

# Optimalno navodnjavanje pšenice primjenom umjetne inteligencije

---

Slijepčević, Edi

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:168:934337>

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

Download date / Datum preuzimanja: **2025-03-13**



Repository / Repozitorij:

[FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repository](#)



SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 60

**OPTIMALNO NAVODNJAVANJE PŠENICE PRIMJENOM  
UMJETNE INTELIGENCIJE**

Edi Slijepčević

Zagreb, lipanj 2024.

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 60

**OPTIMALNO NAVODNJAVANJE PŠENICE PRIMJENOM  
UMJETNE INTELIGENCIJE**

Edi Slijepčević

Zagreb, lipanj 2024.

## DIPLOMSKI ZADATAK br. 60

Pristupnik: **Edi Slijepčević (0036518436)**  
Studij: Informacijska i komunikacijska tehnologija  
Profil: Automatika i robotika  
Mentor: izv. prof. dr. sc. Vinko Lešić

Zadatak: **Optimalno navodnjavanje pšenice primjenom umjetne inteligencije**

### Opis zadatka:

Uz izražene globalne trendove urbanizacije, sve veći fokus stavlja se na kvalitetu prehrane kao jedan od osnovnih čimbenika za povećanje kvalitete života. Današnja tehnologija omogućava značajne mogućnosti u unaprjeđenju uzgoja biljaka osiguravanjem povoljnih klimatskih uvjeta u stakleniku, uzgoj na inače nepovoljnim lokacijama, produljenu sezonu te biološki uzgoj izoliran od zaraza i prema optimalnim recepturama mikroklimе za uzgoj biljaka. U radu je potrebno istražiti i razviti model razvoja pšenice u ovisnosti o vlažnosti tla prema identifikaciji optimalne recepture navodnjavanja pšenice. Pri tome je potrebno obraditi već dostupni skup podataka, provesti naknadnu korekciju prikupljenih mjerenja te odrediti površinu biljke i količinu biomase tijekom vremena iz vremenskog slijeda fotografija. Na temelju obrađenih podataka, potrebno je načiniti model ovisnosti rasta biljke u korelaciji s vlažnosti tla i potencijalnim drugim uvjetima primjenom povratnih neuronskih mreža. Načinjeni model potrebno je testirati na ranije dostupnim stvarnim podacima i uz primjenu vremenske prognoze za odabrano podneblje.

Rok za predaju rada: 28. lipnja 2024.



## Sadržaj

<b>Uvod .....</b>	<b>1</b>
<b>1. Podatci o uvjetima rasta i razvoja pšenice.....</b>	<b>2</b>
1.1. Laboratorijski podatci o mikroklimatskim uvjetima pšenice.....	3
1.2. Mjerenja na terenu .....	7
1.3. Remote – sensing pristup.....	8
<b>2. Obrada laboratorijskih podataka o uvjetima rasta i razvoja pšenice .....</b>	<b>11</b>
2.1. Obrada i segmentiranje fotografija .....	11
2.2. Izračun površine biljke i normaliziranog razlikovnog vegetacijskog indeksa (NDVI) .....	15
2.3. Uparivanje obrađenih podataka o NDVI i površini pšenice s laboratorijskim mjerenjima.....	19
2.4. Analiza uparenih senzorskih podataka i vrijednosti NDVI-a i površine pšenice u pikselima .....	24
<b>3. Predviđanje NDVI-a i površine pšenice korištenjem modela strojnog učenja .</b>	<b>27</b>
3.1. Predprocesiranje podataka .....	27
3.2. Predviđanje NDVI-a i površine pšenice korištenjem modela linearne regresije .	27
3.3. Predviđanje NDVI-a i površine pšenice korištenjem Lasso modela .....	29
3.4. Predviđanje Multi Layer Perceptron (MLP) modelom .....	30
<b>4. Predviđanje površine i NDVI-a pšenice korištenjem povratne neuronske mreže .....</b>	<b>31</b>
4.1. Predprocesiranje podataka .....	31
4.1.1. Prozorska segmentacija podataka .....	32
4.1.2. Standardizacija podataka .....	33
4.2. Povratna neuronska mreža za predviđanje NDVI-a i površine biljke.....	34
4.2.1. LSTM (Long Short-Term Memory) .....	34
4.2.2. Arhitektura i korišteni slojevi u neuronskoj mreži.....	35

4.2.3.	Hiperparametri korišteni u modelu za predikciju NDVI-a i površine pšenice.....	38
4.3.	<b>Rezultati predviđanja NDVI-a i površine pšenice korištenjem povratne neuronske mreže.....</b>	<b>39</b>
5.	<b><i>Usporedba rezultata svih korištenih modela u predikciji NDVI-a i površine u pikselima .....</i></b>	<b>42</b>
6.	<b><i>Testiranje naučenih modela na stvarnim podacima.....</i></b>	<b>45</b>
6.1.	Testiranje LSTM modela na stvarnim podacima s terena .....	45
6.1.	Testiranje Lasso modela na stvarnim podacima s terena .....	47
6.2.	Usporedba rezultata Lasso i LSTM modela .....	47
7.	<b><i>Testiranje utjecaja različitih režima navodnjavanja korištenjem LSTM modela .....</i></b>	<b>48</b>
	<b><i>Zaključak .....</i></b>	<b>50</b>
	<b><i>Literatura .....</i></b>	<b>51</b>
	<b><i>Sažetak.....</i></b>	<b>53</b>
	<b><i>Summary .....</i></b>	<b>54</b>

# Uvod

Uz sve izraženiji utjecaj klimatskih promjena te porasta stanovništva stavlja se sve veći pritisak na poljoprivredu kao najbitnijeg i trenutno najugroženijeg gospodarskog sektora. Ovakav trend najviše će se odraziti na intenzivan uzgoj monokultura kao što su pšenica, soja, kukuruz te smanjiti i u ekstremnim slučajevima posve uništiti prinose. Kao jedan od najučestalijih vremenskih nepogoda suša se pokazala kao i jedna od rijetkih koje se mogu ublažiti te spasiti urod. Umjetna inteligencija ima veliki potencijal u očuvanju prirodnih resursa poput vode i kvalitete tla, te korištenjem automatiziranih sustava moguće je uštediti i do 95% vode za navodnjavanje [1]. Kako bi se estimiralo zdravstveno stanje pšenice najčešći pokazatelj je normalizirani razlikovni indeks (eng. Normalized Difference Vegetation Index, NDVI) s kojim se može estimirati utjecaj suše te adekvatno smanjiti njen utjecaj. Bitan parametar u uzgoju pšenice je i nadzemna masa biljke koju je cilj povećati ili barem očuvati uslijed utjecaja vremenskih nepogoda. Ovaj rad detaljno će obraditi procese uključene u analizu podataka, počevši od segmentacije slika radi dobivanja ciljanih parametara do uparivanja sa senzorskim mjerenjima. Dalje, rad će istražiti primjenu modela linearne regresije i neuronskih neuronske mreža za predviđanje NDVI-a i površine biljke.. Naposljetku, rad će se fokusirati na testiranje naučenog modela na stvarnim terenskim podacima. To uključuje prikupljanje mjerenja korištenjem remote-sensing pristupa pomoću satelitskih snimki polja za prikupljanje stvarnih podataka i vremenskih prognoza te korištenje ranije dostupnih podataka prikupljenih na terenu. Također, biti će testirani režimi navodnjavanja kako vi se pokazao njihov utjecaj na predikciju NDVI-a. Kroz ove korake, istraživanje će doprinijeti boljem razumijevanju i implementaciji naprednih tehnologija u poljoprivredi, omogućujući precizniju analizu rasta pšenice i optimizaciju poljoprivredne proizvodnje.



# 1. Podatci o uvjetima rasta i razvoja pšenice

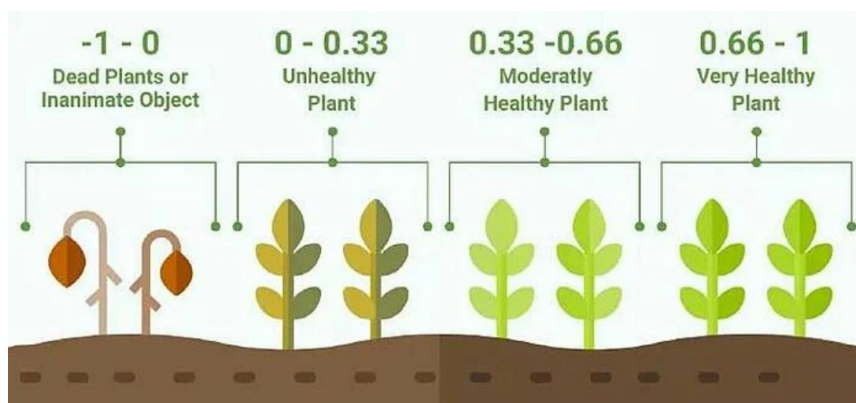
Podaci o rastu i razvoju pšenice prikupljeni su u sklopu AgroSPARC projekta provedenom u vremenskom razdoblju između 2022. i 2023. godine, pri čemu su se prikupljali podaci o uvjetima rasta i razvoja pšenice te snimale fotografije samih biljaka kako bi se kasnije odredili parametri razvoja pšenice u vidu vegetacijskih indeksa i biomase koja je u svrhu ovog rada kao parametar zamijenjena površinom biljke u pikselima [1].

Vegetacijski indeksi biljaka kvantitativni su pokazatelji zdravstvenog stanja, pokrova i produktivnosti vegetacije putem daljinskog istraživanja. NDVI se najčešće koristi zbog svoje jednostavnosti i učinkovitosti. NDVI se računa pomoću reflektiranog zračenja u crvenom (RED) i blisko infracrvenom (NIR) spektralnom području, prema izrazu (1).

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

NDVI vrijednosti se kreću od -1 do +1. Više NDVI vrijednosti (bliže +1) ukazuju na gustu i zdravu vegetaciju s visokim stupnjem fotosinteze, dok niže vrijednosti (bliže 0) ukazuju na rijetku vegetaciju, stresirane biljke ili nevegetirane površine. Negativne vrijednosti obično označavaju vodene površine, oblake ili sjene.

Na slici 1.1 je vizualna reprezentacija koja pokazuje vrijednosti NDVI-a u odnosu na zdravstvenog stanja biljke.



Sl. 1.1 Reprezentacija NDVI vrijednosti u odnosu na zdravlje biljke

Primjene NDVI:

- **Poljoprivreda:** Praćenje usjeva za otkrivanje bolesti, stresa uslijed suše ili potrebe za gnojivom.
- **Ekologija i konzervacija:** Procjena bioraznolikosti i praćenje promjena u ekosustavima.
- **Šumarstvo:** Procjena zdravlja šuma i otkrivanje deforestacije.
- **Upravljanje zemljištem:** Planiranje uporabe zemljišta i procjena degradacije zemljišta.

Površina biljke ili biomasa je dobar pokazatelj razvoja biljke te je cilj da porast biomase pšenice bude što veći kako bi i prinos te kulture bio zadovoljavajuć. Veliku ulogu u optimizaciji uzgoju pšenice ima adekvatno navodnjavanje, te je najčešći problem nedostatak vode u sušnim razdobljima. Posebno je bitno da pšenica ima dovoljnu količinu dostupne vode u razdoblju razvoja lista, stabljike te punjenje zrna.

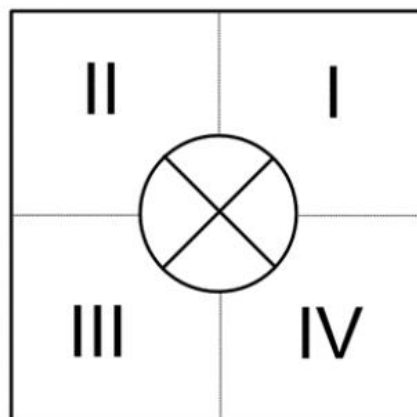
## **1.1. Laboratorijski podatci o mikroklimatskim uvjetima pšenice**

Sušna razdoblja su posebno problematična zbog česte nepredvidljivosti vremenskih prilika općenito te je vrlo bitno kompenzirati nedostatak padalina sa preciznim navodnjavanjem u pravom trenutku. Laboratorijski podatci o mikroklimatskim uvjetima u kojima je pšenica uzgajana prikupljeni su u komorama koje simuliraju stvarne uvjete u kojima pšenica raste. Svaka od četiri korištene komore imaju četiri mikroklimatske zone te su mjerenja prikupljena u svakoj zoni. Jedna od komora sa glavnim pripadnim dijelovima prikazana je na slici 1.2.



Sl. 1.2 Komora s glavnim pripadnim dijelovima [1]

Podjela komore na mikroklimatske zone prikazana je na slici 1.3. U svakom spremniku sa zemljom nalazio se analogni kapacitivni senzor vlage prikazan na slici 1.4.



Sl. 1.3 Prikaz podjele komore na 4 mikroklimatske zone[1]



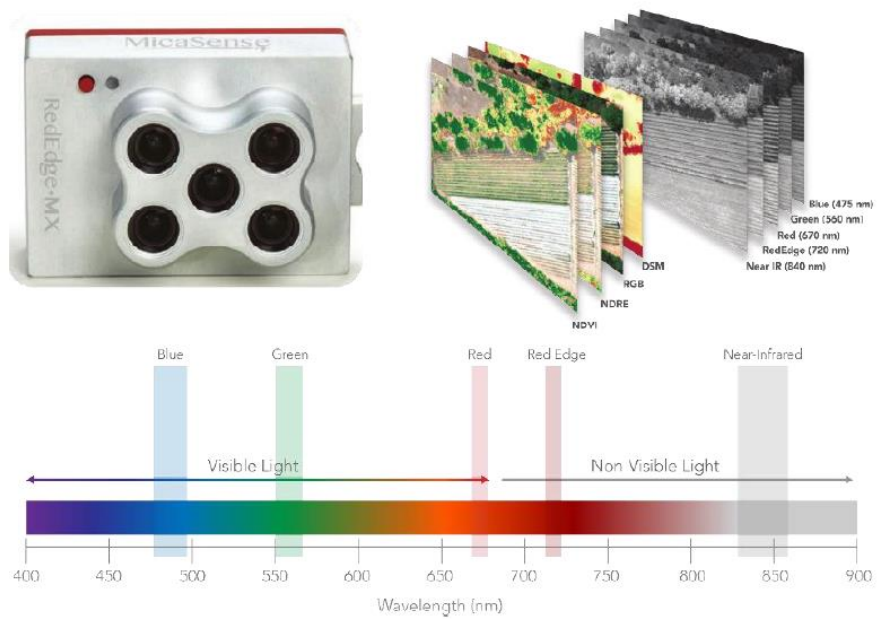
Sl. 1.4 Analogni senzor vlage

Podatci o temperaturi i vlažnosti zraka prikupljeni su s DHT22 senzorima postavljenim na senzorskim „prstenima“ na slici 1.2. Ovaj kapacitivni senzor uz senzor vlažnosti tla je služio i kao povratna veza temperature i vlažnosti zraka u sustavu upravljanja mikroklimatskim uvjetima. Senzor je prikazan na slici 1.5.



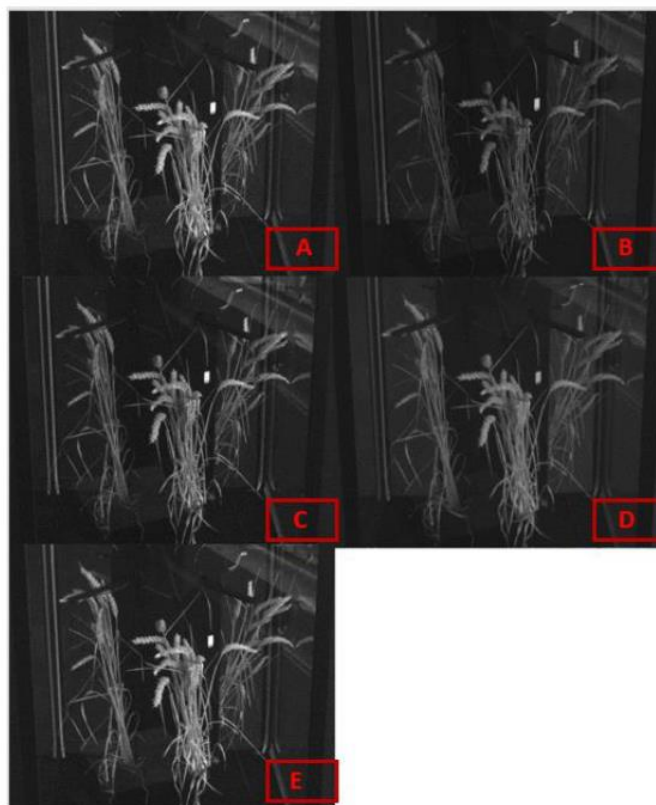
Sl. 1.5 DHT22 senzor

Fotografije komora pomoću kojih se izračunavaju vegetacijski indeksi i površina u pikselima prikupljene su RedEdge-MX senzorom te je njegova spektralna rezolucija prikazana na slici 1.6.



Sl. 1.6 Spektralna rezolucija RedEdge-MX senzora [2]

Prikaz fotografija u svih pet spektara prikazan je na slici 1.7.



Sl. 1.7 Multispektralna slika biljke, A – blue , B – green, C – red, D – near infrared , E - red [1]

## 1.2. Mjerenja na terenu

Mjerenja na terenu su ključna za točnost predviđanja modela a posebice vlažnost tla koju je najpreciznije mjeriti na samom polju gdje se i uzgaja pšenica. Kako bi se naučeni model mogao testirati na podacima sa terena bitno je imati podatke o fenozazi pšenice te preračunati mililitre (iz skupa podataka na kojem se treniralo) u milimetre po metru kvadratnom što je uobičajena jedinica za padaline. Za površinu zemlje u posudama uzeta je vrijednost od 150 cm<sup>2</sup> te je količinu vode u mililitrima potrebno podijeliti sa faktorom 15 kako bi se dobila količina padalina koja odgovara kiši.

S druge strane vrijednosti NDVI-a dobivaju se spektrometrom koji se uobičajeno koristi u agronomiji tako da se mjeri reflektirana svjetlost od površine biljke u različitim spektrima te se kasnije izračunavaju vegetacijski indeksi.

Podaci s terena prikupljeni su na lokaciji Odranski Obrež (Sl 1.8) u blizini grada Zagreba te uključuju mjerenja s meteorološke stanice (padaline u mm/m<sup>2</sup>, temperaturu i vlažnost zraka) te mjerenja tla što uključuje vlažnost tla na dubini od 15 cm.



Sl 1.8 Satelitska snimka lokacije od interesa

### 1.3. Remote – sensing pristup

NDVI vrijednosti također se mogu prikupiti sa satelita koristeći Copernicus Browser. Prikupljena mjerenja se nakon toga mogu upariti sa vremenskim prognozama prikupljenih na Open – Meteo platformi te stvarnim vremenskim prilika na terenu i kombinacijom ovih parametara dobiti novi model koji bi učinkovitije predviđao NDVI.

Podatci sa Open-Meteo platforme nisu pronađeni za testni skup od 23.5.2023 – 07.07.2023. te je odabran deterministički pristup prognozi što ujedno smanjuje grešku predikcije. Na slici 1.9 prikazana je Open-Meteo platforma sa upisanim podacima o lokaciji za koju se provodi testiranje modela.

