

Modeli strojnog učenja za predviđanje metastaza u limfne čvorove u bolesnika s rakom dojke

Vrdoljak, Josip

Doctoral thesis / Disertacija

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Split, School of Medicine / Sveučilište u Splitu, Medicinski fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:171:497333>

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

Download date / Datum preuzimanja: **2024-11-29**



Repository / Repozitorij:

[MEFST Repository](#)



**SVEUČILIŠTE U SPLITU
MEDICINSKI FAKULTET**

JOSIP VRDOLJAK, dr. med.

**MODELI STROJNOG UČENJA ZA PREDVIĐANJE METASTAZA
U LIMFNIM ČVOROVIMA U BOLESNIKA S RAKOM DOJKE**

DOKTORSKA DISERTACIJA

SPLIT, 2023.

Ovaj rad izrađen je pri Katedri za patofiziologiju Medicinskog fakulteta Sveučilišta u Splitu

Voditelj rada: izv. prof. dr. sc. Joško Božić, dr. med.

ZAHVALA

Želim zahvaliti svim suradnicima i kolegama koji su sudjelovali u izradi ovoga istraživanja.

Osobito zahvaljujem Jošku Božiću na pomoći, ukazanim prilikama i savjetima, a iznad svega na prijateljstvu.

Hvala mojim roditeljima na podršci kroz sve godine studija koje su kulminirale ovom disertacijom.

Velika hvala mojoj supruzi Miji i sinu Pavlu što su mi glavni izvor motivacije i nadahnuća – zbog vas sve što radim ima smisla!

GRACIAS AGO TIBI DOMINE QUIA FUI IN HOC MUNDO.

SADRŽAJ

1. UVOD	1
1.1. Rak dojke.....	2
1.1.1. Epidemiologija	2
1.1.2. Etiologija	3
1.1.3. Patofiziologija.....	4
1.1.4. Patohistološka klasifikacija	5
1.1.4. TNM klasifikacija.....	7
1.1.5. Imunohistokemijska klasifikacija	7
1.1.6. Podjela prema gradusu.....	8
1.1.7. Dijagnostičke metode	9
1.1.8. Liječenje raka dojke.....	10
1.1.9. Neoadjuvantno liječenje	12
1.2. Strojno učenje.....	13
1.2.1. Uvod u strojno učenje	13
1.2.2. Strojno učenje u medicini i onkologiji	15
1.2.3. Stablo odluke.....	16
1.2.4. K-najbliži susjedi („k-nearest neighbors“)	17
1.2.5. Slučajne šume.....	19
1.2.6. Stabla s podizanjem gradijenta („Gradient Boosted Trees“).....	20
1.2.7. TabNet (artificijalna neuronska mreža za tabularne podatke)	22
1.2.8. Validacija modela strojnog učenja	23
1.2.9. Objašnjivo strojno učenje	25
1.3. Strojno učenje u klasifikaciji limfnih čvorova u raku dojke	26
1.3.1. Modeli koji koriste radiološke slike/parametre (Radiomika).....	26
1.3.2. Modeli na kliničko-patološkim podacima.....	27
1.3.3. Kombinirani pristup	28
2. CILJEVI I HIPOTEZE ISTRAŽIVANJA	30
3. ISPITANICI I POSTUPCI	33
3.1. Ispitanici	34
3.2. Postupci	34
3.2.1. Eksploratorna analiza podataka i čišćenje podataka.....	34
3.2.2. Modeliranje i optimizacija.....	35
3.2.3. Performanse modela	38
3.2.4. Analiza važnosti prediktora i objašnjivost modela	38
4. REZULTATI	39
4.1. Karakteristike ispitanika.....	40
4.2. Performanse modela strojnog učenja	42
4.2.1. Performanse za NST- skupinu.....	42
4.2.2. Performanse za ukupnu populaciju.....	44
4.2.3. Analiza važnosti prediktora	46
5. RASPRAVA	50
6. ZAKLJUČCI	57
7. SAŽETAK	60
8. SUMMARY	63
9. LITERATURA	66
10. ŽIVOTOPIS	75
11. DODATAK	80

POPIS OZNAKA I KRATICA

COVID-19 – bolest uzrokovana virusom SarsCov2 (engl. *Coronavirus Disease „2019“*)

