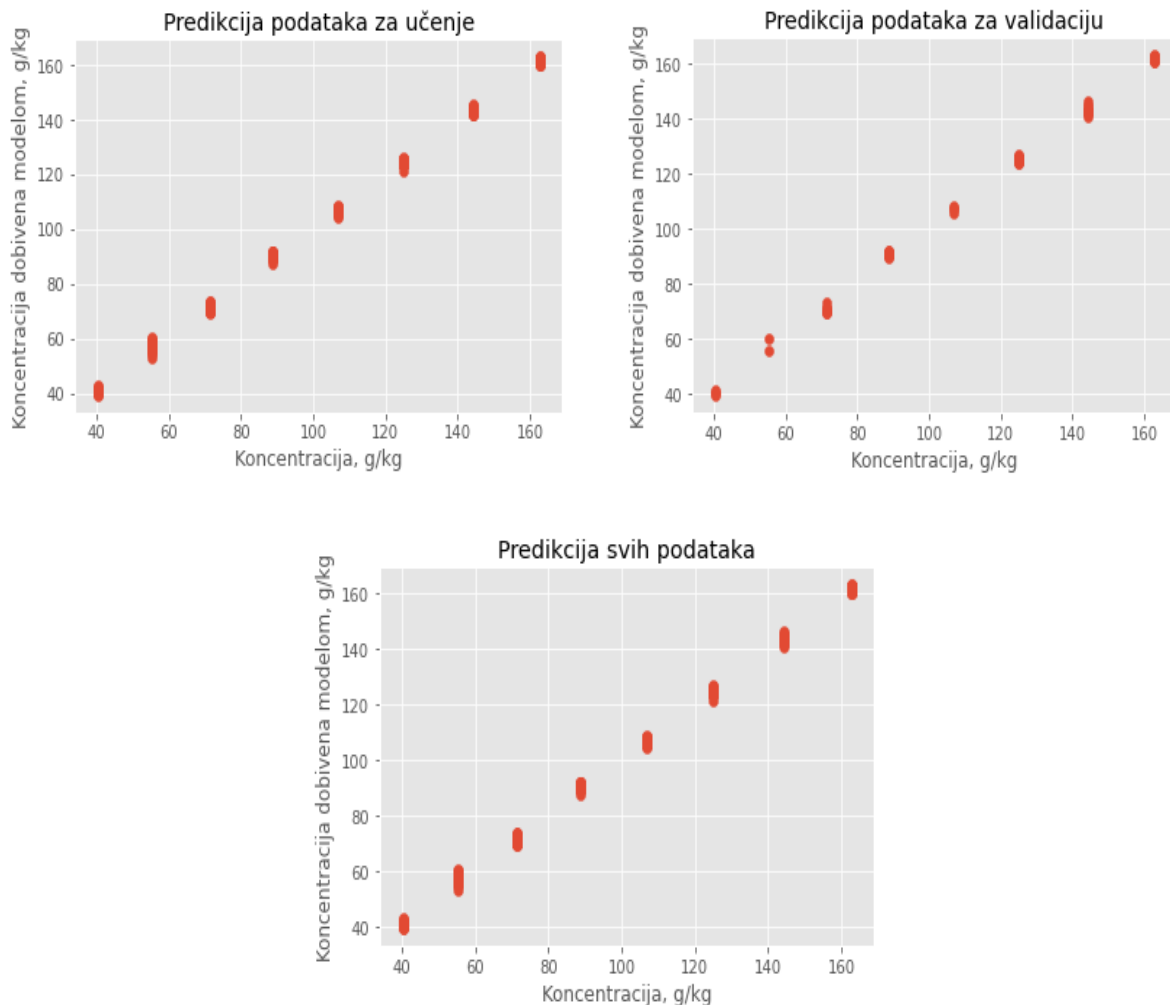
Tablica 4.9. *Najbolji model neuronske mreže za skup podataka D4*

Broj neurona u 1. i 2. skrivenom sloju	Aktivacijska funkcija – 1. skriveni sloj	Aktivacijska funkcija – 2. skriveni sloj	Aktivacijska funkcija – izlazni sloj	Loss funkcija	Algortiam učenja	Broj epoha	<i>Batch size</i>
18-20	<i>ReLU</i>	<i>ReLU</i>	<i>LINEAR</i>	<i>MSE</i>	NADAM	40	10

Tablica 4.10. *Kriterij uspješnosti za dobiveni model neuronske mreže za skup podataka D4*

R^2 skupa za učenje	R^2 skupa za validaciju	R^2 svi podaci	MSE skupa za učenje	MSE skupa za validaciju
0,9984	0,9980	0,9984	0,0016	0,0018

Usporedbe stvarnih koncentracija i onih dobivenih modelom prikazane su na *Slici 4.15.* Kao i kod prethodna tri skupa podataka, jasno se vidi linearan trend što ukazuje na jako dobro slaganje eksperimenta i modela.