Podatci o NDVI-u u sklopu Copernicus Browsera prikupljeni su sa Sentinel – 2 satelitom te su dobivene vrijednosti NDVI-a za lokaciju od interesa prikazani na slici 4.2

Location and Time

Location: **Coordinates** List

Latitude: 45.735

Longitude: 15.9516

Time: Forecast Length Time Interval

Start Date: 2023-05-23

End Date: 2023-07-07

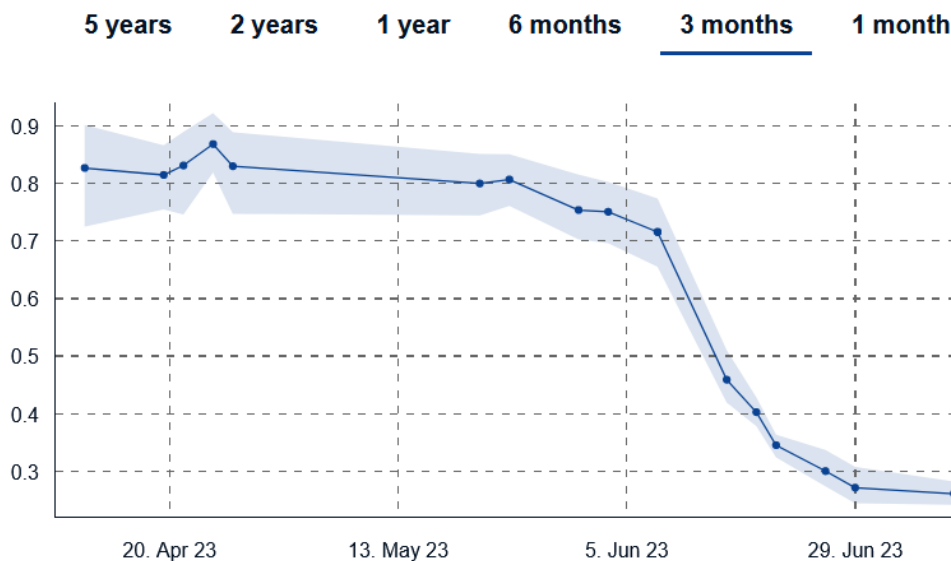
Hourly Weather Variables

- Temperature (2 m)
- Relative Humidity (2 m)
- Dewpoint (2 m)
- Apparent Temperature
- Precipitation Probability
- Precipitation (rain + showers + snow)
- Rain
- Showers
- Snowfall
- Snow Depth
- Weather code
- Sealevel Pressure
- Surface Pressure
- Cloud cover Total
- Cloud cover Low
- Cloud cover Mid
- Cloud cover High
- Visibility
- Evapotranspiration
- Reference Evapotranspiration (ET<sub>0</sub>)
- Vapour Pressure Deficit

Sl. 1.9 Open – Meteo platforma

## Sentinel-2 L2A - NDVI

● — ☁ 4% ✕



📄 Export CSV

Sl. 1.10 Prozor u Copernicus Browser-u koji prikazuje vrijednosti NDVI-a za odabranu lokaciju

Kao što je vidljivo na slici 1.10 postoji mogućnost podešavanja pokrivenosti oblaka koji su inače smetnja u snimanju površine zemlje te je vrlo bitno odabrati adekvatan prag jer ako odaberemo preveliki prag nećemo imati pouzdan prikaza vrijednosti što je također slučaj ako odaberemo preveliku vrijednost.

Slika 1.11 prikazuje problematiku postavljanja previsokog praga pokrivenosti oblacima.



# Sentinel-2 L2A - NDVI

100% ×



[Export CSV](#)

Sl. 1.11 Rezultat odabira prevelikog praga za pokrivenost oblaka

## **2. Obrada laboratorijskih podataka o uvjetima rasta i razvoja pšenice**

Dohvaćanje fotografija podataka iz baze podataka je izvršeno pomoću WinSCP programa te su za pojedine komore ovisno o orijentaciji odabrane fotografije snimane ili ujutro(orijentacija prema zapadu) ili poslijepodne(orijentacija prema istoku) zbog utjecaja refleksije na staklenoj stjenki komore.

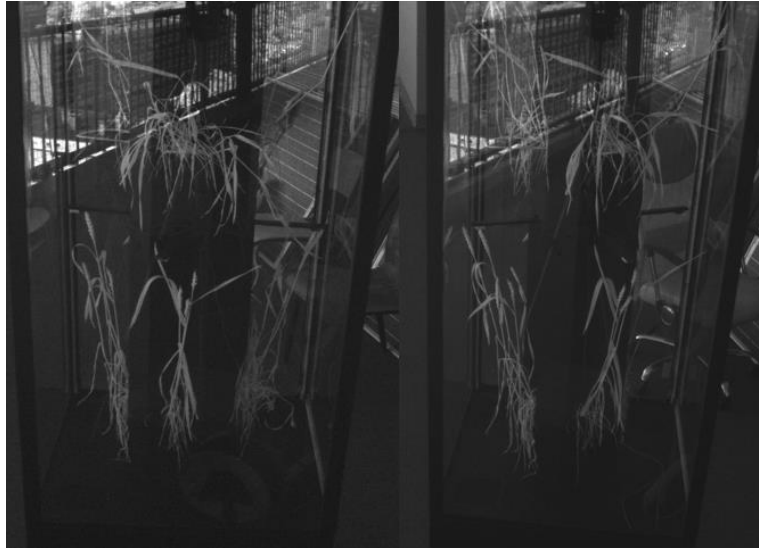
Obrada fotografija za izračun NDVI indeksa i površine biljke u pikselima izvršena je u programu MATLAB, za uparivanje tih vrijednosti sa ostalim mjerenim parametrima te daljnju obradu korišten je Jupyter Notebook te *pandas* biblioteka zbog svoje jednostavnosti upravljanja skupovima podataka.

### **2.1. Obrada i segmentiranje fotografija**

Prvi korak bio je izdvojiti tri spektra od interesa u odvojene datoteke radi preglednosti. Kao što je ranije navedeno spektri od interesa su RED i NIR te RE (eng. red egde) zbog velikog kontrasta s pozadinom, lakše segmentacije te na kraju izračuna površine biljke.

Slijedeći korak bio je olakšati segmentaciju fotografija izdvajanjem ciljane regije od interesa te ostatak piksela fotografije postaviti u crnu boju kako pozadina i ostale biljke na slici utjecale na segmentiranje.

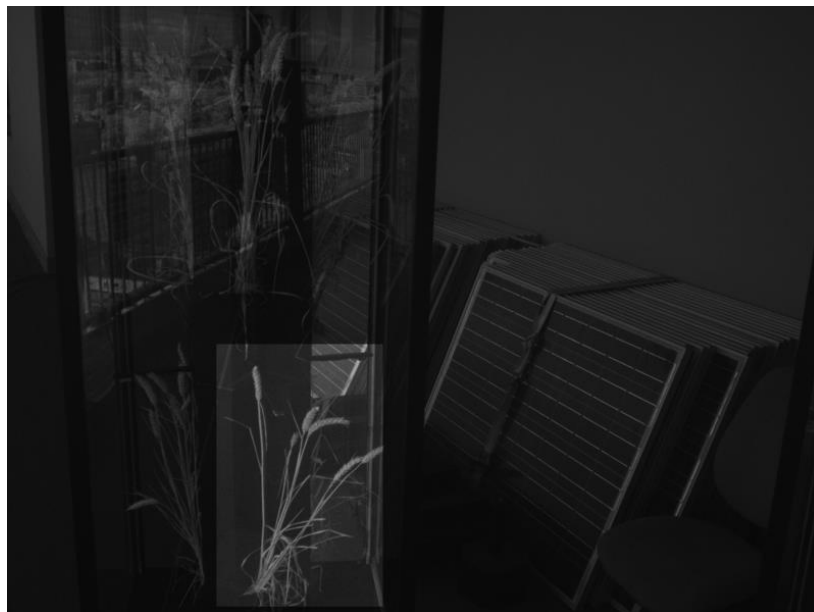
Inicijalni pristup ovom problemu bio je iteriranje kroz datoteku te interaktivno crtanje pravokutnika na svakoj fotografiji. Ovaj pristup se činio najpraktičniji s obzirom na broj fotografija i kvalitetu snimanja (česte promjene kuteva slikanja biljke zbog trošenja zupčanika u sustavu za okret kamere i komore) prikazane na slici 2.1.



Sl. 2.1 Prikaz promjene kuta slikanja komore

Na slici 2.1 također se vidi veliki utjecaj refleksije koja je otežavala segmentaciju a za gornju polovicu svake komore u potpunosti je onemogućila.

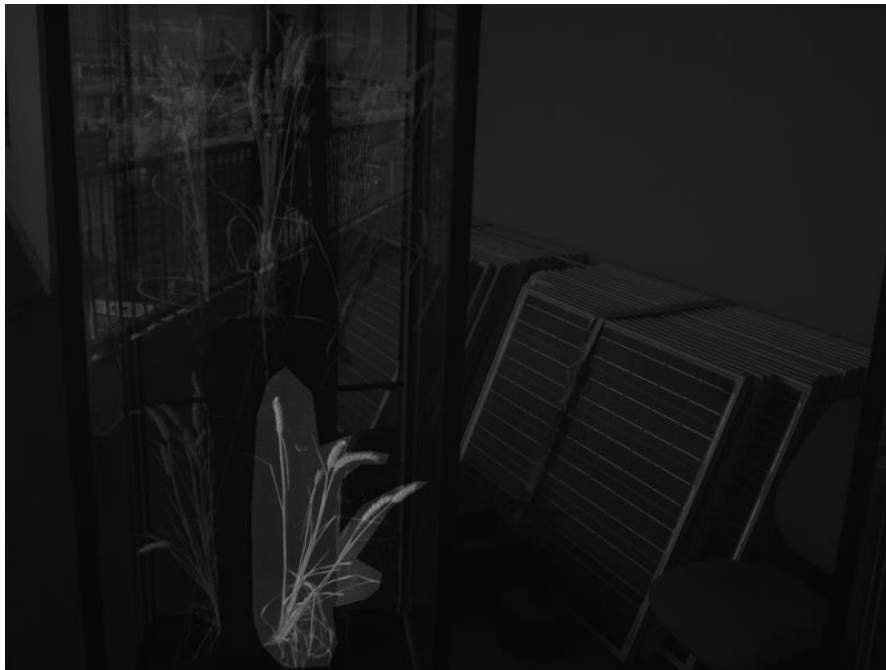
Na slici 2.2 je prikazan rezultat pripreme fotografije za segmentiranje gdje su spojena izvorna i obrađena fotografija za segmentaciju.



Sl. 2.2 Inicijalna metoda eliminacije utjecaja pozadine

Ovaj pristup se pokazao neadekvatnim iz razloga što su biljke gusto raspoređene te je značajan utjecaj ostalih smetnji kao što su refleksija od susjednih površina i listovi biljaka u drugim zonama.

Drugi pristup koji uključuje crtanje regije od interesa sa proizvoljnim brojem točaka se pokazao puno boljim iz razloga što je moguće puno preciznije eliminirati pozadinske smetnje svake fotografije pšenice kao što se da zaključiti iz slike 2.3.



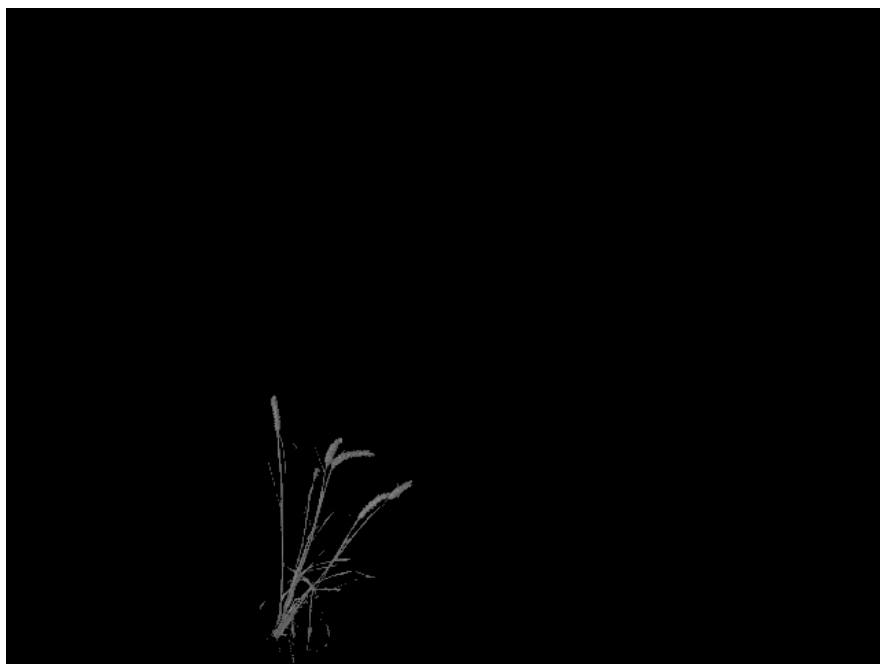
Sl. 2.3 Odabrani pristup eliminacije pozadinskih smetnji za segmentiranje biljke

Za segmentaciju fotografija biljaka korištena je MATLAB aplikacija Image Segmenter koja služi za ekstrakciju piksela fotografije koji odgovaraju površini biljaka.

Koraci segmentiranja pomoću aplikacije Image Segmenter su sljedeći:

- **Učitavanje slike:** Prvo, aplikacija učitava sliku na kojoj se nalazi biljka.
- **Primjena thresholding-a:** Thresholding je tehnika koja koristi određenu vrijednost praga da bi se odlučilo koji pikseli pripadaju biljci, a koji pozadini. Na primjer:
  - Svi pikseli čija je vrijednost veća od praga se smatraju kao dio biljke.
  - Svi pikseli čija je vrijednost manja ili jednaka pragu se smatraju kao dio pozadine.

- **Stvaranje binarne maske:** Rezultat thresholding-a je binarna maska, gdje pikseli koji pripadaju biljci imaju vrijednost 1 (ili bijelo), a pikseli koji pripadaju pozadini imaju vrednost 0 (ili crno).
- **Primjena maske na originalnu sliku:** Konačno, binarna maska se koristi za izdvajanje biljke sa originalne slike. To se radi tako što se originalna slika pomnoži sa binarnom maskom, zadržavajući samo piksele biljke, dok se svi ostali pikseli (pozadina) postavljaju na crnu (ili neku drugu boju pozadine) kao što je vidljivo na slici 2.4.



Sl. 2.4: Prikaz fotografije na koju je primjenjena maska  
dobivena aplikacijom Image Segmenter

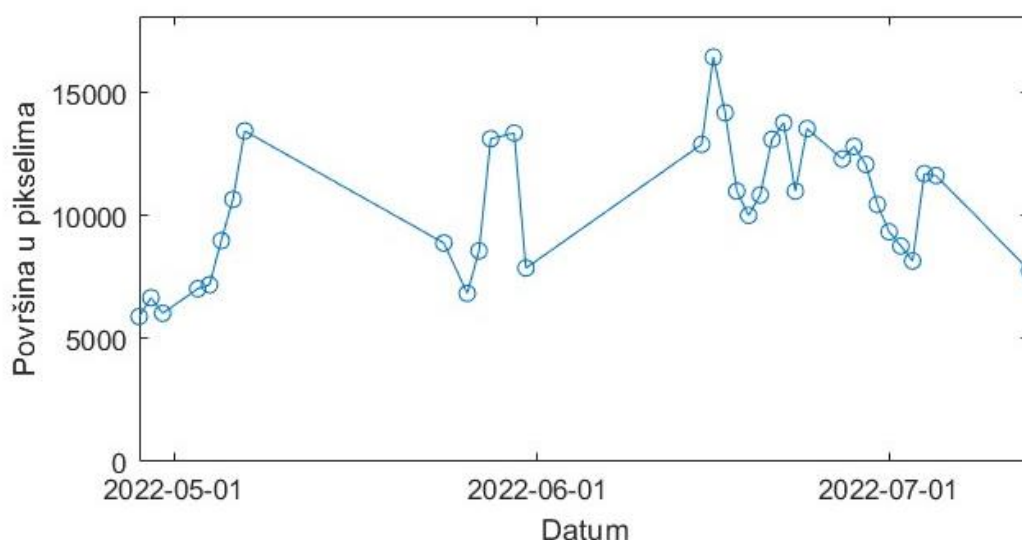
U vidu ubrzanja procesa segmentiranja na skupu fotografija u aplikaciji Image Segmenter moguće je odabrati i generiranje funkcije koja se kasnije koristi u petlji koja iterira po datoteci u kojoj su fotografije.

Ukoliko se obrađuje veći skup fotografija sniman u dužem periodu potrebno je uzeti u obzir mijenjanje osvjetljenja (fotografije snimane od početka svibnja do sredine srpnja) uzrokovane naoblakom ili mijenjanjem položaja sunca, pa tako i odabradi adekvatni prag po kojem će se vršiti segmentiranje.

## 2.2. Izračun površine biljke i normaliziranog razlikovnog vegetacijskog indeksa (NDVI)

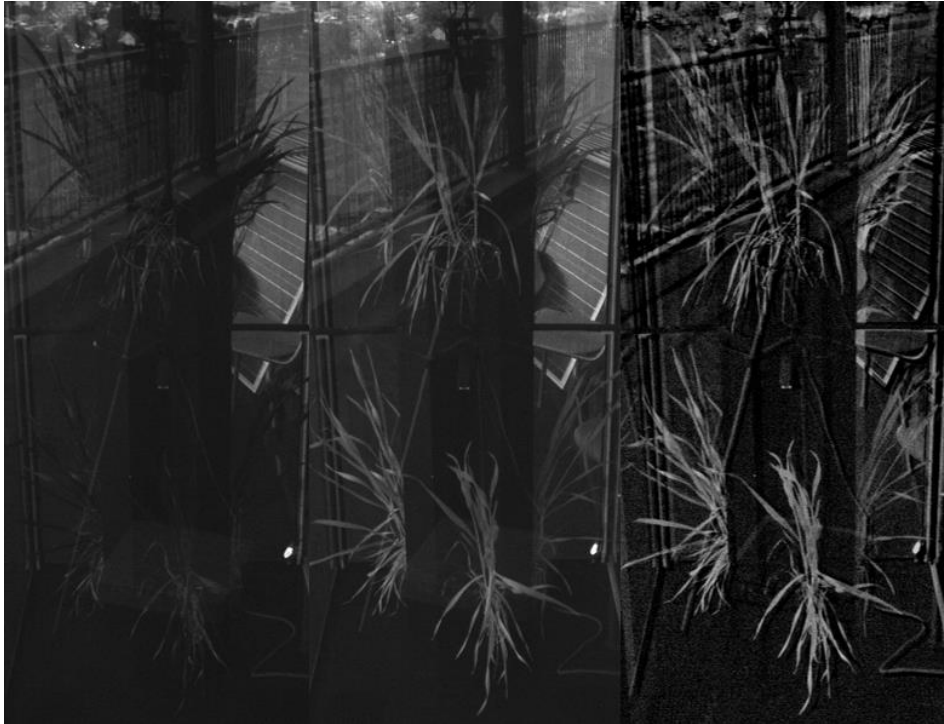
Površina u pikselima je izračunata tako da se izračunala suma jedinica iz maske.

Odabir praga je izvršen tako da se primjenio jedan prag za sve fotografije. Ovaj pristup se pokazao neadekvatan zbog ranije spomenutih problema u osvjetljenju. Rezultat su vrlo neprecizne vrijednosti koje nerealno prikazuju promjene NDVI i površine biljke kroz vrijeme što se vidi na slici 2.5.

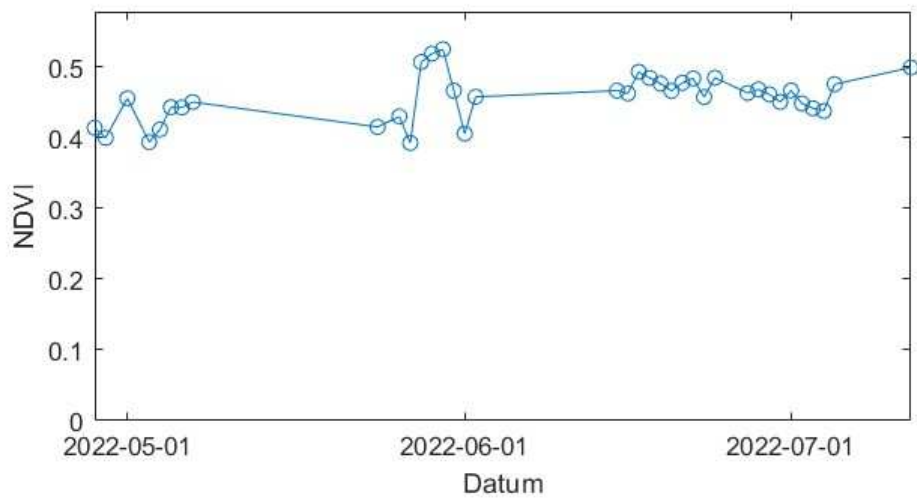


Sl. 2.5 Graf površine u pikselima promatran u blisko crvenom spektru

NDVI se računao tako da se jednačba (1) primjenila na fotografije u R i NIR spektru te se zatim na tim fotografijama vršila segmentacija. Na slici 2.6 vidljiv je rezultat primjene jednačbe (1) na fotografije R i NIR spektra. Jednaka problematika primjenjivanja istog praga za segmentiranje svih fotografija vidljiva je na slici 2.7.