ITM – indeks tjelesne mase

BRCA1 i BRCA2 – tumor supresorski geni (engl. *BReast CAncer Genes 1 and 2*)

DNK – deoksiribonukleinska kiselina

P53 – gen s funkcijom tumor supresora i onkogen

PI3K/AKT – fosfatidilinozitol 3-kinazni / protein kinaza B put

PTEN – fosfataza i tenzin homolog deletiran na desetom kromosomu (engl. *Phosphatase and tensine homolog deleted at chromosome ten*)

DCIS – duktalni in situ karcinom

LCIS – lobularni in situ karcinom

NOS – nespecificirani tip karcinoma dojke (engl. *Not Otherwise Specified*)

IDC – invazivni duktalni karcinom (engl. *Invasive Ductal Carcinoma*)

TNM klasifikacija – tumor, limfni čvor, metastaza klasifikacija (engl. *Tumour, Node, Metastasis Classification*)

ER – estrogenski receptor

PR – progesteronski receptor

HER-2 – receptor za humani epidermalni čimbenik rasta 2

Ki67 indeks – nuklearni protein koji se koristi kao marker proliferacije stanica, indeks proliferacije stanica

UZV – ultrazvuk

MRI – magnetska rezonancija

CT – kompjutorska tomografija

SLN – limfni čvor čuvar (engl. *Sentinel Lymph Node*)

LČ – limfni čvorovi

DFS – preživljenje bez povrata bolesti (engl. *Disease Free Survival*)

OS – ukupno preživljenje (engl. *Overall Survival*)

NST – neoadjuvantno sustavno liječenje

NCCN – (engl. *National Comprehensive Cancer Network*)

KNN – algoritam K- najbližih susjeda (engl. *K-Nearest Neighbors*)

ANN – artifičijalna neuronska mreža

TabNet – artifičijalna neuronska mreža za tabularne podatke (engl. *Tabular Network*)

ROC – krivulja međuodnosa osjetljivosti i specifičnosti (engl. *Receiver Operating Characteristic*)

AUC – površina ispod ROC krivulje (engl. *Area Under the Curve*)