Slika 4.15. Usporedba stvarnih koncentracija i onih dobivenih modelom za skup podataka D4 na setu podataka za učenje, validaciju i za cjelokupni set podataka

4.5. Usporedba najboljih modela

U Tablici 4.11. prikazani su dobiveni rezultati, odnosno kriteriji uspješnosti najboljih modela za sva četiri skupa podataka na kojima su provedena ispitivanja. Koeficijent determinacije na svim je skupovima podataka dosta visok, u rasponu vrijednosti 0,9979-0,9989, dok je srednja kvadratna pogreška 0,0020-0,0014. Vrijednosti kriterija uspješnosti na skupu za učenje i na skupu za validaciju vrlo su slične, a kod skupa podataka D3 čak i iste. Iz tih je vrijednosti vidljivo kako je na svim četirima skupovima dobiveno odlično slaganje stvarnih koncentracija te onih određenih modelom. Razlike u vrijednostima između modela za određeni skup podataka su vrlo male te se razlikuju tek u trećoj decimali. Iz toga razloga nema potrebe za razvijanjem kompleksnijeg modela neuronske mreže od onoga s jednim

skrivenim slojem. Također, predobrada podataka u ovoj studiji nije imala velikog utjecaja na konačne rezultate.

Tablica 4.11. *Kriterij uspješnosti najboljih modela neuronske mreže za sve skupove podataka*

R^2 skupa za učenje	R^2 skupa za validaciju	R^2 svi podaci	MSE skupa za učenje	MSE skupa za validaciju
D1				
0,9985	0,9981	0,9985	0,0015	0,0017
D2				
0,9980	0,9979	0,9980	0,0020	0,0019
D3				
0,9989	0,9985	0,9989	0,0011	0,0014
D4				
0,9984	0,9980	0,9984	0,0016	0,0018

5. ZAKLJUČAK

Cilj ovoga rada bio je razviti i ispitati kemometrijske modele primjenom neuronskih mreža za modeliranje ovisnosti koncentracije djelatne tvari, ksilometazolin hidroklorida, u n-butanolu, o temperaturi i spektralnim podacima dobivenih mjerenjima ATR-FTIR spektrometrom. U radu je ispitan utjecaj predobrade spektralnih podataka, analiziran je broj skrivenih slojeva neuronskih mreža, broj neurona u prvome i drugom skrivenom sloju, vrsta primijenjenog algoritma učenja kao i prijenosne funkcije na kvalitetu razvijenih neuronskih mreža.

Ispitivanja su provedena na četiri skupa podataka, D1, D2, D3 i D4, koja su se međusobno razlikovali po predobradi ulaznih podataka te broju skrivenih slojeva. Razvijeni modeli neuronske mreže na sva četiri skupa podataka dali su jako dobre rezultate.

Iznosi koeficijenta determinacije i srednje kvadratne pogreške pokazali su kako predobrada podataka u ovome slučaju nema velikoga utjecaja na konačne rezultate. To se može pripisati preciznome mjerenju i dobro odabranome rasponu valnih brojeva u kojemu je sadržana informacija o sustavu.

Također, istraživanje je pokazalo da prilikom razvoja modela nema potrebe za dodavanjem drugoga skrivenog sloja neuronske mreže, jer prije svega nema velike razlike u krajnjim rezultatima.

Primjena procesne analitičke tehnologije (PAT) u praksi omogućila bi kontinuirano praćenje kvalitete djelatne tvari, a razvoj kalibracijskoga modela ključan je korak pri implementaciji iste. Takvim pristupom izbjegava se zaustavljanje procesa, uzorkovanje i *off-line* analiza, a omogućuje se pravovremeno utvrđivanje procesnih odstupanja kako bi se mogli održati uvjeti koji donose proizvode željene kvalitete. Ovaj bi način praćenja i vođenja procesa rezultirao bržom i kvalitetnijom proizvodnjom, što u konačnici omogućuje farmaceutskoj industriji brže plasiranje proizvoda na tržište te manju vjerojatnost za pojavom pogreške.