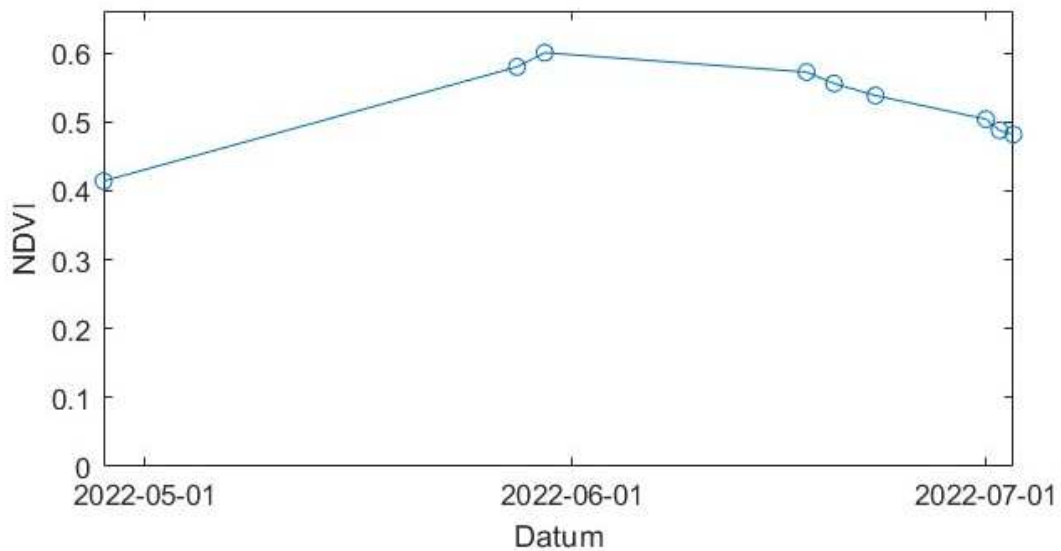


Sl. 2.6 R spektar - lijevo , NIR spektar – u sredini , NDVI - desno

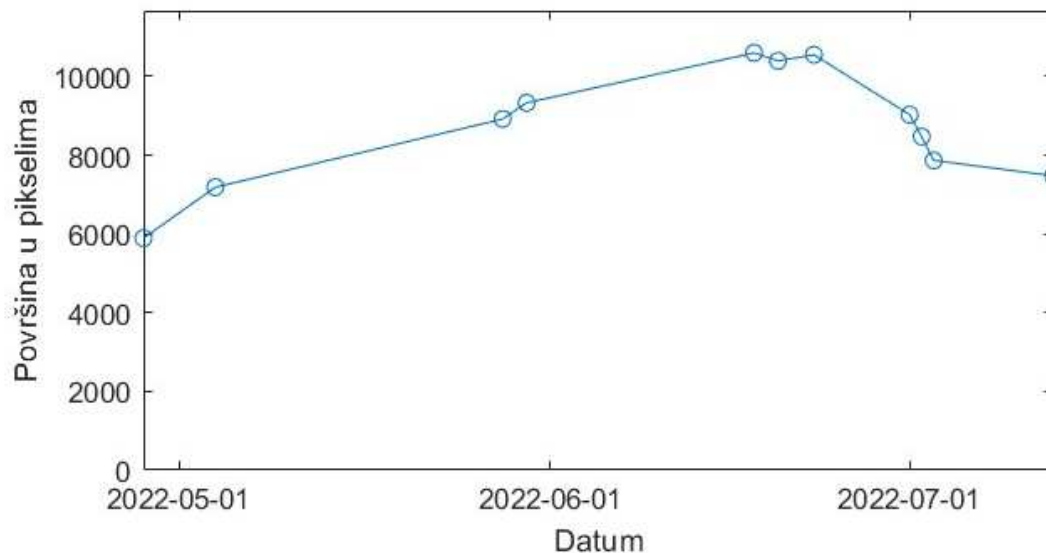


Sl. 2.7 Rezultat segmentiranja sa jednim pragom za sve fotografije promatrane u blisku crvenom spektru

Empirijski je utvrđeno da je dovoljno napraviti četiri skupa fotografija ovisno o kontrastu i osvjetljenju te generirati četiri funkcije za segmentiranje, također su eliminirane fotografije tj. vrijednosti NDVI-a i površine u pikselima koje odudaraju od očekivanih vrijednosti zbog utjecaja smetnji. Ovaj pristup se pokazao zadovoljavajućim za ovaj slučaj te slika 2.8 i slika 2.9 pokazuju rezultat primjene četiri funkcije za segmentiranje.



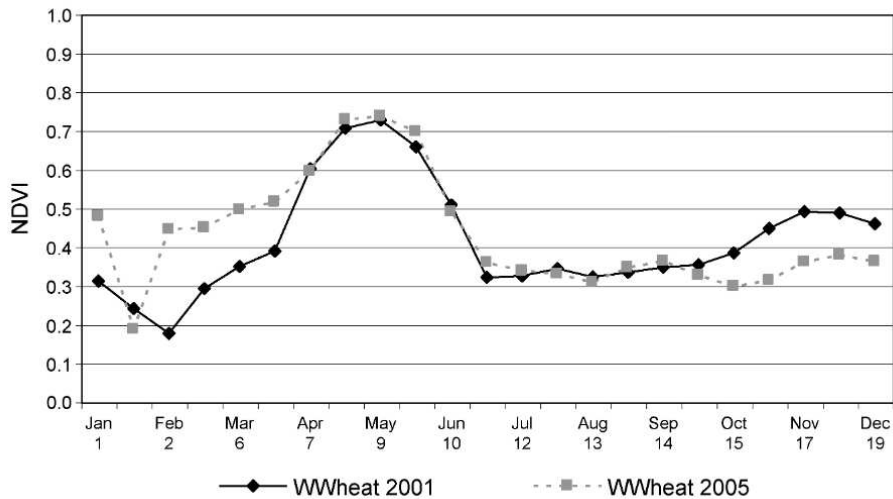
Sl. 2.8 Graf NDVI vrijednosti korištenjem četiri funkcije za segmentiranje



Sl. 2.9 Graf površine u pikselima korištenjem četiri funkcije za segmentiranje



Kada se usporede grafovi na slikama 2.6 i 2.9 sa drugim radovima na ovu temu vidljivo je da dobivene vrijednosti odgovaraju očekivanom kretanju ovih parametara kroz vrijeme u razdoblju od kraja travnja do polovice srpnja. Na slici 2.10 i 2.11 je prikazano očekivano kretanje NDVI i biomase kroz vrijeme.



Sl. 2.10 Graf kretanja NDVI indeksa kroz vrijeme, crna krivulja – ozima pšenica 2001. godine, siva krivulja – ozima pšenica 2005. godine [3]



Sl. 2.11 Graf biomase pšenice kroz vrijeme[4]

Vrijedi napomenuti da graf na slici 2.11 prikazuje ukupnu nadzemnu suhu tvar biljke, a za potrebe ovoga rada primjenjena je aproksimacija te biomase kao broj piksela koji ne reprezentira u potpunosti ukupnu biomasu (promjene u kutu slikanja, sušenje lišća, promjene u postotku vlage u biljci itd.).

### 2.3. Uparivanje obrađenih podataka o NDVI i površini pšenice s laboratorijskim mjerenjima

U ovom poglavlju opisan je postupak uparivanja NDVI vrijednosti i površine pšenice u pikselima korištenjem Jupyter Notebook-a. Ovaj proces uključuje nekoliko ključnih koraka:

- dohvat podataka
- interpolaciju podataka
- dodavanje varijabli za navodnjavanje i fenofazu biljke
- korelacijsku analizu.

CSV datoteka sa senzorskim mjerenjima sadrži mjerenja za svaku biljku. Podatci o svakoj biljci se mogu razdvojiti po zonama koje se kreću ovako:

- Komora 1: 1, 2, 3, 4
- Komora 2: 5, 6, 7, 8
- Komora 3: 9, 10, 11, 12
- Komora 4: 13, 14, 15, 16

Na slici 2.12 prikazana su učitana senzorska mjerenja.

	#	measurement_id	batch_timestamp	timestamp	air_temperature	air_humidity	lighting	soil_moisture	zone_id
0	1	1	2022-01-03 14:29:48.000000	2021-11-09 15:31:05.000000	14.0	89.3	99.9	6.2	1
1	2	2	2022-01-03 14:29:48.000000	2021-11-09 15:31:05.000000	13.3	95.4	99.9	6.3	2
2	3	3	2022-01-03 14:29:48.000000	2021-11-09 15:31:05.000000	13.4	84.2	99.9	5.3	3
3	4	4	2022-01-03 14:29:48.000000	2021-11-09 15:31:05.000000	13.1	77.7	99.9	4.1	4
4	5	5	2022-01-03 14:29:48.000000	2021-11-09 15:31:21.000000	14.0	68.5	99.9	6.1	1
...	...	...	...	...	...	...	...	...	...
582103	582104	582141	2023-07-05 01:00:21.000000	2023-07-04 12:42:31.000000	25.9	87.4	85.4	36.7	16
582104	582105	582142	2023-07-05 01:00:21.000000	2023-07-04 12:57:31.000000	25.7	97.9	0.2	75.7	13
582105	582106	582143	2023-07-05 01:00:21.000000	2023-07-04 12:57:31.000000	25.5	91.9	4.4	43.8	14
582106	582107	582144	2023-07-05 01:00:21.000000	2023-07-04 12:57:31.000000	25.8	95.8	0.9	-27.5	15
582107	582108	582145	2023-07-05 01:00:21.000000	2023-07-04 12:57:31.000000	25.8	87.5	85.5	36.7	16

582108 rows x 9 columns

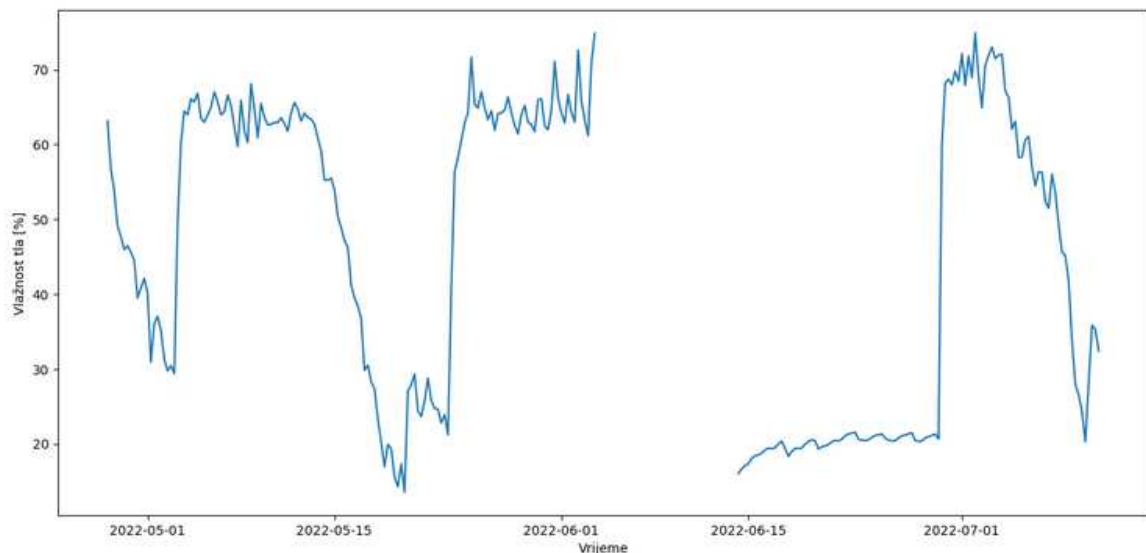
Sl. 2.12 Okvir podataka o senzorskim mjerenjima

Prvo je bilo ključno eliminirati senzorska mjerenja izvan vremenskog intervala za koji posjedujemo podatke o NDVI vrijednostima i površini biljaka. Također je iznimno važno eliminirati vrijednosti izvan raspona <10, 100> za vlagu tla i vlagu zraka, budući da takve vrijednosti ukazuju na pogreške u mjerenjima. Vrijednosti vlage tla ispod 10% su se događala kada se senzor fizički odvoji od tla te ne mjeri stvarnu vrijednost. Vremenski interval mjerenja je postavljan na šest sati. Na slici 2.13 prikazan je rezultirajući okvir podataka.

	timestamp	air_temperature	air_humidity	lighting	soil_moisture	zone_id	Date
0	2022-04-28 00:00:00	12.095833	70.958333	0.058333	63.145833	1.0	2022-04-28
1	2022-04-28 06:00:00	18.687500	47.633333	0.291667	56.750000	1.0	2022-04-28
2	2022-04-28 12:00:00	21.137500	37.641667	0.079167	53.825000	1.0	2022-04-28
3	2022-04-28 18:00:00	16.333333	47.241667	0.037500	49.125000	1.0	2022-04-28
4	2022-04-29 00:00:00	14.783333	48.787500	0.062500	47.695833	1.0	2022-04-29
...	...	...	...	...	...	...	...
303	2022-07-12 18:00:00	21.350000	65.225000	0.004167	NaN	1.0	2022-07-12
304	2022-07-13 00:00:00	19.420833	76.383333	0.016667	NaN	1.0	2022-07-13
305	2022-07-13 06:00:00	27.187500	40.254167	0.125000	NaN	1.0	2022-07-13
306	2022-07-13 12:00:00	30.895833	26.629167	0.041667	NaN	1.0	2022-07-13
307	2022-07-13 18:00:00	25.416667	48.433333	0.000000	NaN	1.0	2022-07-13

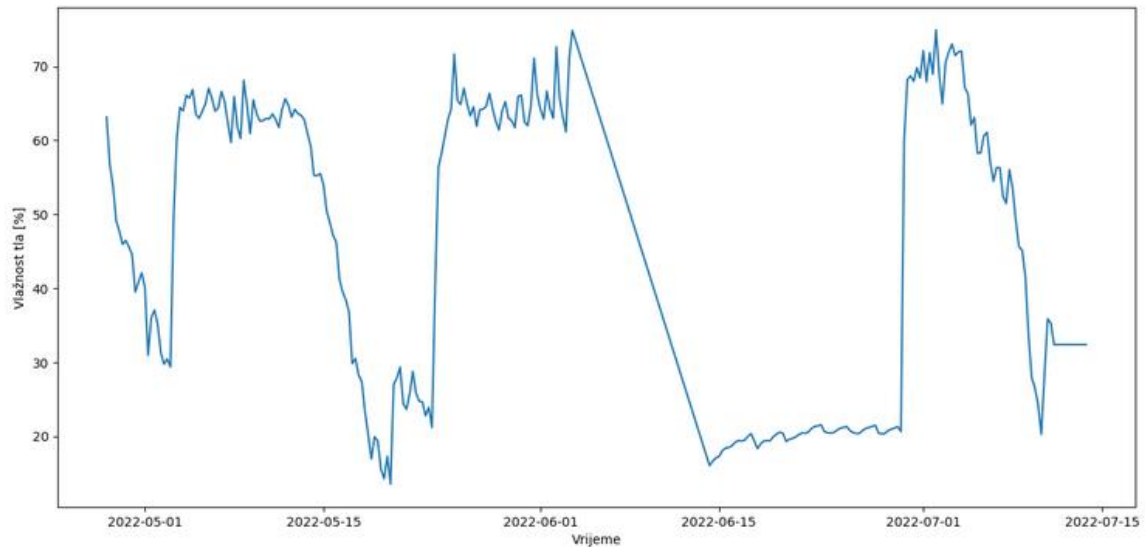
308 rows x 7 columns

Sl. 2.13 Okvir podataka sa odgovarajućim vremenskim intervalom i izbačenim pogrešnim mjerenjima



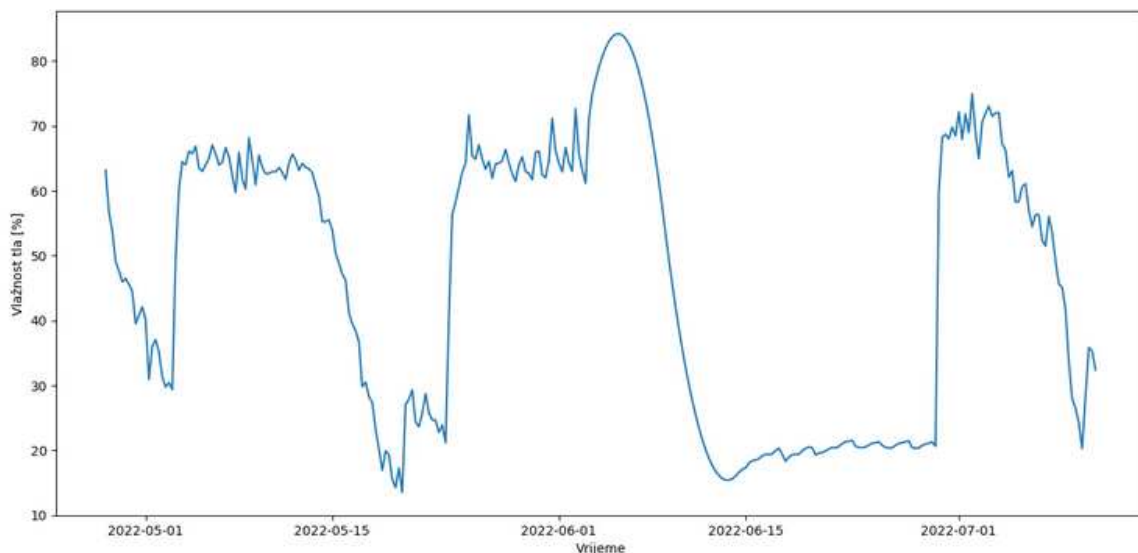
Sl. 2.14 Senzorska mjerenja vlažnosti tla za biljku 1 u komori 1 (zone\_id =1)

Iz slike 2.14 vidljivo je da postoje „rupe“ u mjerenjima čiji su uzrok previsoke/preniske vrijednosti ili pak nedostatak mjerenja u potpunosti. U svakom slučaju ove vrijednosti potrebno je interpolirati, te je za to najpraktičnija metoda linearne interpolacije ili čak ovisno o slučaju backward-fill (dopunjavanje rupa u mjerenjima sa zadnjom poznatom vrijednosti) ili forward-fill metoda (dopunjavanje rupa u mjerenjima sa slijedećom poznatom vrijednosti). Na slici 2.15 prikazan je interpolirani graf vrijednosti vlažnosti tla korištenjem metode linearne interpolacije.



Sl. 2.15 Linearna interpolacija vrijednosti vlažnosti tla

Polinomna interpolacija nije odabrana iz razloga što ukoliko nedostaje previše uzastopnih mjerenja rezultat je da funkcija po kojoj se interpolira unese preveliku grešku i nerealno prikazuje kretanje npr. vlažnosti tla. Opisana situacija je vidljiva na slici 2.16.