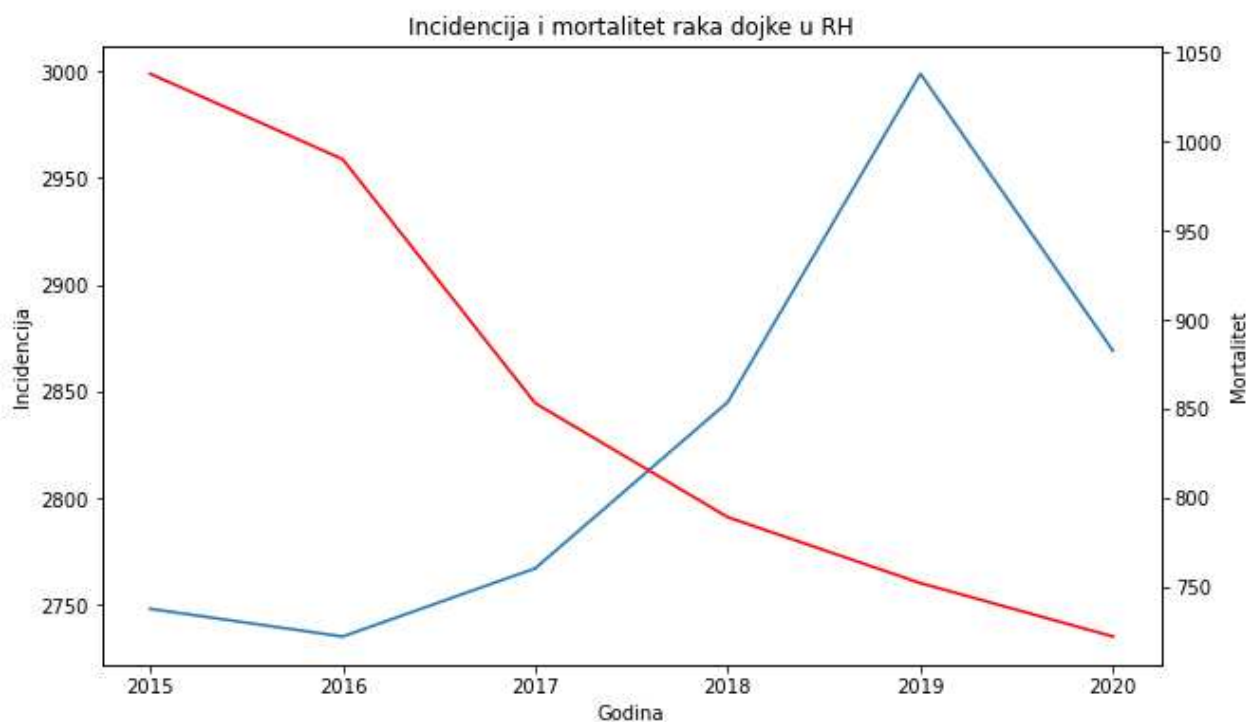
1. UVOD

1.1. Rak dojke

1.1.1. Epidemiologija

Rak dojke najčešći je maligni tumor u žena te najviše pridonosi smrtnosti od raka što definira veliku javno-zdravstvenu važnost ove bolesti (1). Prema posljednjim podacima Američkog društva za rak, incidencija raka dojke porasla je tijekom protekla četiri desetljeća. Od 2010. do 2019. godine stopa se povećala za 0.5% godišnje, uglavnom potaknuta lokaliziranim stadijima i hormonski pozitivnom bolesti (2). Smrtnost od raka dojke pala je za 43% u razdoblju od 1989. do 2020. godine što odražava značajne napretke u dijagnostičkim i terapijskim modalitetima za rak dojke.

Incidencija raka dojke u Hrvatskoj posljednjih nekoliko godina u stalnom je porastu, osim 2020. godine kada je zabilježen pad incidencije, vjerojatno zbog smanjenog broja dijagnostičkih postupaka uzrokovanih pandemijom bolesti COVID-19. Točnije, prema podacima iz 2020. godine u Hrvatskoj dijagnosticirano je 2869 slučajeva raka dojke, dok je u istoj godini od raka dojke umrlo 722 bolesnika (3).



Slika 1. Trend incidencije i mortaliteta raka dojke u RH (2015. - 2020.) (3.)

1.1.2. Etiologija

Rak dojke složena je i multifaktorijska bolest, a njezina etiologija uključuje kombinaciju genetskih, hormonalnih i okolišnih čimbenika.

1.1.2.1. Genetski čimbenici

Poznato je da nasljedne mutacije gena, kao što su BRCA1 i BRCA2, značajno povećavaju rizik od raka dojke (4). Ostali značajni geni uključuju TP53, PTEN i CHEK2 (4). Nadalje, obiteljska anamneza raka dojke može povećavati rizik za oboljenje, posebno ako je oboljeli bliži srodnik (5).

1.1.2.2. Hormonalni čimbenici

Doživotna izloženost estrogenima igra ključnu ulogu u razvoju raka dojke. Čimbenici koji doprinose izloženosti estrogenima uključuju ranu menarhu, kasnu menopauzu i hormonsku nadomjesnu terapiju (6). Također, reproduktivni čimbenici poput nulipariteta, kasne dobi pri prvom porodu i nedostatka dojenja mogu povećati rizik od raka dojke (6).

1.1.2.3. Okolišni čimbenici

Nekoliko čimbenika životnoga stila povezano je s rizikom od raka dojke, kao što su konzumacija alkohola, pušenje, tjelesna neaktivnost i pretilost (kao i postmenopauzalni indeks tjelesne mase) (7).

Zatim, izloženost ionizirajućem zračenju (zračenje prsišta poradi liječenja drugog tumora npr. limfoma), posebno tijekom adolescencije, može povećati rizik od raka dojke kasnije u životu (8). Suprotno tomu, kod žena starijih od 40. godina izloženost zračenju nije povezana s povećanim rizikom za oboljenje (9).

Među ostale faktore rizika spada starija životna dob (gdje većina slučajeva raka dojke nastaje oko 50. godine starosti), kao i povećana gustoća dojke (žene s gušćim tkivom dojke imaju veću vjerojatnost da će razviti rak) te prisutnost benignih čvorova na dojci s atipijom (10).