6. LITERATURA

- [1] A. Sander, Jedinične operacije u ekoinženjerstvu I dio, Zagreb, 2011, str. 80-113
- [2] T. Jednačak i P. Novak: Procesne analitičke tehnike temeljene na vibracijskoj spektroskopiji in-line, Kem. Ind. 62 (3-4) 71–80 (2013)
- [3]https://en.wikipedia.org/wiki/Fouriertransform_infrared_spectroscopy#Attenuated_total_reflectance (pristup 20.2.2022.)
- [4] Lawrence X. Yu, Robert A. Lionbergera i sur., Applications of process analytical technology to crystallization processes, Advanced Drug Delivery Reviews 56 (2004) 349–369
- [5] E. Simone, Application of process analytical technology (PAT) tools for the better understanding and control of the crystallization of polymorphic and impure systems, Loughbourgh University, 2015
- [6] N. Bolf, H. Dorić, Procesna analitička tehnologija, Kem. Ind. 67 (9-10) (2018) 466–469
- [7] Fangkun Zhang i sur., Comparative study on ATR-FTIR calibration models for monitoring solution concentration in cooling crystallization, Journal of Crystal Growth Volume 459, 1 February 2017, str. 50-55
- [8] A. Borissova i sur., In Situ Measurement of Solution Concentration during the Batch Cooling Crystallization of L-Glutamic Acid using ATR-FTIR Spectroscopy Coupled with Chemometrics, Crystal Growth & Design 2009 Vol. 9, No. 2 692–706
- [9] Ž. Ujević Andrijić, Umjetne neuronske mreže, Kem. Ind. 68 (5-6) (2019) 219–220
- [10] L. Perlovsky, Introduction to deep learning, Harvard University, USA, 2018
- [11] N. Bolf, Strojno učenje, Kem. Ind. 70 (9-10) (2021) 591–593
- [12] S. Raschka, Python Machine Learning, Packt Publishing Ltd., UK, 2015.
- [13] Ž. Ujević Andrijić, N. Bolf, Metode umjetne inteligencije u kemijskom inženjerstvu, Umjetne neuronske mreže, Zagreb, 2021.
- [14] C. Cui, T. Fearn, Modern practical convolutional neural networks for multivariate regression: applications to NIR calibration, Chemometrics and Intelligent Laboratory Systems, 2018.

[15] Ž. Ujević Andrijić, Metode umjetne inteligencije u kemijskom inženjerstvu, Statistička analiza i predobrada podataka, validaciju rezultata, Zagreb, 2021.

[16] <https://www.medchemexpress.com/xylometazoline-hydrochloride.html> (pristup 7.6.2022.)

[17] [https://en.wikipedia.org/wiki/Python_\(programming_language\)](https://en.wikipedia.org/wiki/Python_(programming_language)) (pristup 20.6.2022.)

[18] <https://www.spyder-ide.org/> (pristup 20.6.2022.)

[19] https://keras.io/guides/sequential_model/ (pristup 20.6.2022.)

[20] Ž. Ujević Andrijić, N. Rimac, Metode umjetne inteligencije u kemijskom inženjerstvu Primjer razvoja i vrednovanja modela neuronskih mreža, Vodič kroz razvoj neuronskih mreža u Pythonu, Zagreb, 2021.

7. POPIS SIMBOLA

Δc	prezasićenost
c	koncentracija
c^*	ravnotežna koncentracija
S	omjer prezasićenosti
σ	relativna prezasićenost
x_{norm}^i	normalizirana i-ta vrijednost značajki
x^i	i-ta vrijednost značajki prije skaliranja
x_{min}	najmanja vrijednost značajke skupa za učenje
x_{max}	najveća vrijednost značajke skupa za učenje
x_{std}^i	standardizirana i-ta vrijednost značajki
x^i	nestandardizirana i-ta vrijednost značajki
μ_x	srednja vrijednost značajki skupa za učenje
σ_x	standardna devijacija značajki skupa za učenje
w_{ij}^{novi}	nova vrijednost težinskog faktora neurona i povezanog s neuronom j
w_{ij}^{stari}	stara vrijednost težinskog faktora neurona i povezanog s neuronom j
Δw_{ij}	korektivni faktor
η	koeficijent učenja
y_i	izlaz iz prethodnog neurona (neurona i)
ε_j	pogreška slijedećeg neurona (neurona j)
E	globalna pogreška mreže
k	broj neurona u izlaznom sloju
d_k	željena vrijednost, dobivena eksperimentom
y_k	vrijednost dobivena neuronskom mrežom
MSE	srednja kvadratna pogreška
R^2	koeficijent determinacije
y_i	izmjerena vrijednost
\hat{y}_i	modelom procijenjena vrijednost
\bar{y}	srednja vrijednost mjerene veličine
n	broj podataka