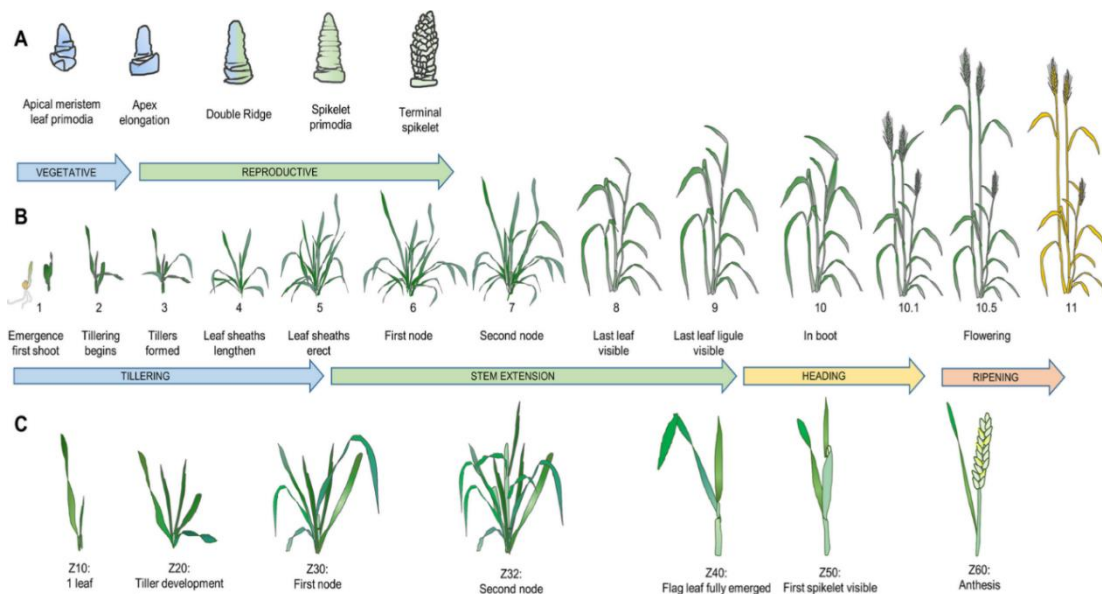
Sl. 2.16 Polinomna interpolacija drugog stupnja

Slijedeći korak bio je napraviti nove varijable koje bi bile od koristi za buduće terensko testiranje. Jedna od tih varijabli je navodnjavanje u čiju kategoriju spadaju i padaline uzrokovane kišom. Navodnjavanje kao varijabla dobivena je iz promjene vlažnosti tla jer drugi parametri o navodnjavanju nisu reprezentativni (upravljačke veličine vodene pumpe ne odgovaraju navodnjavanju zbog povremenog nedostatka vode u spremniku).

Za promjenu vlage koja će označavati minimalno navodnjavanje odabrana je promjena od 10%. Ova promjena se zatim pomnožila sa ukupnim volumenom zemlje te se tako dobila gruba procjena navodnjavanja u mililitrima. Za potrebe testiranja na terenu potrebno je ove mililitre pretvoriti u milimetre po metru kvadratnom.

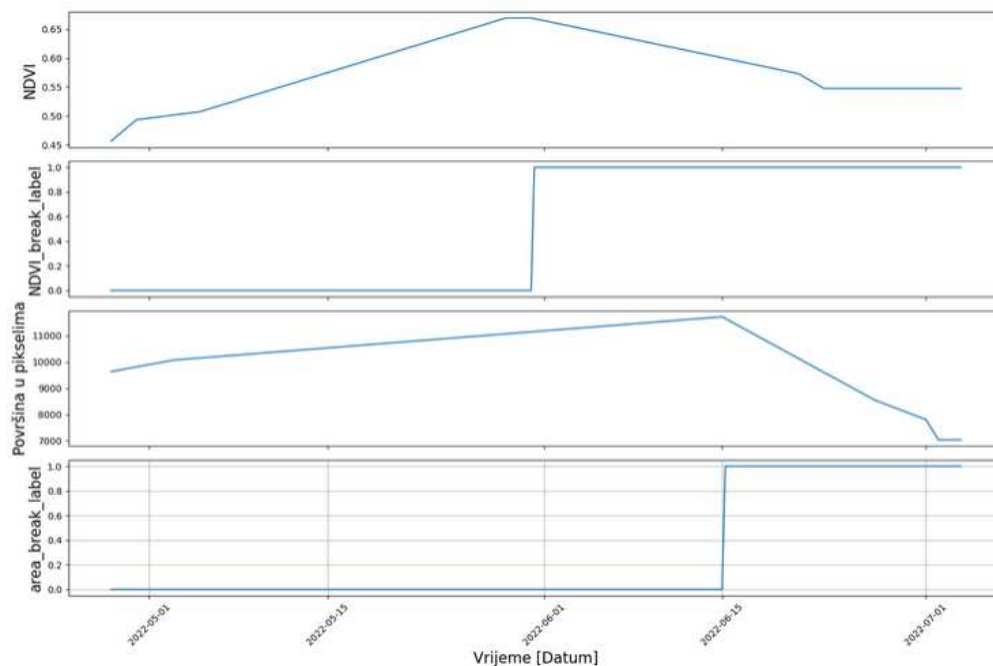
Kao dva dodatna parametra odabrana je fenofaza biljke te kumulativno navodnjavanje po fenofazi.

Slika 2.17 prikazuje korištenu reprezentaciju fenofaza pšenice čije se vrijednosti u ovom slučaju kreću od 6 do 11.



Sl. 2.17 Fenofaze pšenice, korištena reprezentacija pod B[5]

Kako bi model što bolje predviđao ciljane parametre dodane su nove varijable koje označavaju je li došlo do prijeloma funkcije. Na slici 2.18 prikazani su spomenuti parametri čija je vrijednost 0 ili 1 u slučaju ako se dogodio prijelom te zrno pšenice počinje sazrijevati.



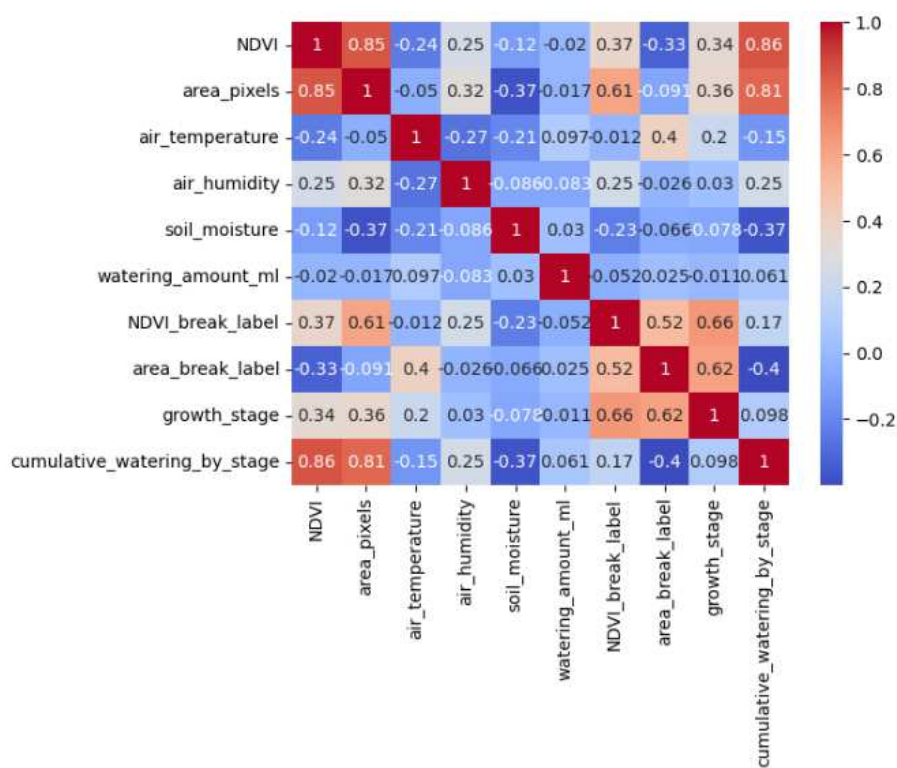
Sl. 2.18 Prikaz varijabli area\_break\_label i NDVI\_break\_label koje daju informaciju je li došlo do loma grafa površine biljke i NDVI vrijednosti

## 2.4. Analiza uparenih senzorskih podataka i vrijednosti NDVI-a i površine pšenice u pikselima

Kako bi se odredilo koje parametre je najbolje koristiti u predikciji budućih vrijednosti NDVI i površine pšenice (biomase) bitno je odrediti koji parametri zadovoljavajuće koreliraju sa istim.

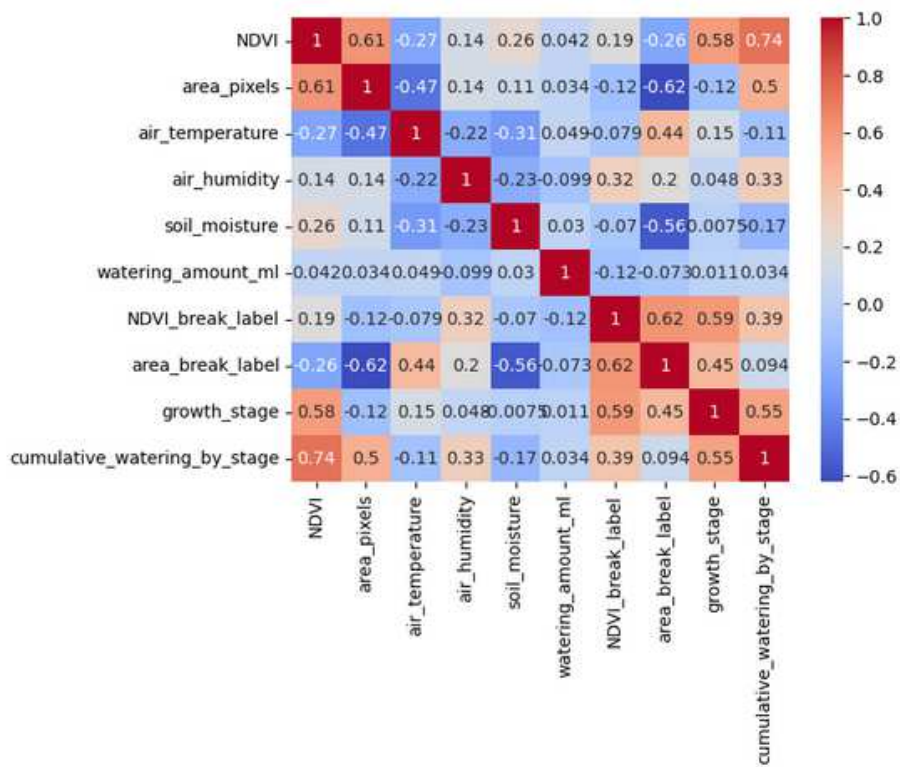
Za reprezentaciju koreliranja je korištena Pearsonova korelacija koja kvantificira stupanj linearnog odnosa između dviju varijabli.

Na slikama 2.19, 2.20, 2.21 i 2.22 prikazana je korelacijska matrica za 4 obrađene biljke.

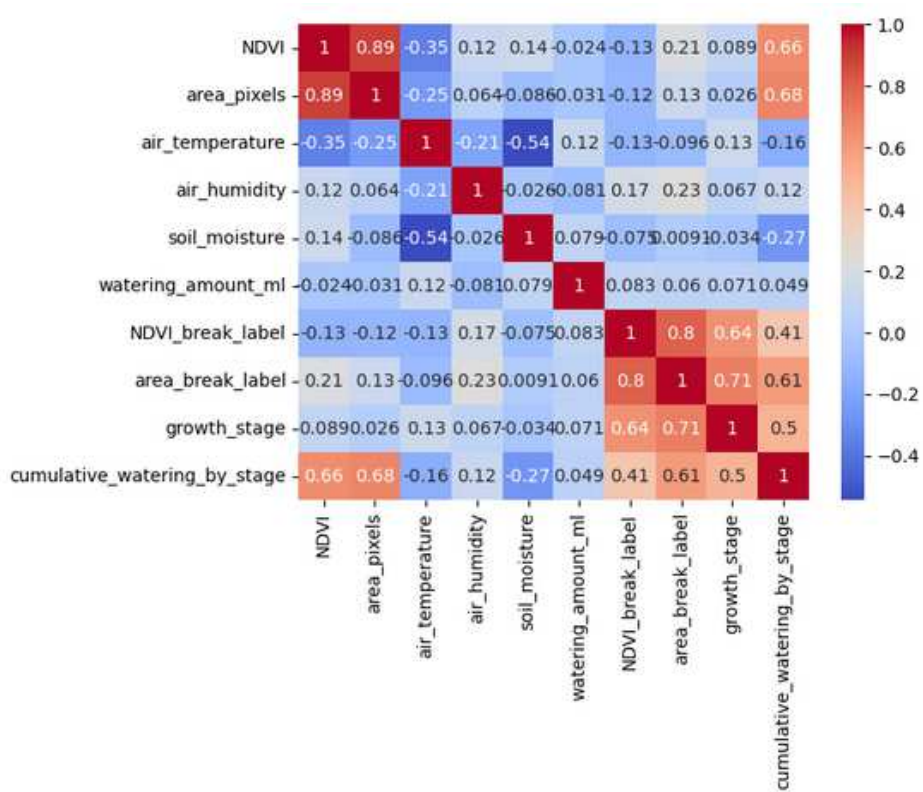


Sl. 2.19 Korelacijska matrica biljke sa identifikacijskim brojem 1



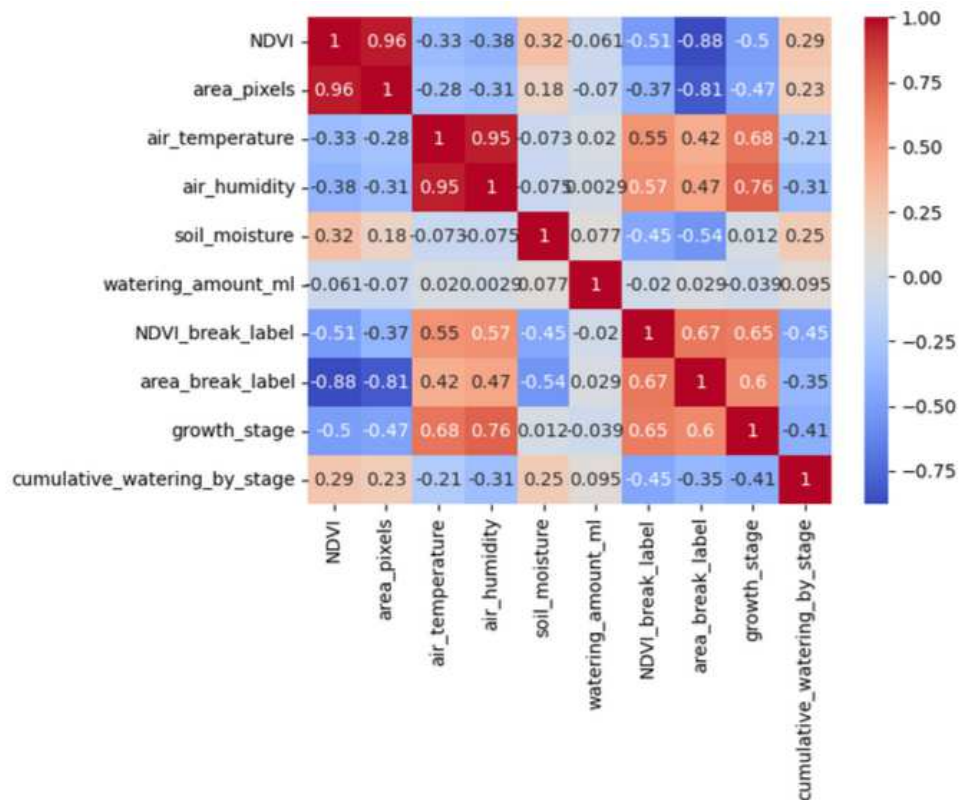


Sl. 2.20 Korelacijska matrica biljke sa indentifikacijskim brojem 2



Sl. 2.21 Korelacijska matrica s indentifikacijskim brojem 3





Sl. 2.22 Korelacijska matrica biljke s identifikacijskim brojem 14

Kao što je vidljivo iz korelacijskih matrica s NDVI i površinom u pikselima najbolje korelira vlažnost tla, broj fenofaze biljke te kumulativno navodnjavanje po fenofazi u mililitrima te su ti parametri odabrani za treniranje neuronske mreže i predikciju. Uz navedene parametre u model su dodani još i količina navodnjavanja u mililitrima te parametri koji označava je li došlo do prijeloma funkcije NDVI i površine u pikselima (NDVI\_break\_label i area\_break\_label).

### **3. Predviđanje NDVI-a i površine pšenice korištenjem modela strojnog učenja**

Kod problematike predviđanja parametara bitno je krenuti s jednostavnijim modelima, posebno ako je podatkovni skup relativno malen. U ovom poglavlju biti će obrađeni modeli linearne regresije i jednostavna neuronska mreža kako bi se odredilo koji model najbolje predviđa buduće vrijednosti NDVI-a i površine pšenice.

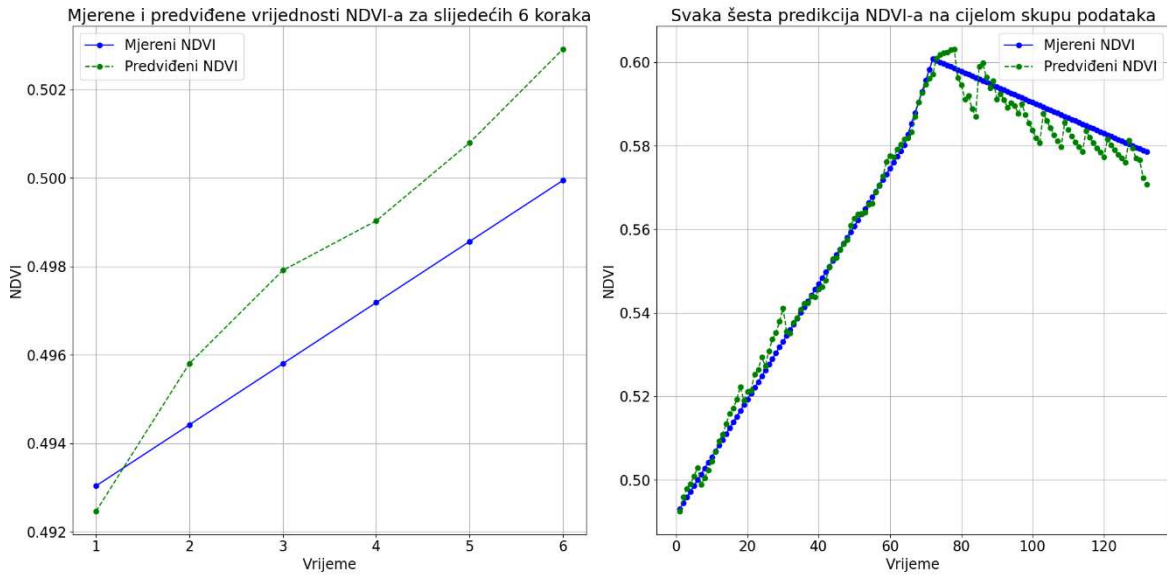
#### **3.1. Predprocesiranje podataka**

Kako bi se adekvatno pripremili podatci za modele strojnog i dubokog učenja bitno je pripremiti skup podataka. Za svaki od korištena tri modela format podataka je isti. Zbog jednostavnosti određen je prozor od šest prošlih vrijednosti korištenih značajki i šest budućih predviđanih vrijednosti.