1.1.3. Patofiziologija

Poput ostalih tumora, i patofiziologija raka dojke fundamentalno se tiče poremećaja u regulaciji dijeljenja stanica. Kako bi se stanica transformirala u tumorsku stanicu, geni koji reguliraju stanični rast i diferencijaciju moraju biti nefunkcionalni/manjkavi (kao što je slučaj kod tumor supresorskih gena), ili pak pretjerano eksprimirani (kod onkogeni). Genske se promjene mogu dogoditi na više razina, od kromosoma do mutacija koje zahvaćaju samo jedan nukleotid DNK (12). Geni koji sudjeluju u patofiziologiji tumora mogu se podijeliti u dvije glavne kategorije: 1) onkogene i 2) tumor supresorske gene. Onkogeni mogu biti geni bez mutacije koji su pretjerano izraženi (npr. poradi pojačane aktivnosti promotorskih regija), ili pak promijenjeni geni koji su zadobili novu funkciju. U oba slučaja, rezultat je pretjerano promoviranje staničnoga rasta. S druge strane, tumor supresorski geni odgovorni su za promoviranje anti-proliferacijskih signala koji zaustavljaju mitozu i stanični rast. To su uglavnom transkripcijski čimbenici aktivirani u slučaju staničnog stresa i oštećenja DNK. Funkcija tumor supresorskih gena je zaustavljanje progresije staničnog ciklusa kako bi se mogao odviti adekvatan popravak DNK, čime se sprječava prenošenje mutacije na novonastale stanice. Već spomenuti BRCA1 i BRCA2 geni primjer su tumor supresora čija mutacija dovodi do pojave raka dojke i raka jajnika u ranijoj životnoj dobi (11).

Još 1971. godine ustanovljena je teorija o nastanku raka koja ovisi o dva mutacijska događaja, Knudsonova hipoteza o „dvostrukom hitcu“ (12). U Knudsonovoj hipotezi, prvi hitac predstavlja naslijeđena mutacija u tumor supresorskom genu koja sama po sebi neće dovesti do raka jer je drugi alel tumor supresorskog gena još uvijek funkcionalan. Tek ukoliko tijekom života dođe i do mutacije drugog alela, razvija se tumor. Stoga, tumor supresorski geni predstavljaju recesivne gene jer sadrže takozvane mutacije s gubitkom funkcije (engl. *loss of function*). Suprotno, onkogeni predstavljaju gene s mutacijom dobitka funkcije (engl. *gain of function*) te stoga imaju dominantan fenotip. Kasnija istraživanja dovela su Knudsonov model

u pitanje jer se pokazalo kako je i mutacija jednog alela tumor supresorskog gena dovoljna za nastanak tumora. Ovaj se fenomen naziva haploinsuficijencija i pokazan je kroz više eksperimentalnih pristupa. Tumori uzrokovani haploinsuficijencijom obično nastaju u starijoj životnoj dobi u usporedbi s onima nastalim zbog procesa dvostrukog hitca (13).

Gen povezan s nastankom većine tumora, pa tako i rakom dojke, je p53 koji djeluje kao transkripcijski čimbenik koji sudjeluje u regulaciji staničnog ciklusa i apoptoze. Također je pokazana uloga p53 u prebačanju staničnog metabolizma s respiratorne fosforilacije na glikolizu, što po Warburgovoj hipotezi podržava rast tumorskih stanica (14, 15).

Poremećaji u procesu apoptoze, poglavito mutacije u PI3K/AKT kinaznom putu uključene su u patofiziologiji raka dojke. Naime, u nekim tumorima dojke pokazana je promjena u PTEN proteinu, što dovodi do pretjerane aktivacije PI3K/AKT i naposljetku do izbjegavanja apoptoze (16). Nadalje, novonastale tumorske stanice uspijevaju izbjeći imunološki odgovor mehanizmima poput antigenskog prikrivanja i otpuštanja imunosupresorskih molekula što je ključno za preživljenje tumora, njegov daljni rast i širenje (20).

Zaključno, kombinacija nasljednih čimbenika i okolišnih čimbenika dovode do nekontrolirane stanične proliferacije, izbjegavanja imunološkog odgovora, izostanka apoptoze te nastanka raka dojke (20).