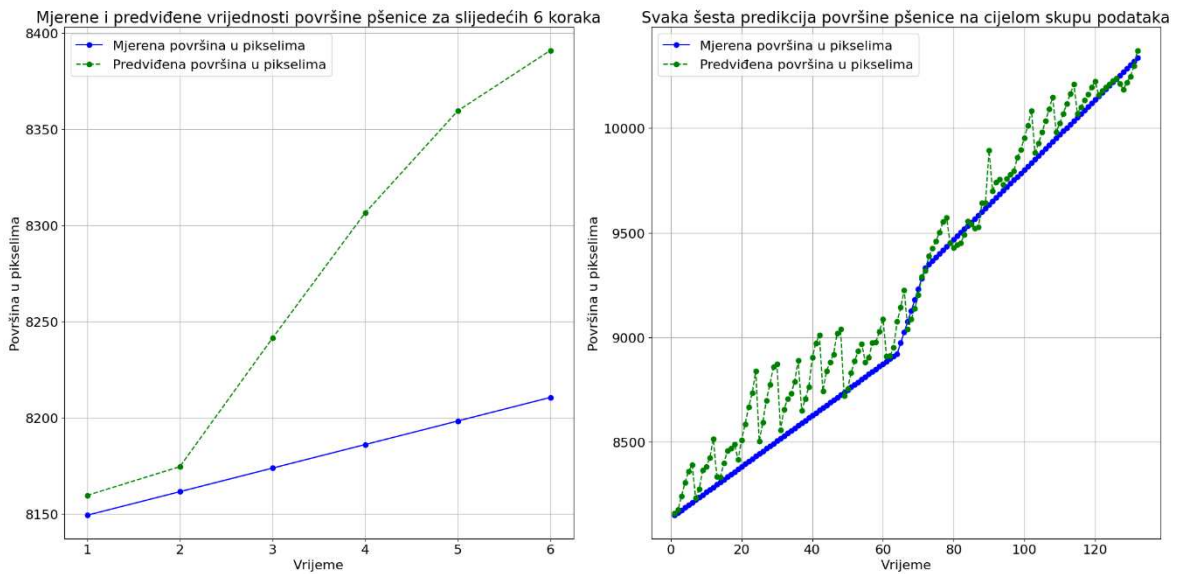
Standardizacije podataka obavljena je korištenjem StandardScaler modula sklearn biblioteke.

#### **3.2. Predviđanje NDVI-a i površine pšenice korištenjem modela linearne regresije**

Model linearne regresije minimizira rezidualni zbroj kvadrata između promatranih ciljeva u skupu podataka i ciljeva predviđenih linearnom aproksimacijom[6]. Rezultati predikcije NDVI-a i površine u pikselima za slijedećih šest koraka prikazani su na slici 3.1 i 3.2



Sl. 3.1 Graf predikcije NDVI-a za slijedećih šest koraka korištenjem modela linearne regresije

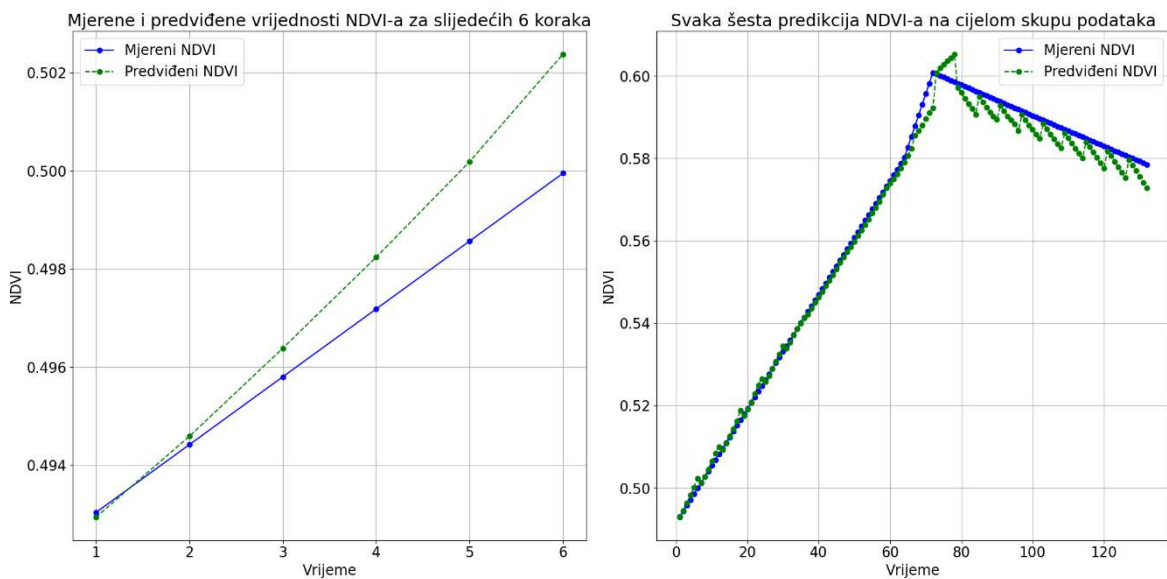


Sl. 3.2 Grafovi predikcija površine u pikselima za slijedećih šest koraka korištenjem modela linearne regresije

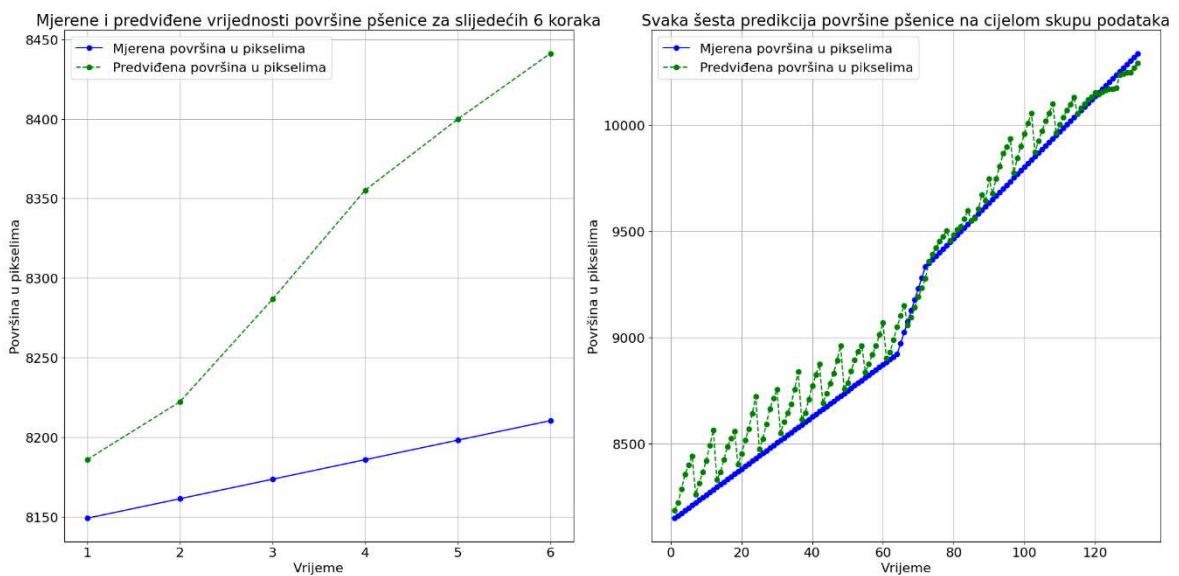
### 3.3. Predviđanje NDVI-a i površine pšenice korištenjem Lasso modela

Lasso model je također model linearne regresije no uz to je dodana regularizacija koja sprječava prenaučenosť što je vrlo bitno kada se radi s malim skupovima podataka[7].

Finim podešavanjem faktora regularizacije utvrđeno je da je optimalna vrijednost ovog hiperparametra  $\alpha=0.01$ . Grafovi koji prikazuju rezultate predikcije prikazani su na slikama 3.3 i 3.4.



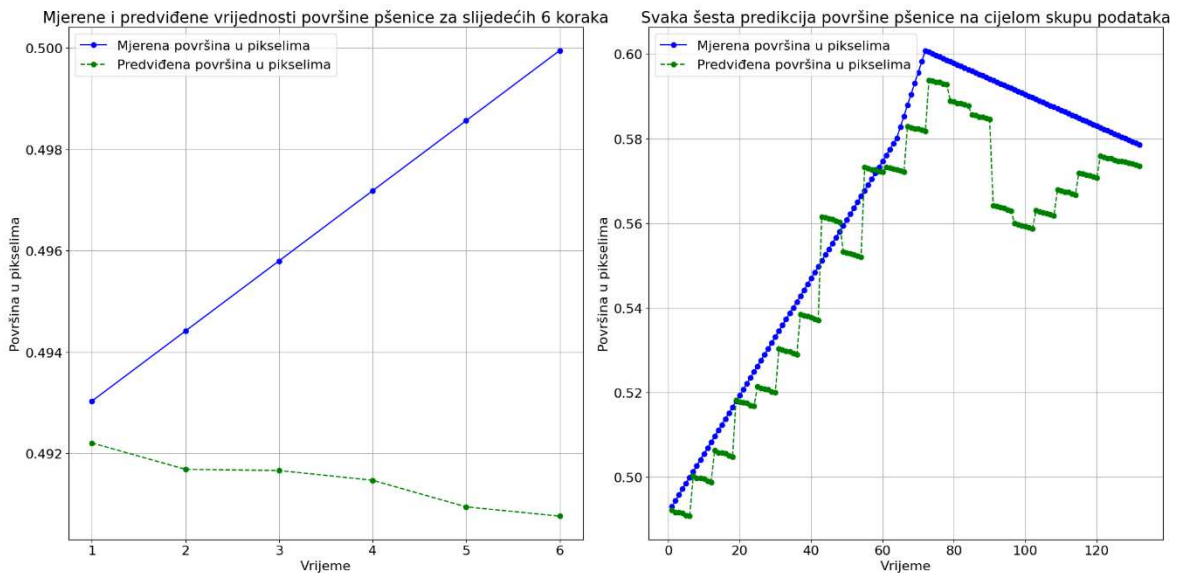
Sl. 3.3 Graf predikcija NDVI-a korištenjem Lasso modela



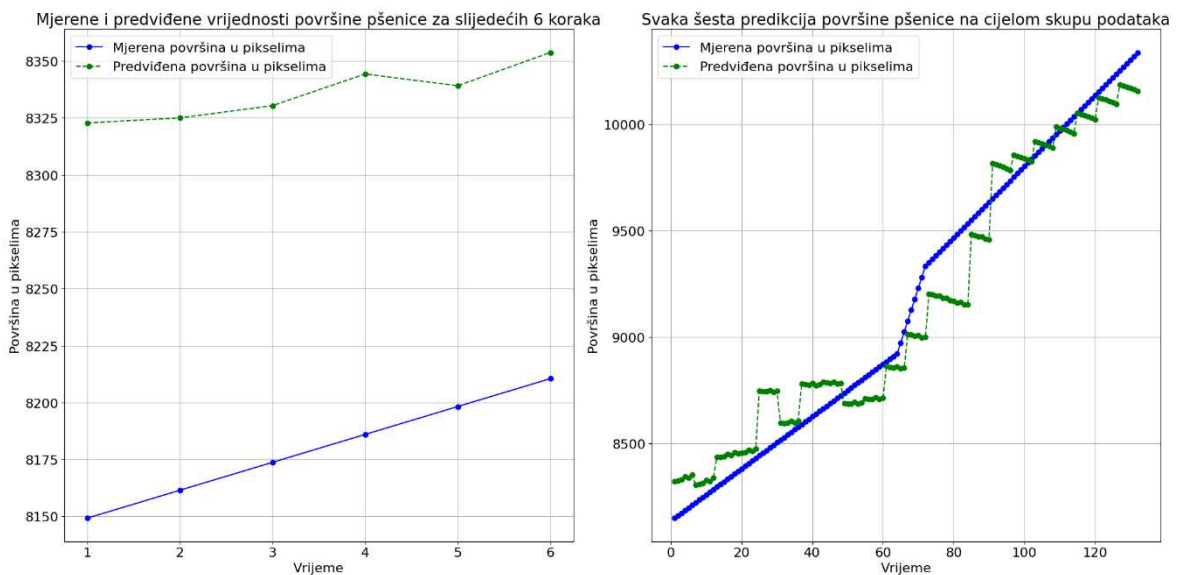
Sl. 3.4 Grafovi predikcije površine u pikselima korištenjem Lasso modela

### 3.4. Predviđanje Multi Layer Perceptron (MLP) modelom

Prirodno je nakon jednostavnijih modela poput linearne regrecije odabrati neki malo složeniji model kako bi vidjeli ponašanje na ovom skupu podataka. U ovu svrhu odabrana je potpuno povezana neuronska mreža (eng. Multi Layer Perceptron, MLP)[8]. Rezultati predikcije MLP-om prikazani su na slikama 3.5 i 3.6.



Sl. 3.5 Graf stvarnih i predviđenih vrijednosti NDVI-a



Sl. 3.6 Graf stvarnih i predviđenih vrijednosti površine u pikselima

## 4. Predviđanje površine i NDVI-a pšenice korištenjem povratne neuronske mreže

U ovom odeljku biti će opisana metodologija korištenja povratne neuronske mreže (eng. Recurrent Neural Networks , RNN) za predviđanje NDVI-a i površine pšenice na temelju vremenskih serija podataka. Kroz proces će biti pokriveni koraci od predprocesiranja podataka, preko arhitekture i treniranja neuronske mreže, do evaluacije rezultata.

### 4.1. Predprocesiranje podataka

Predprocesiranje podataka je ključan korak za osiguravanje da ulazni podaci budu u odgovarajućem formatu i kvaliteti za treniranje modela povratne neuronske mreže. Ovaj odjeljak sastoji se od prozorske segmentacije podataka te standardizacije podataka.

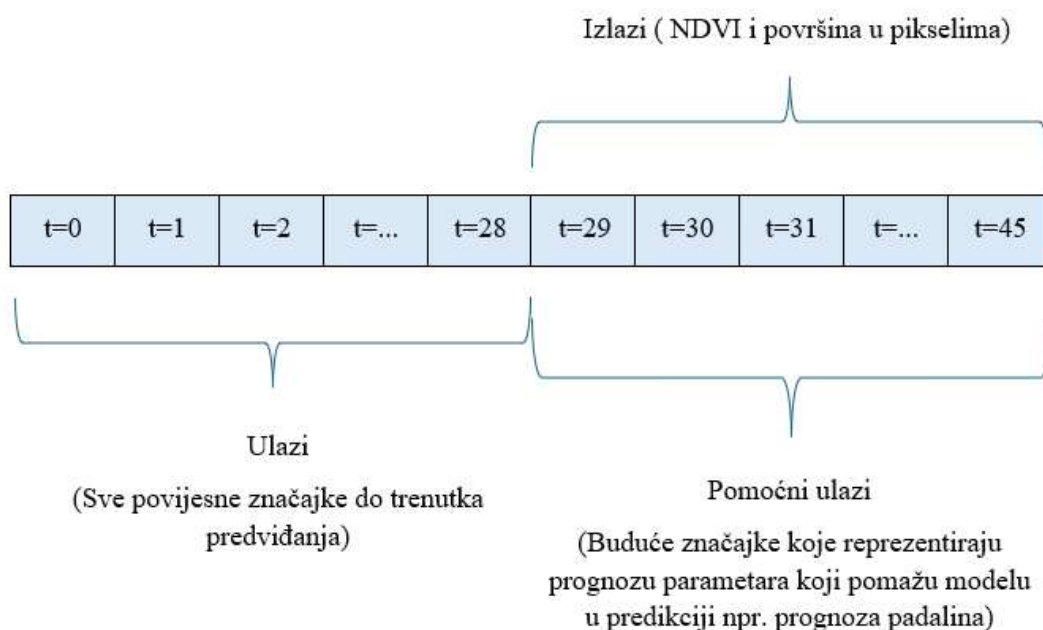
Kao prvi korak prije prozorske segmentacije podataka bilo je potrebno učitati CSV datoteke koje su dobivene obradom u Jupyter Notebook-u te ih adekvatno spojiti u jedan okvir podataka koji je prikazan na slici 3.1.

	NDVI	area_pixels	air_temperature	air_humidity	soil_moisture	watering_amount_ml	NDVI_break_label	area_break_label	growth_stage	cumulative_watering_by_stage
0	0.414215	5890.000000	12.095833	70.958333	63.145833	0.0	0	0	8	0.000000
1	0.415598	5943.958333	18.687500	47.633333	56.750000	0.0	0	0	8	0.000000
2	0.416981	5997.916667	21.137500	37.641667	53.825000	0.0	0	0	8	0.000000
3	0.418363	6051.875000	16.333333	47.241667	49.125000	0.0	0	0	8	0.000000
4	0.419746	6105.833333	14.783333	48.787500	47.695833	0.0	0	0	8	0.000000
...	...	...	...	...	...	...	...	...	...	...
1187	0.372818	8664.000000	22.576388	22.579745	55.141667	0.0	1	1	11	2.605443
1188	0.372818	8664.000000	22.004721	22.465967	55.729167	0.0	1	1	11	2.605443
1189	0.372818	8664.000000	22.639721	22.446022	60.595833	0.0	1	1	11	2.605443
1190	0.372818	8664.000000	23.847638	22.613411	69.866667	0.0	1	1	11	2.605443
1191	0.372818	8664.000000	24.190786	22.807078	61.869565	0.0	1	1	11	2.605443

Sl. 4.1 Prikaz korištenog okvira podataka

### 4.1.1. Prozorska segmentacija podataka

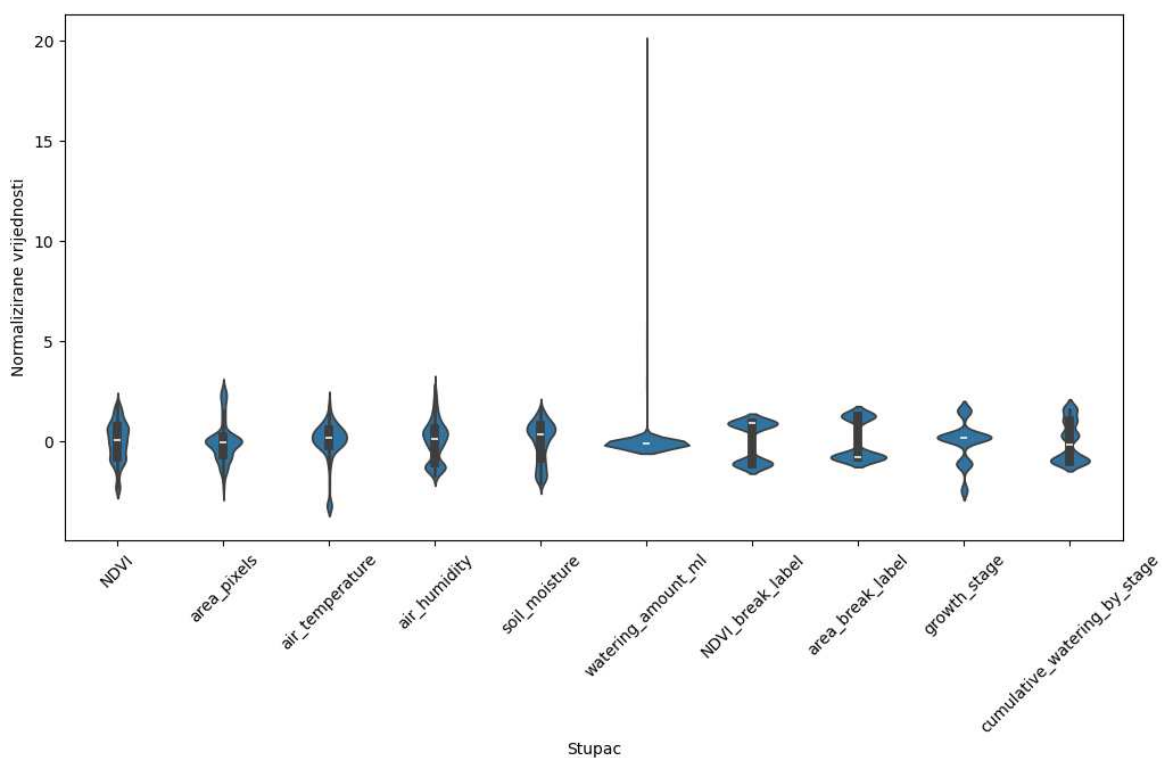
Prozorska segmentacija podataka uključuje segmentiranje vremenskih serija na manje podsekvence koje mreža može koristiti kao ulazne podatke. Ovo omogućuje mreži da bolje uči vremenske obrasce i odnose u podacima. U ovom koraku određujemo veličinu prozora, odnosno broj vremenskih točaka koje će se koristiti kao ulaz za mrežu, i generiramo odgovarajuće prozore iz vremenskih serija NDVI indeksa i površine pšenice. Na slici 3.2 prikazana je vizualna reprezentacija prozorska segmentacija ulaznih podataka za mrežu gdje se pomoću 28 zadnjih vremenskih trenutaka (7 dana) predviđa sekvenca od 28 budućih vrijednosti uz dodatnu pomoć budućih vrijednosti nekih značajki (temperatura zraka, vlažnost zraka i padaline/navodnjavanje) koje odgovaraju prognozi nekih parametara te je u ovom slučaju odabran deterministički pristup (zbog dostupnosti podataka) toj prognozi što bi značilo da je ta prognoza u potpunosti točna (stvarni budući podatci).



Sl. 4.2 Vizualna reprezentacija prozorske segmentacije podataka[9]

## 4.1.2. Standardizacija podataka

Standardizacija podataka podrazumijeva skaliranje podataka kako bi se postigla konzistentnost i ubrzala konvergencija mreže. Ovaj korak je bitan jer različite mjerne jedinice i rasponi vrijednosti mogu utjecati na performanse modela. Za standardizaciju podataka korištena je sklearn biblioteka, odnosno njezin StandardScaler modul, kojim smo standardizirali podatke za NDVI i površinu pšenice tako da smo oduzeli srednju vrijednost i podijelili sa standardnom devijacijom za svaku seriju podataka. Reprerzentacija normaliziranih podataka prikazana je na slici 3.3.



Sl. 4.3 Prikaz normaliziranih podataka za treniranje neuronske mreže



## 4.2. Povratna neuronska mreža za predviđanje NDVI-a i površine biljke

U ovom odjeljku opisana je struktura i funkcionalnost povratne neuronske mreže koja će se koristiti za predviđanje sekvence NDVI-a i površine pšenice. Poseban naglasak stavlja se na korištenje LSTM (Long Short-Term Memory) ćelija, koje su ključne za učinkovito hvatanje dugoročnih ovisnosti u podacima i neizostavne za predviđanje vremenskih sekvenci.

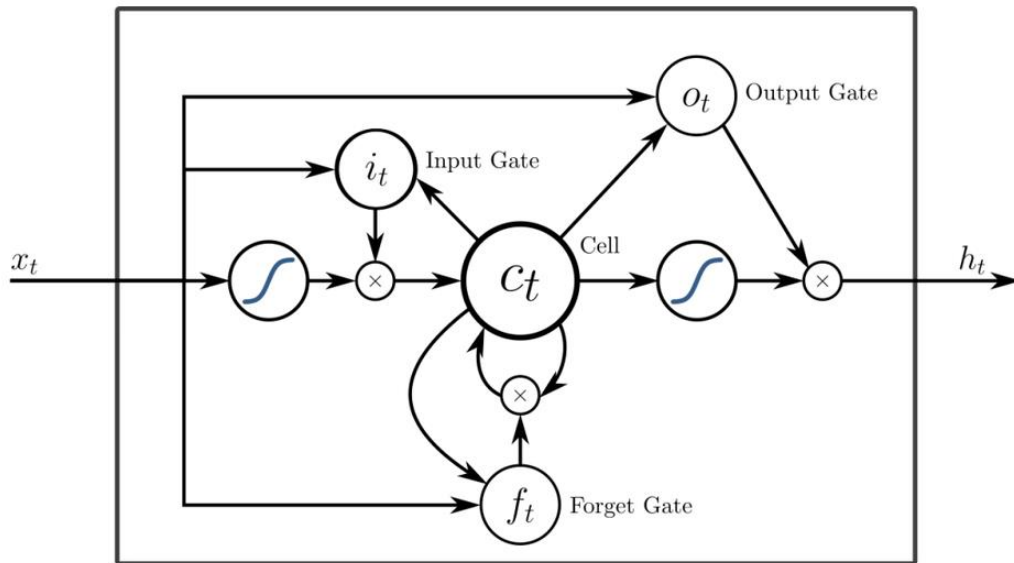
### 4.2.1. LSTM (Long Short-Term Memory)

Stanice dugog kratkoročnog pamćenja (LSTM) su vrsta arhitekture RNN-a koja rješava problem dugih nizova podataka u RNN-ovima. Tradicionalni RNN-ovi pate od problema nestajanja ili eksploziranja gradijenata, što ograničava njihovu sposobnost hvatanja dugoročnih ovisnosti. LSTM ćelije uvode memorijsku ćeliju koja omogućuje selektivno pohranjivanje i pristup informacijama kroz dulja vremenska razdoblja.

Memorijska ćelija LSTM-a sastoji se od tri glavne komponente: ulaznih vrata, zaboravljenih vrata i izlaznih vrata. Ulazna vrata određuju koliko novog ulaza treba pohraniti, zaboravna vrata kontroliraju koliko prethodnog stanja treba zaboraviti, a izlazna vrata odlučuju koliko stanja memorijske ćelije poslati u sljedeće skriveno stanje. Ova vrata koriste sigmoidnu aktivacijsku funkciju za regulaciju protoka informacija.

Korištenjem ovih komponenti, LSTM stanice učinkovito rješavaju problem nestajanja ili eksploziranja gradijenata, omogućujući RNN-ovima hvatanje dugoročnih ovisnosti u podacima. To je ključno za aplikacije poput obrade prirodnog jezika, prepoznavanja govora i predviđanja vremenskih serija [10].

Ilustracija LSTM ćelije prikazana je na slici 3.4.



Sl. 4.4 Prikaz LSTM ćelije[11]

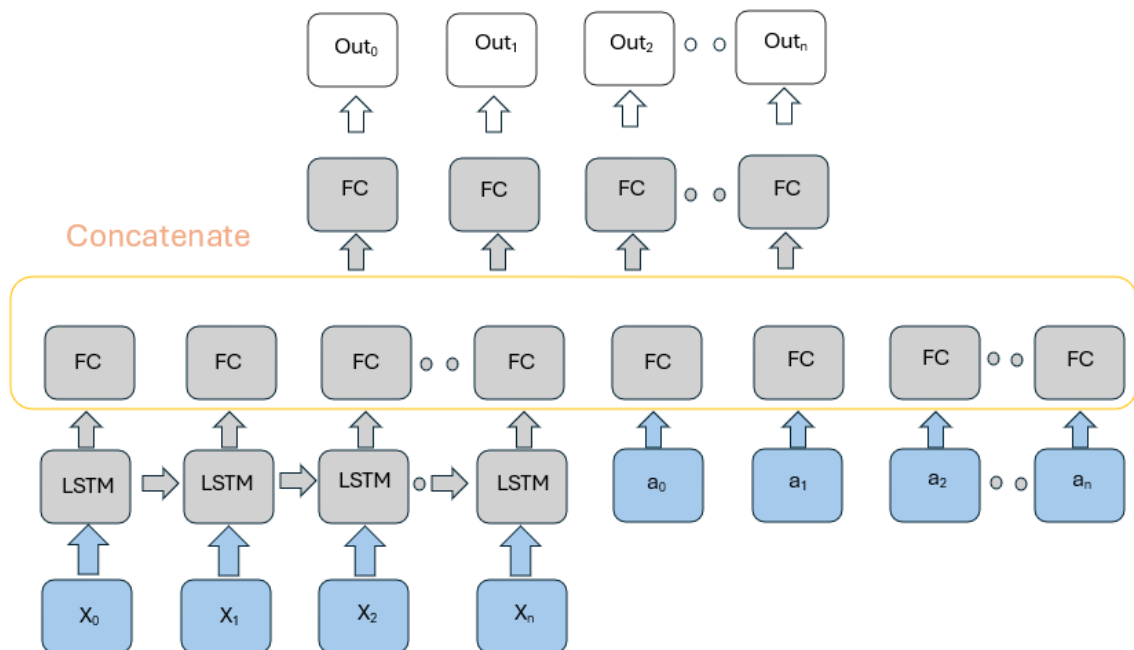
#### 4.2.2. Arhitektura i korišteni slojevi u neuronskoj mreži

U ovom odjeljku opisani su slojevi korišteni u povratnoj neuronskoj mreži za predviđanje NDVI-a i površine pšenice:

- **Ulazni sloj za prošle podatke (past\_input):** Prihvaća vremenske serije podataka o NDVI-u, površini pšenice, vlažnosti tla, etiketama prekida NDVI-a i površine, fazama rasta, kumulativnom navodnjavanju po fazi te količini navodnjavanja.
- **LSTM sloj:** Ovaj sloj s 256 jedinica omogućuje mreži učenje vremenskih obrazaca u podacima. Koristi L2 regularizaciju kako bi se spriječilo prekomjerno učenje i povratna veza za obradu sekvencijalnih podataka.
- **Dense sloj:** S 128 neurona koji ima veliku fleksibilnost i može naučiti različite oblike podataka jer je potpuno povezan s prethodnim slojem.
- **Ulazni sloj za buduće podatke (future\_input):** Prihvaća vremenske serije podataka o budućim fazama rasta, količini navodnjavanja te etiketama prekida NDVI-a i površine.

- **TimeDistributed Dense sloj:** Ovaj sloj s 64 neurona primjenjuje isti Dense sloj na svaki vremenski korak neovisno, omogućujući mreži učenje značajki iz budućih vremenskih koraka te koristi L2 regularizaciju.
- **Concatenate sloj:** Spaja izlaze iz prošlih i budućih ulaznih slojeva, kombinirajući informacije kako bi se postigla bolja predikcija.
- **Dense sloj:** S 2 neurona što odgovara formatu izlaznih podataka.

Vizualna reprezentacija ove arhitekture je prikazana na slici 3.4.



Sl. 4.5 Vizualna reprezentacija arhitekture neuronske mreže koja koristi LSTM sloj za predikciju sekvenci uz pomoć dodatnih ulaza[12]

L2 regularizacija dodana je u model neuronske mreže iz nekoliko razloga:

1. **Sprječavanje prenaučnosti (overfitting):** L2 regularizacija pomaže u sprječavanju prenaučnosti tako što penalizira velike težine u mreži. Kada su težine prevelike, model postaje previše prilagođen trenutačnim podacima i gubi sposobnost generalizacije na nove, neviđene podatke.

2. **Ujednačavanje težina:** L2 regularizacija nastoji održati težine malima i ujednačenima. Na taj način pomaže u stabilizaciji učenja i čini model otpornijim na promjene u podacima.
3. **Smanjenje varijance:** Regularizacija smanjuje varijancu modela, što znači da model postaje manje osjetljiv na fluktuacije u podacima i bolje generalizira na nove primjere.
4. **Poboljšanje robusnosti modela:** Uključivanjem L2 regularizacije, model postaje robusniji na šum u podacima, jer težine koje previše rastu i previše osciliraju mogu dovesti do osjetljivosti na male promjene u ulazima.

Na slici 3.5 prikazana je implementacija ove povratne neuronske mreže u Google Colab-u.

```
past_timesteps = 28
future_timesteps = 28
num_features_past = 8
num_features_future = 3
num_targets = 1

# Input past
past_input = Input(shape=(past_timesteps, num_features_past), name="past_input")
x = LSTM(128, return_sequences=True, bias_regularizer=regularizers.l2(0.1))(past_input)
x = Dense(64)(x)

# Input future
future_input = Input(shape=(future_timesteps, num_features_future), name="future_input")
y = TimeDistributed(Dense(64, kernel_regularizer=regularizers.l2(0.01), bias_regularizer=regularizers.l2(0.1)))(future_input)

#Concat
concatenated = Concatenate(axis=-1)([x, y])
#Output modela
output = Dense(num_targets, bias_regularizer=regularizers.l2(0.01))(concatenated)

model = Model(inputs=[past_input, future_input], outputs=output)
es = EarlyStopping(monitor='root_mean_squared_error', patience=15, verbose=1, restore_best_weights=True)

model.compile(loss=MeanSquaredError(), optimizer=Adam(learning_rate=0.01), metrics=[RootMeanSquaredError()])
model.summary()
```

Sl. 4.6 Implementacija korištene neuronske mreže

### 4.2.3. Hiperparametri korišteni u modelu za predikciju NDVI-a i površine pšenice

Hiperparametri su posebne vrste parametara u strojnom učenju koji se ne uče iz podataka, već ih postavlja korisnik prije procesa treniranja modela. Oni kontroliraju ponašanje algoritma i proces učenja. U osnovi, hiperparametri su "parametri za parametre" modela. Nekoliko ključnih hiperparametara u strojnom učenju uključuju:

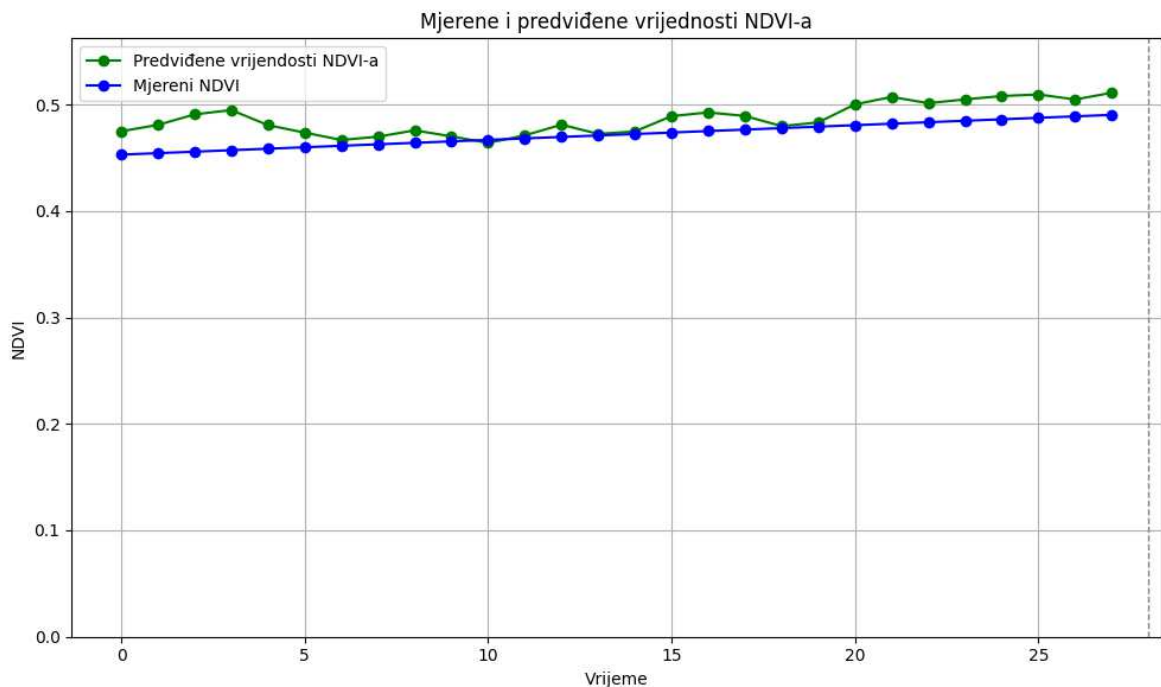
- **Brzina učenja (eng. Learning Rate):** Ovaj hiperparametar određuje korak kojim algoritam učenja ažurira težine modela. Ako je brzina učenja prevelika, model može preskočiti optimalno rješenje. Ako je premala, učenje može biti vrlo sporo i možda neće konvergirati.
- **Veličina serije (eng. Batch Size):** Ovo je broj uzoraka podataka koji se koriste za ažuriranje težina modela u jednom prolazu kroz mrežu. Manje serije mogu dovesti do bržeg učenja, ali s više šuma, dok veće serije mogu rezultirati stabilnijim procjenama gradijenata, ali sporijim treniranjem. U ovom slučaju odabran je batch size 16 jer je on rezultirao najboljom predikcijom.
- **Broj epoha (eng. Number of Epochs):** Jedna epoha je jedan prolaz kroz cijeli skup podataka za treniranje. Više epoha može poboljšati performanse modela, ali može dovesti i do pretreniranja (overfitting) no kako je korištena callback funkcija EarlyStopping treniranje je zaustavljeno na epohi za koju je bila najmanja pogreška modela.
- **Regularizacija (eng. Regularization):** Tehnike kao što su L1, L2 regularizacija koriste se za sprječavanje pretreniranja tako što penaliziraju velike težine ili isključuju određene neurone tijekom treniranja.
- **Struktura modela (eng. Model Architecture):** Hiperparametri ovdje uključuju broj slojeva, broj neurona po sloju, vrste aktivacijskih funkcija itd.

Svi spomenuti hiperparametri i njihove vrijednosti prikazani su na slici 4.6.

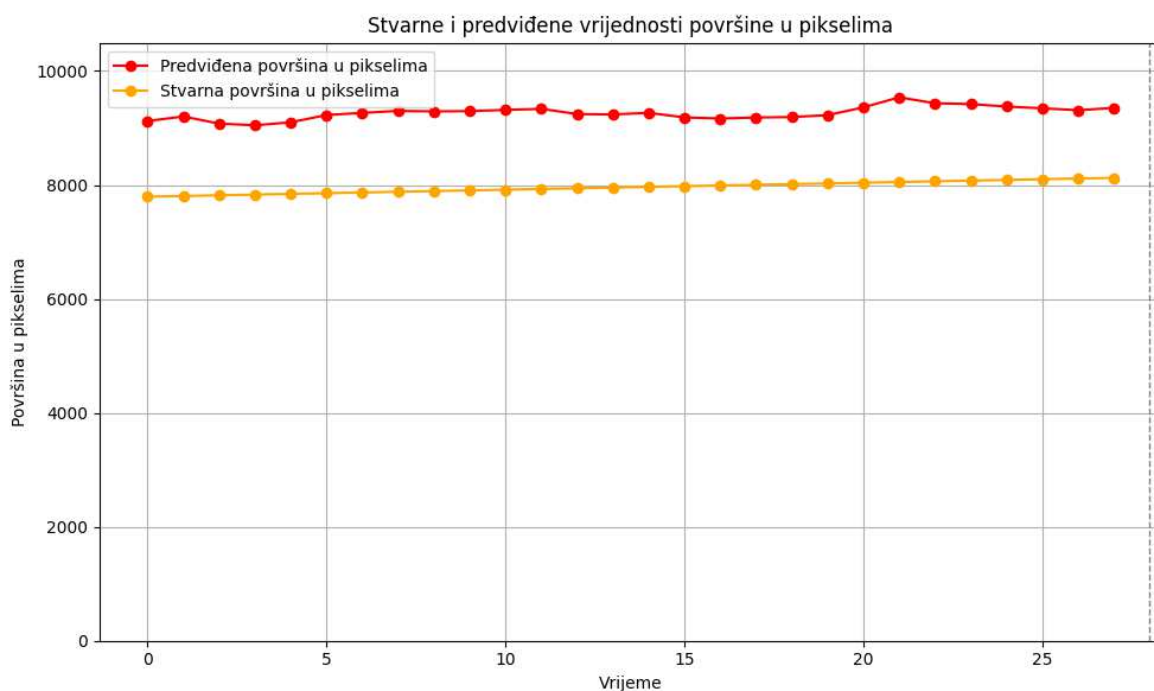
### 4.3. Rezultati predviđanja NDVI-a i površine pšenice korištenjem povratne neuronske mreže

Na slikama 4.7 i 4.8 prokazana je usporedba stvarnih i predviđenih vrijednosti NDVI-a i površine u pikselima za horizont od 28 vremenskih koraka koristeći 28 prethodnih vremenskih koraka.

Kao što je vidljivo iz grafova model puno preciznije predviđa vrijednosti NDVI-a nego površine. Ovo ponašanje modela može se pripisati malom skupu podataka, gdje je vrlo teško podesiti hiperparametre neuronske mreže kako bi se dobio optimalan rezultat za oba predviđena parametra.