1.1.4. Patohistološka klasifikacija

Prema klasifikaciji Svjetske zdravstvene organizacije iz 2019. godine, epitelne tumore dojke možemo podijeliti na benigne epitelne tvorbe i prekursorske lezije (tipična i atipična duktalna hiperplazija, lezije kolumnarnih epitelnih stanica), adenozne lezije i benigne sklerozirajuće lezije (sklerozirajuća adenoza, apokrina adenoza i adenom, mikroglandularna adenoza, radijalni ožiljak), adenome (tubularni, laktacijski i duktalni adenom), epitelno-

mioepitelne tumore (pleomorfni adenom, adenomioepiteliom i maligni adenomioepiteliom), papilarne neoplazme (intraduktalni papilom, papilarni duktalni karcinom in situ, inkapsulirani papilarni karcinom, solidni papilarni karcinom i invazivni papilarni karcinom), ne-invazivne lobularne neoplazme (atipična lobularna hiperplazija i lobularni karcinom in situ - LCIS), duktalni karcinom in situ (DCIS), na invazivne karcinome dojke (bez osobitih značajki – engl. *No Special Type* – NST, mikroinvazivni, lobularni, tubularni, kribiformni, mucinozni, mucinozni cistadenokarcinom, invazivni mikropapilarni s apokrinom diferencijacijom te metaplastični) te naposljetku na rijetke tumore i tumore po tipu salivatorne žlijezde (sekretorni, adenoid-cistični, mukoepidermoidni, polimorfni) kao i na neuroendokrine neoplazme (17, 18).

Također, po novoj klasifikaciji, mikroinvazivni karcinomi označavaju invazivne karcinome gdje veličina karcinoma nije prešla 1mm (17, 18).

Kod invazivnih karcinoma dojke, veliku većinu (oko 80%) novodijagnosticiranih slučajeva sačinjavaju karcinomi bez posebnih osobitosti (engl. *No Special Type*- NST, prethodno zvani *Not Otherwise Specified*- NOS) (18). Tu kategoriju sačinjavaju tumori koji ne prikazuju dovoljno karakteristika da bi bili morfološki klasificirani u neki od specifičnih histoloških podtipova. U prethodnim klasifikacijama oni su se zvali invazivnim duktalnim karcinomima (IDC), što je originalno trebalo reflektirati pretpostavljeno podrijetlo karcinoma (isto vrijedi i za lobularni karcinom), ali se poslije pokazalo krivim (iako je zbog tradicije invazivni lobularni karcinom zadržao svoje ime) (19). Drugi po učestalosti je invazivni lobularni karcinom (5-15%) te su ostali histološki podtipovi značajno rjeđi (22).

1.1.4. TNM klasifikacija

Rak dojke također svrstavamo prema TNM klasifikaciji (veličina tumora, status limfnih čvorova i status udaljenih presadnica) (20). Prema TNM klasifikaciji određujemo stadij bolesti. Po veličini karcinome dojke dijelimo u T1 (≤ 2 cm), T2 (≥ 2 cm ali ≤ 5 cm), T3 (≥ 5 cm) te T4 (tumor bilo koje veličine koji doseže kožu). S obzirom na status limfnih čvorova tumore dijelimo na N0 (bez presadnica u regionalnim limfnim čvorovima), N1 (pomični limfni čvorovi s presadnicama), N2 (nepomični limfni čvorovi s presadnicama), N3 (presadnice prisutne u unutarnjim mamarnim limfnim čvorovima). Naposljetku, s obzirom na prisutnost udaljenih presadnica dijelimo rak dojke na M0 (bez udaljenih presadnica), M1 (s udaljenim presadnicama) (20).

Prema tome, stadij I raka dojke smatramo T1-N0-M0; stadij IIA (T0-N1, T1-N1, T2-N0); stadij IIB (T2-N1-M0, T3-N0-M0); stadij IIIA (T0-N2-M0, T1-N2-M0, T2-N2-M0, T3-N1/2-M0); stadij IIIB (T4-bilo koji N- M0), IIIC (bilo koji T-N3-M0), IV (bilo koji T, bilo koji N, M1) (20).