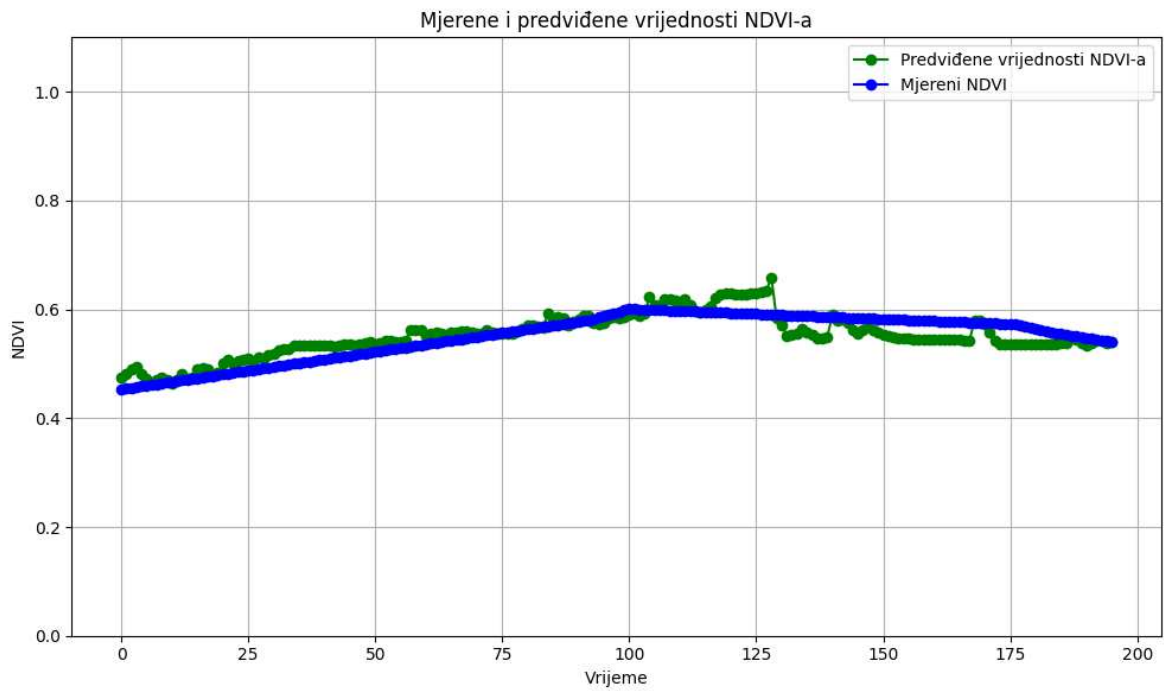
Sl. 4.7 Graf predviđenih i mjerenih vrijednosti NDVI-a za idućih sedam dana s vremenskim korakom od šest sati



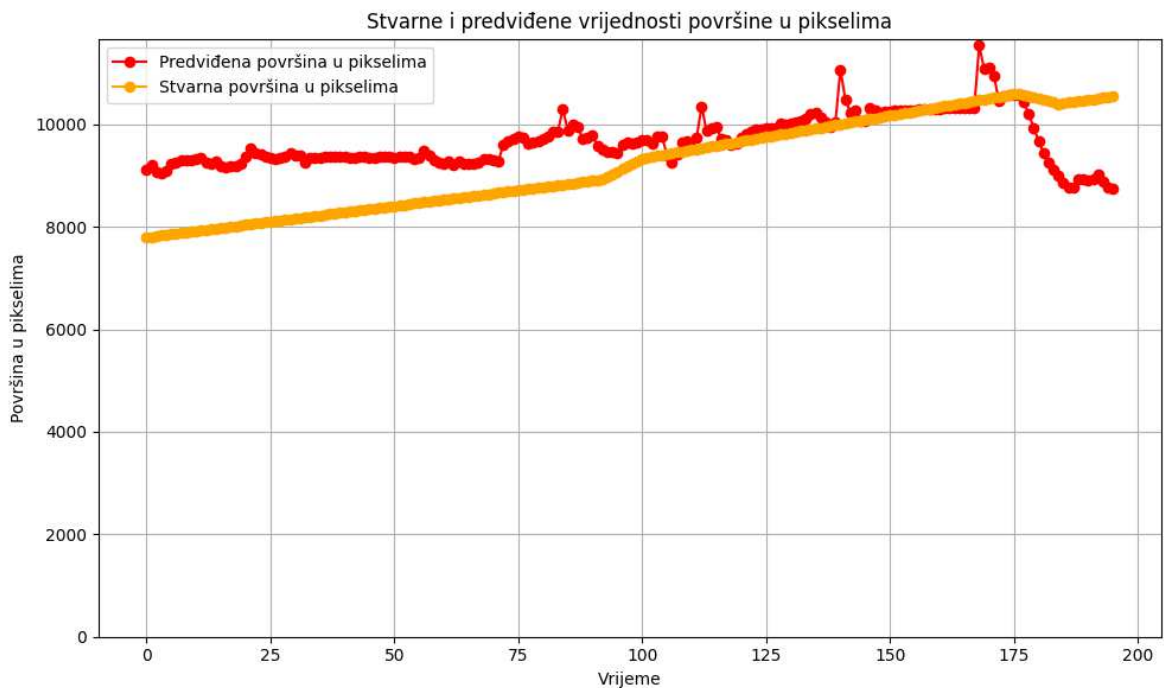
Sl. 4.8 Graf predviđenih i mjerenih vrijednosti površine u pikselima za idućih sedam dana s vremenskim korakom od šest sati

Kako bi evaluirali model na cijelom skupu podataka za jednu biljku korisno je usporediti točnost predviđanja modela u svim fenofazama tj. u svim dostupnim razvoja biljke.

U pojedinoj fenofazi vlažnost tla i padaline različito utječu na razvoj pšenice, npr. velika je razlika utjecaja ovih parametara u fazi razvoja stabljike i sazrijevanju zrna gdje povećane padaline više ne utječu na povećanje biomase. Ove razlike u predviđanju se vide na slikama 4.8 i 4.9 gdje je za predviđanja NDVI-a točnost modela po cijelom skupu podataka približno ista no u predviđanju površine biljke to nije slučaj. Predviđanje površine biljke puno je preciznije u blizini tjemena funkcije što se može pripisati različitosti rasta biomase svake biljke u prethodnim fenofazama, gdje bi se moglo naslutiti da je ovaj problem također uzrokovan nedostatkom podataka kako bi model općenitije predviđao ciljane parametre te da ne bi došlo do prenaučavanja modela.



Sl. 4.9 Graf usporedbe mjerenih i predviđenih vrijednosti NDVI-a za cijeli testni skup podataka



Sl. 4.10 Graf usporedbe mjerenih i predviđenih vrijednosti površine u pikselima za cijeli testni skup podataka



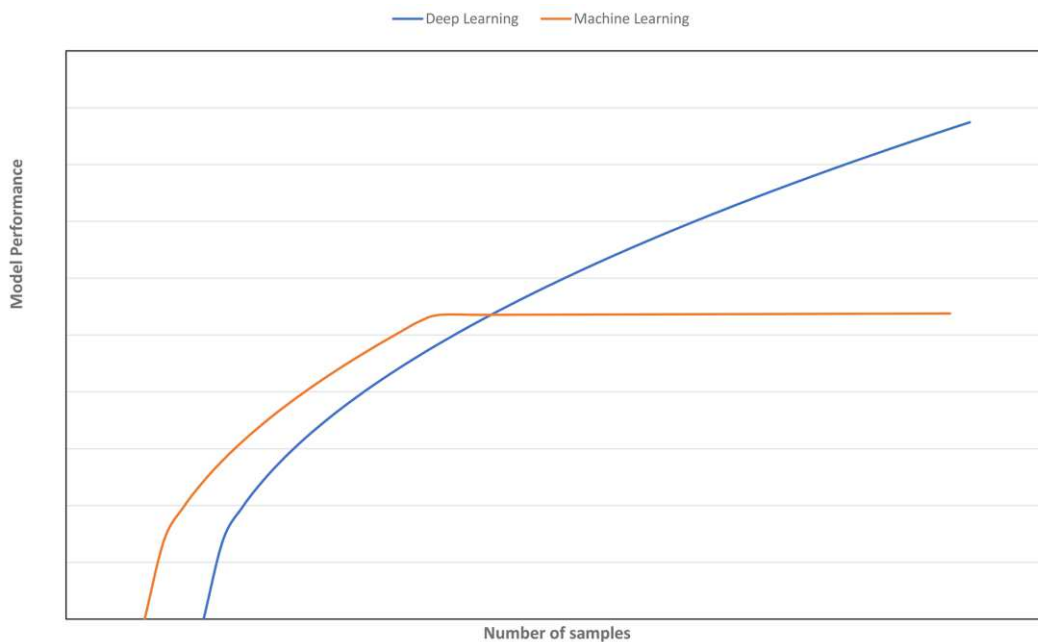
## 5. Usporedba rezultata svih korištenih modela u predikciji NDVI-a i površine u pikselima

Kako bi se što bolje predvidjele vrijednosti NDVI-a i površine pšenice na stvarnim podacima s terena potrebno je evaluirati i usporediti rezultate predikcije svih korištenih modela strojnog učenja. U tablici 5.1 prikazane su metrike svih korištenih modela.

Metoda	NDVI				Površina u pikselima			
	MAE	RMSE	MAPE [%]	R2	MAE	RMSE	MAPE [%]	R2
Linearna regresija	0.0022	0.0029	0.3841	0.9917	135	172	1.5043	0.9362
Lasso	0.0017	0.0024	0.2999	0.9947	106	135	1.1913	0.9599
MLP	0.0028	0.0038	0.5018	0.9862	179	223	2.0595	0.893
LSTM	0.0084	0.0098	1.7712	0.2246	541	634	6.2303	0.517

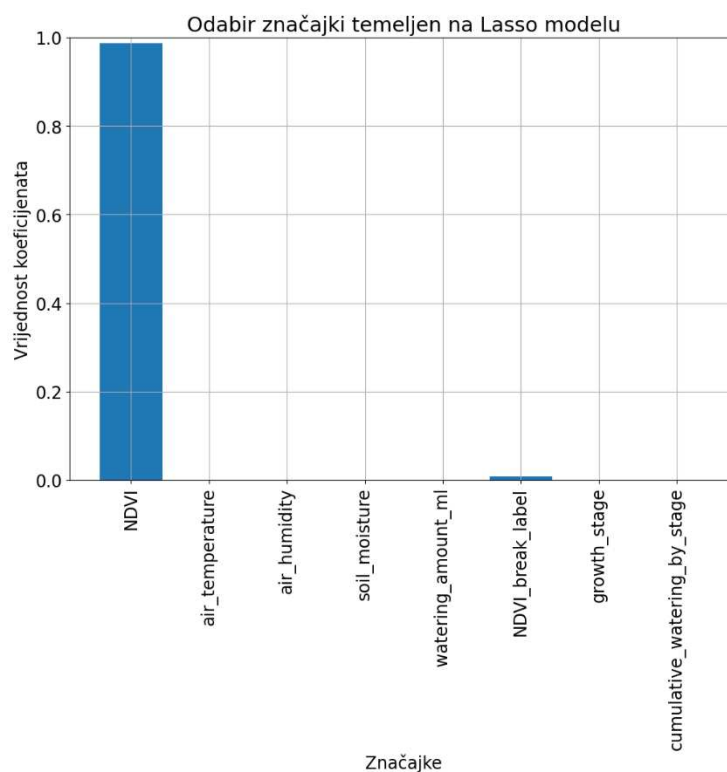
Tablica 5.1 Usporedba metrika korištenih metoda

Kao što je vidljivo iz tablice 5.1 najbolje metrike ima Lasso model, ovo je i za očekivati s obzirom na veličinu skupa podataka na kojima su se trenirali modeli. Također modeli očito puno lošije predviđaju površinu u pikselima što bi bilo objašnjeno s time da vrijednosti površine u pikselima u testnom skupu više variraju u odnosu na ostale biljke dok je graf vrijednosti NDVI-a svih biljki relativno sličan. U vidu teme ovoga rada predviđanje NDVI-a na stvarnim podacima s terena i testiranje režima navodnjavanja biti će obavljeno korištenjem LSTM i Lasso modelom (zbog najboljih metrika). Slika 5.1 prikazuje preformanse modela s obzirom na veličinu podataka i razliku između modela strojnog i dubokog učenja[12].



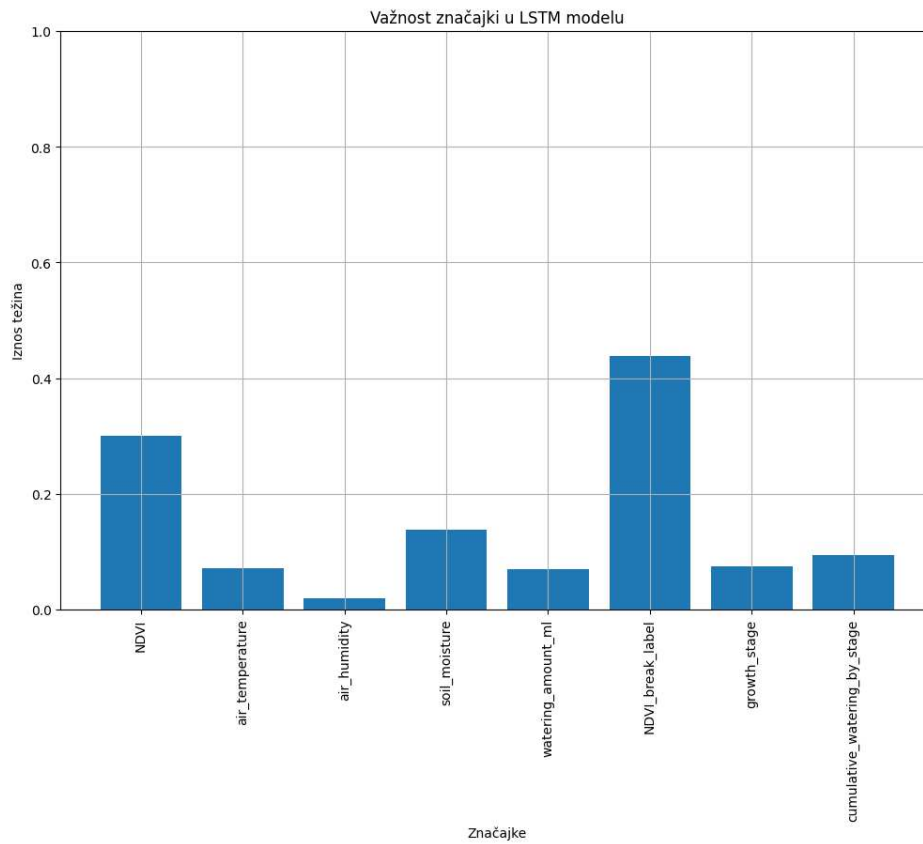
Sl. 5.1 Graf performansi modela s obzirom na veličinu skupa podataka [13]

Iako Lasso model daje najtočnije predikcije NDVI-a i površine u pikselima potrebno je provjeriti utjecaj svih značajki na predikcije modela. U ovu svrhu napravljena je analiza Lasso modela u vidu predikcije NDVI-a kojeg je na terenu jedino moguće testirati. Grafikon koji pokazuje utjecaj korištenih značajki na predikciju Lasso modela prikazan je na slici 5.2.



Sl. 5.2 Utjecaj pojedinih značajki na Lasso model

LSTM model s druge strane radi lošije predikcije no uzima u obzir puno više značajki s odgovarajućim težinama. Grafikon koji prikazuje težine po značajkama prikazan je na slici 5.3.



Sl. 5.3 Iznos težina po značajkama modela

Prema ovoj usporedbi kasnije je odabran LSTM model kako bi se testirao utjecaj pojedinih režima navodnjavanja na predikciju NDVI-a u poglavlju 7.

## **6. Testiranje naučenih modela na stvarnim podacima**

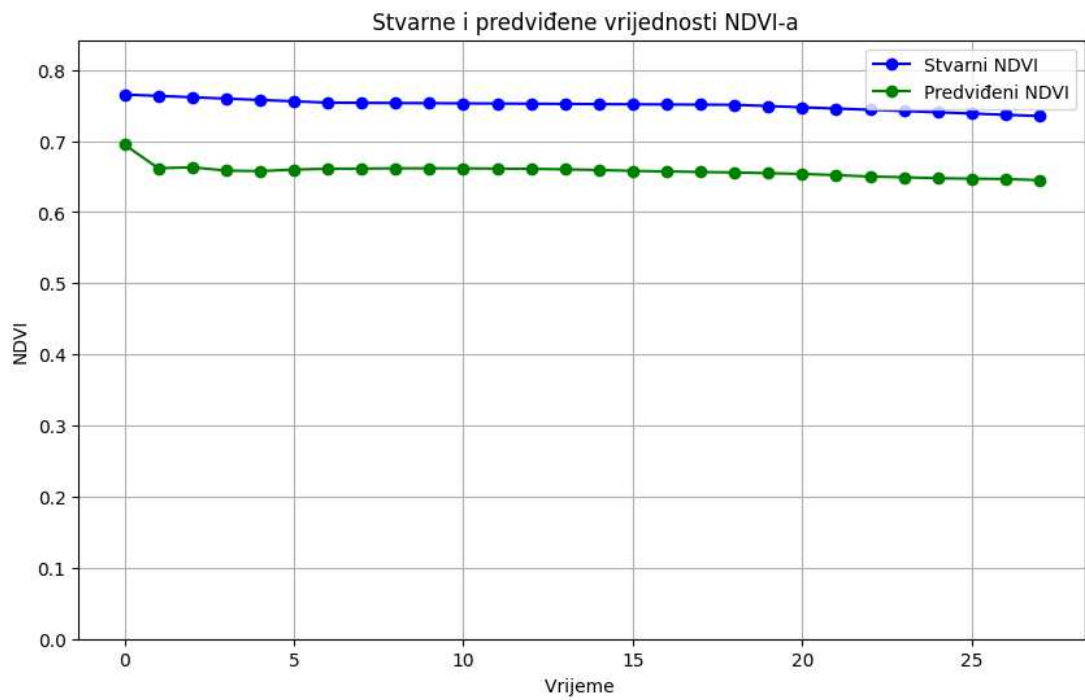
Kako bi se zapravo mogla odrediti učinkovitost korištenih modela potrebno je naučeni model testirati na stvarnim podacima s terena koristeći vremensku prognozu kao dodatan ulaz u model (u procesu učenja modela korišten je deterministički pristup vremenskoj prognozi npr. padaline). Podatci o biomasi odnosno površini biljke u pikselima nisu dio ranije dostupnih podataka sa terena te sa u ovu svrhu naučeni model ne može testirati.

S druge strane predviđanje NDVI vrijednosti moguće je ostvariti koristeći „remote – sensing pristup. Predviđanje pomoću podataka s terena obavljeno je korištenjem modela s najboljom metrikom (Lasso) i povratne neuronske mreže (LSTM) čije je korištenje tema ovoga rada.

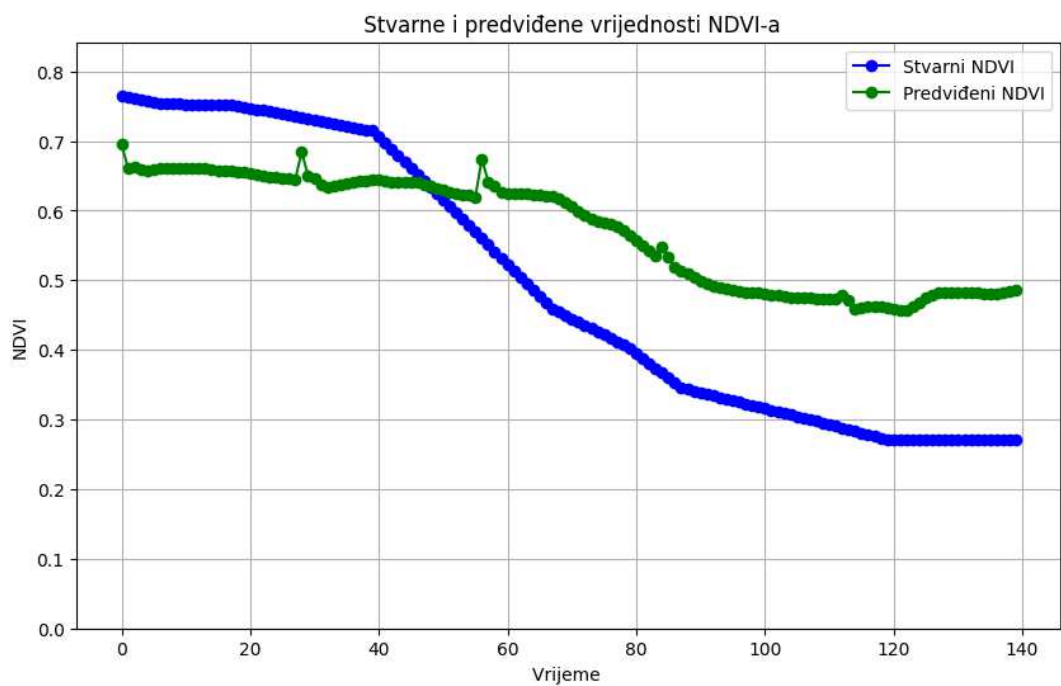
### **6.1. Testiranje LSTM modela na stvarnim podacima s terena**

Graf na slici 6.1 prikazuje usporedbu predviđanja idućih tjedana dana vrijednosti NDVI-indeksa s vremenskom rezolucijom od 6 sati. Na slici je vidljivo da u odnosu na testni skup koji sadrži podatke iz komora model predviđa niže vrijednosti NDVI-a što bi se moglo priroditi prenaučenosti tj. nedostatku podataka gdje model ne može dovoljno općenito predviđati ciljane vrijednosti.

Na slici 6.2 prikazan je graf testiranja na cijelom skupu podataka svakih 28 vremenskih koraka.



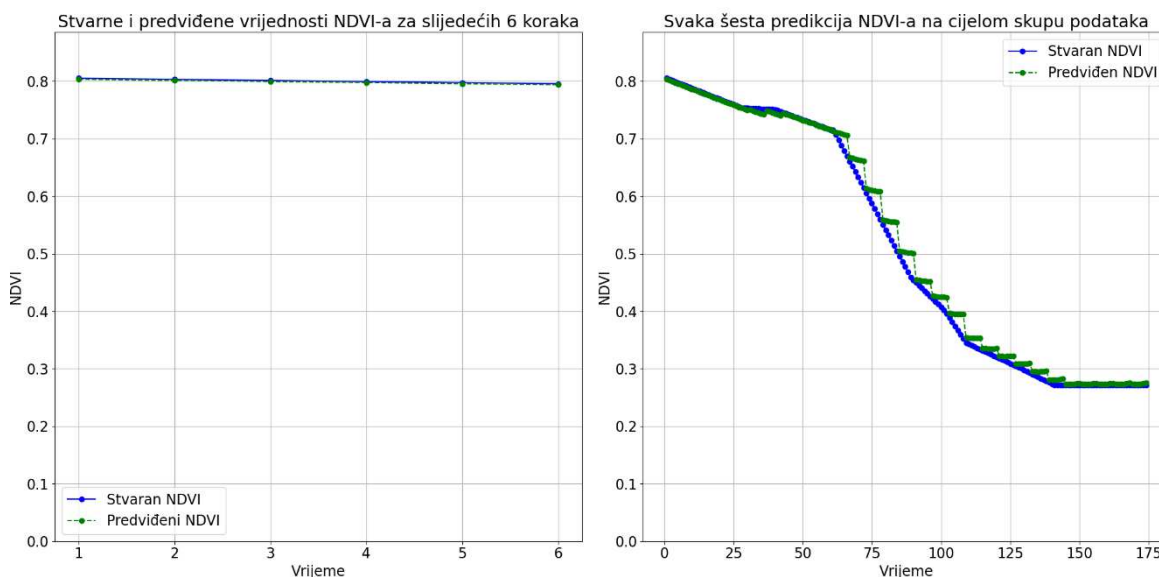
Sl. 6.1 Predviđene i stvarne vrijednosti NDVI-a za horizont od 28 vremenskih koraka (tjedan dana)



Sl. 6.2 Graf predviđenih i stvarnih vrijednosti NDVI-a na cijelom skupu podataka

## 6.1. Testiranje Lasso modela na stvarnim podacima s terena

Kako se Lasso metoda pokazala najpreciznija s obzirom na metrike testirana je i na stvarnim podacima s terena. Usporedba stvarnih i predviđenih vrijednosti prikazana je na slici 6.3.



Sl. 6.3 Graf koji prikazuje usporedbu stvarnih i predviđenih vrijednosti NDVI-a

## 6.2. Usporedba rezultata Lasso i LSTM modela

S obzirom na relativno malen skup podataka prednost u predviđanju daje se Lasso modelu u odnosu na LSTM. Rezultirajuće metrike oba modela testirana na podacima s terena prikazane su u tablici 6.1 gdje se opet vidi da Lasso model ima bolje performanse u odnosu na LSTM model.

Metoda	NDVI			
	MAE	RMSE	MAPE [%]	R2
Lasso	0.0098	0.0151	2.2697	0.9948
LSTM	0.15	0.1867	42.0136	0.2268

Tablica 6.1 Tablica metrika korištenih metoda

## 7. Testiranje utjecaja različitih režima navodnjavanja korištenjem LSTM modela

Kao što je istaknuto u petom poglavlju, odabran je LSTM model u svrhu testiranja utjecaja različitih režima navodnjavanja jer on kod predikcije uzima u obzir više značajki nego Lasso model koji uzima samo prošle vrijednosti NDVI-a i informaciju o tome je li došlo do loma funkcije (NDVI\_break\_label). Kao simulacija navodnjavanja tj. utjecaj promjene vlažnosti tla, temperature i padalina kreirani su „kišni“ dani sa umjerenom količinom padalina u trajanju od šest sati u iznosu od 2.5 mm/m<sup>2</sup> po satu.

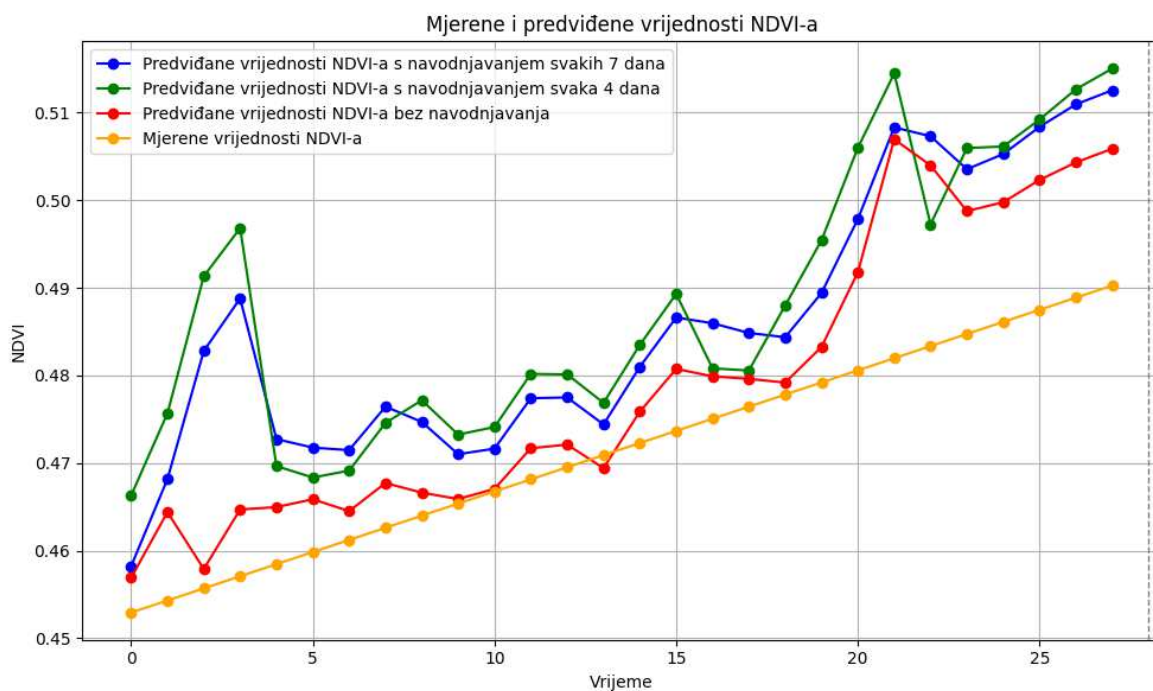
- Temperatura zraka smanjena za 20%
- Srednja vlažnost zraka postavljena na 95%
- Vlažnost tla postavljena na 80%
- Količina navodnjavanja u mililitrima postavljena na 56.25 ml
- Kumulativno navodnjavanje po fenofazi na cijelom skupu je uvećano za 168.85 ml za režim kiše svaka četiri dana i 112.5 ml za režim kiše svakih sedam dana što odgovara otprilike broju kišnih dana pomnoženih sa količinom navodnjavanja

Valja napomenuti da su vrijednosti padalina zbog lakše manipulacije podacima pretvorene u mililite jer je model treniran i u ovom slučaju testiran na laboratorijskim podacima.

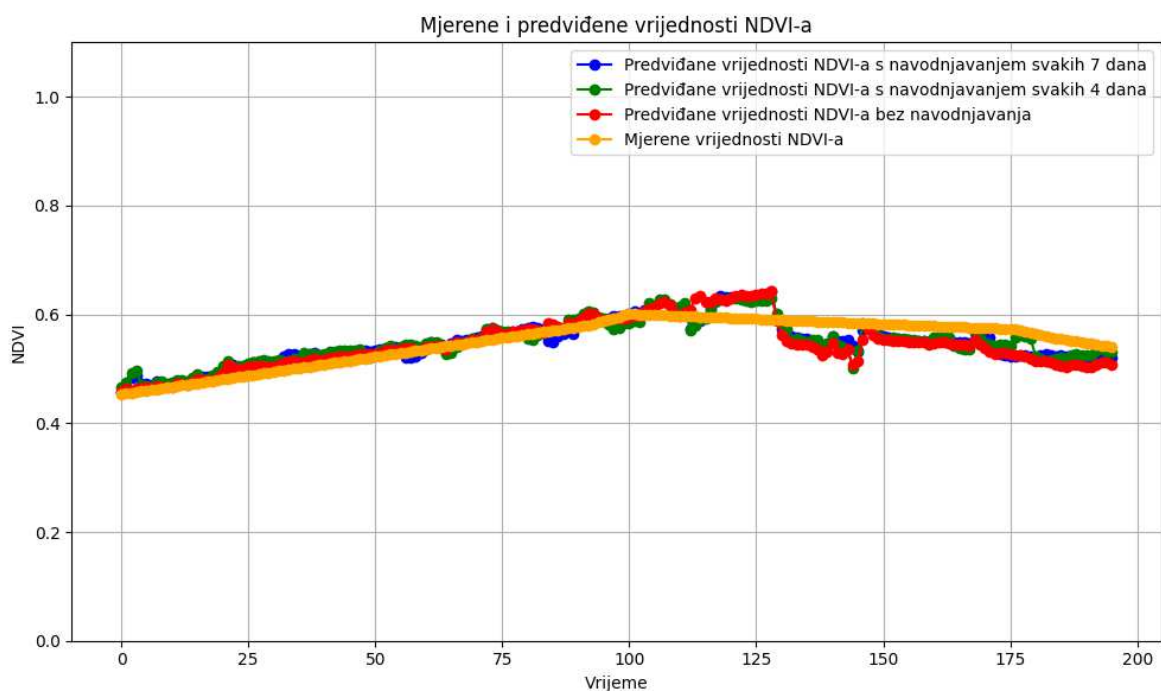
Utjecaj „kišnih“ dana nije testiran na podatkovnom skupu s terena jer kako je i pokazano u prošlom poglavlju niti jedan model ne predviđa dovoljno dobro vrijednosti s terena zbog neopćenitosti predikcija uzrokovanih uskim skupom podataka.

Testirani režimi su se sastojali od „navodnjavanja“ svaka četiri i svakih sedam dana što bi odgovaralo da je svaki četvrti ili sedmi dan šest sati padala umjerena kiša.

Gravofi koji pokazuju utjecaj ova dva režima prikazani su na slici 7.1 i 7.2.



Sl. 7.1 Usporedba dva režima navodnjavanja sa referentnom predikcijom za slučaj bez navodnjavanja za šest budućih koraka



Sl. 7.2 Usporedba dva režima navodnjavanja sa referentnom predikcijom za slučaj bez navodnjavanja na cijelom skupu podataka



## Zaključak

Podaci prikupljeni u okviru AgroSPARC projekta pružaju dubok uvid u rast i razvoj pšenice kroz analizu NDVI indeksa i površine biljke u pikselima. Metodologija koja obuhvaća korištenje daljinskog istraživanja, segmentaciju fotografija te analizu senzorskih podataka omogućila je precizno praćenje uvjeta rasta i razvoja biljaka. Važno je istaknuti da su dobiveni rezultati odgovarali očekivanjima, što potvrđuje valjanost korištenih metoda i alata za analizu. Integracija podataka o vlažnosti tla, fenofazi biljke te kumulativnom navodnjavanju ključna su za razumijevanje utjecaja tih faktora na NDVI i površinu biljke. Pokazano je da se može raditi predikcija NDVI-a i biomase s adekvatnim skupovima podataka te se pokazao utjecaj navodnjavanja na predikcije modela. Predikcije modela daju dobre rezultate samo u okoilini slučajeva koji su se dogodili u povijesti tj. u laboratorijskom okruženju jer se u podatkovnom skupu korištene samo „zdrave“ biljke. Daljnji rad u ovom smjeru mogao bi uključivati proširenje uzorka podataka kako bi se uključilo više scenarija s terena i iz laboratorijskog okruženja te naposljetku razvoj novih tehnologija za kontrolu navodnjavanja i uspješniji uzgoj poljoprivrednih kultura.

## Literatura

- [1] V. Lešić (voditelj), M. Zovko, M. Orsag, "AgroSPARC - Napredna i prediktivna poljoprivreda za otpornost klimatskim promjenama", projekt poziva EFRR, br. KK.05.1.1.02.0031, 2020-2023, Poveznica: <https://www.icent.hr/projekti/agrosparc>.
- [2] MicaSense. RedEdge-MX Professional Multispectral Sensor for Agriculture. Poveznica: <https://micasense.com/rededgemx/>
- [3] Iwake Masialeti, Stephen L. Egbert, Brian D Wardlow, *A Comparative Analysis of Phenological Curves for Major Crops in Kansas*. GIScience & Remote Sensing, 2010,47.2: 241-259. Poveznica: [https://www.researchgate.net/publication/228649433\\_A\\_Comparative\\_Analysis\\_of\\_Phenological\\_Curves\\_for\\_Major\\_Crops\\_in\\_Kansas](https://www.researchgate.net/publication/228649433_A_Comparative_Analysis_of_Phenological_Curves_for_Major_Crops_in_Kansas)
- [4] *Biomass Growth and Dry Matter Accumulation in Wheat*. AHDB Knowledge Library. Poveznica: <https://ahdb.org.uk/knowledge-library/biomass-growth-and-dry-matter-accumulation-in-wheat>.
- [5] Hyles, J., Bloomfield, M.T., Hunt, J.R. et al. Phenology and related traits for wheat adaptation. *Heredity* 125, 417–430 (2020). Poveznica: <https://doi.org/10.1038/s41437-020-0320-1>
- [6] *LinearRegression*, Scikit learn, Poveznica: [https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LinearRegression.html](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html)
- [7] *Lasso*, Scikit-learn, Poveznica: [https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.Lasso.html](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html)
- [8] *Sequential*, TensorFlow, Poveznica: [https://www.tensorflow.org/api\\_docs/python/tf/keras/Sequential](https://www.tensorflow.org/api_docs/python/tf/keras/Sequential)
- [9] *Time Series Forecasting with TensorFlow*, TensorFlow Tutorials. Poveznica: [https://www.tensorflow.org/tutorials/structured\\_data/time\\_series](https://www.tensorflow.org/tutorials/structured_data/time_series).
- [10] *Kako stanice dugog kratkoročnog pamćenja (LSTM) rješavaju problem dugih nizova podataka u RNN-ovima?*, ETICA Academy, (2023, kolovoz). Poveznica: <https://hr.eitca.org/artificial-intelligence/eitc-ai-dltf-deep-learning-with-tensorflow/recurrent-neural-networks-in-tensorflow/recurrent-neural-networks-rnn/examination-review-recurrent-neural-networks-rnn/how-do-long-short-term-memory-lstm-cells-address-the-issue-of-long-sequences-of-data-in-rnns/>
- [11] *Long short-term memory*, Wikipedia, (2024). Poveznica: [https://en.wikipedia.org/w/index.php?title=Long\\_short-term\\_memory&action=history](https://en.wikipedia.org/w/index.php?title=Long_short-term_memory&action=history)
- [12] Emmanuel Sirimal Silva, Hossein Hassani, Mansi Ghodsi, Zara Ghodsi, *Forecasting with auxiliary information in forecasts using multivariate singular spectrum analysis*. Information Sciences. Volume 479, (2019), Stranice: 214-230, ISSN 0020-0255, Poveznica: <https://doi.org/10.1016/j.ins.2018.11.053>

- [13] Ng, W., Minasny, B., Mendes, W. D. S., and Demattê, J. A. M.: The influence of training sample size on the accuracy of deep learning models for the prediction of soil properties with near-infrared spectroscopy data, *SOIL*, 6, (2020), 565–578, Poveznica:<https://doi.org/10.5194/soil-6-565-2020>

## Sažetak

Kako bi bila osigurana veća sigurnost opskrbe hranom potrebno je razvijati nove metode upravljanja i kontrole rasta i razvoja usjeva s naglaskom na navodnjavanje. U ovu svrhu razvijen je model povratne neuronske mreže koji predviđa vrijednosti NDVI-a i površine pšenice koja je korištena kao alternativa za biomasu. Model je testiran na stvarnim podacima s terena te se pokazao kao zadovoljavajuć s obzirom na opseg dostupnih podataka. Istražen je i remote-sensing pristup koji obuhvaća prikupljanje podataka o NDVI vrijednostima korištenjem Sentinel 2 satelita te podataka o prognozi korištenjem Open-Weather platforme te uparivanje tih vrijednosti sa stvarnim mjerenjima na terenu kako bi se dobila što bolja predikcija ciljanih parametara. Napravljeno je testiranje utjecaja navodnjavanja na NDVI te je donešen zaključak da navodnjavanje ima očekivani utjecaj na predikcije NDVI-a no potrebni su veći skupovi podataka kako bi predikcije modela bile općenitije.

**Ključne riječi:** NDVI, biomasa, segmentacija fotografija, povratne neuronske mreže, Lasso, remote-sensing, navodnjavanje

## Summary

In order to ensure greater security of food supply, it is necessary to develop new methods of managing and controlling the growth and development of crops with an emphasis on irrigation. For this purpose, a recurrent neural network model was developed that predicts the NDVI values and the area of wheat plant that was used as an alternative for biomass. The model was tested on real data from the field and proved to be satisfactory considering the extent of available data. A remote-sensing approach was also investigated, which includes collecting data on NDVI values using the Sentinel 2 satellite and forecast data using the Open-Weather platform, and matching these values with actual measurements in the field in order to obtain the best possible prediction of target parameters. Testing of different irrigation models concluded that irrigation does have an expected effect on NDVI predictions but bigger datasets are needed to better generalize the predictions of the models.

**Keywords:** NDVI, biomass, image segmentation, recurrent neural networks, Lasso, remote-sensing, irrigation