1.1.5. Imunohistokemijska klasifikacija

Nakon uzimanja patohistološkog uzorka, imunohistokemijski određuje se prisutnost hormonskih receptora (ER i PR), receptora za humani epidermalni čimbenik rasta (HER2) i indeks proliferacije stanica (Ki67 indeks) (21). Ki67 indeks (ili proliferacijski indeks) određuje se prema ekspresiji Ki67 proteina koji se nalazi u staničnoj jezgri. Ki67 se eksprimira u svim fazama staničnog ciklusa osim G0 faze. Indeks se računa prema sljedećoj formuli (22):

$$PI = \frac{N_{Ki-67(+)}}{N_{Ki-67(+)} + N_{Ki-67(-)}} \times 100\%$$

PI- Proliferacijski indeks, NKi-67(+)- broj Ki67-pozitivnih stanica, NKi-67(-)- broj Ki67-negativnih stanica

Na temelju imunohistokemijske analize rak dojke možemo klasificirati u 5 različitih imunohistokemijskih podtipova. Luminalni A podtip označava karcinome koji sadrže ER i PR te je Ki67 < 20%. Luminalni B podtip označava karcinome koji također sadrže ER i PR, ali je Ki67 > 20%. Luminalne B dalje možemo podijeliti na Luminal B/HER2 +, s obzirom na prisutnost HER2 receptora (23). HER-2 pozitivnim tumorima označavamo tumore s negativnim hormonskim receptorima koji sadrže HER2 receptore. Naposljetku, trostruko negativne tumore označavamo uz izostanak ER, PR i HER2 receptora (21, 23).

Tablica 1. Imunohistokemijska podjela raka dojke (21)

Imunohistokemijska podjela raka dojke	
Luminalni A tip	ER +, PR +, Ki67 < 20%
Luminalni B tip	ER +, PR +, Ki67 > 20%.
Luminalni B, HER2+	ER +, PR +, HER2 +, Ki67 > 20%.
HER2 pozitivni	ER -, PR-, HER2 +
Trostruko negativni (engl. <i>Basal like</i>)	ER -, PR -, HER2 -

1.1.6. Podjela prema gradusu

Tumore dojke također dijelimo u tri gradusa; prema stupnju diferencijacije stanica, pleomorfizmu i razini mitotičke aktivnosti (24). Često se koristi „*Nottingham Histologic Score*“ sustav za gradiranje koji uzima u obzir sljedeće značajke: 1) količina stvaranja žlijezda, 2) značajke stanične jezgre (stupanj pleomorfizma), 3) mitotička aktivnost (stupanj proliferacije). Svaka od ovih značajki boduje se od 1 - 3, a zatim se rezultati zbrajaju kako bi se dobio konačni ukupni rezultat u rasponu od 3 - 9. Tumori gradusa I imaju ukupni zbroj od 3 - 5, tumori gradusa II imaju ukupni zbroj 6 - 7, dok tumori gradusa III imaju ukupni zbroj 8 - 9 (24).

1.1.7. Dijagnostičke metode

1.1.7.1. Dijagnostika raka dojke

Dijagnosticiranje raka dojke kreće od kliničkog pregleda (uz specifične znakove i simptome) do primjene radioloških tehnika (mamografija, UZV, CT, MRI) u svrhu vizualiziranja tumorske tvorbe te potom i uzimanja biopsije te određivanja patohistološkog nalaza (25). Završna dijagnoza postavlja se određivanjem patohistološkog i imunohistokemijskog tipa bolesti, nakon čega slijedi stupnjevanje bolesti (engl. *staging*). Razlikujemo „kirurško“ stupnjevanje, gdje se status metastatskih limfnih čvorova odredio nakon kirurškog zahvata i „slikovno“ stupnjevanje, gdje se status metastatskih limfnih čvorova odredio slikovnim metodama (UZV, MRI, CT). Status udaljenih metastaza se također prvotno određuje slikovnim metodama, nakon čega se može potvrditi uzimanjem patohistološkog uzorka (25).

1.1.7.2. Određivanje statusa limfnih čvorova

Kod bolesnika određenih za kirurško liječenje status limfnih čvorova (LČ) određuje se biopsijom limfnog čvora „čuvara“ (SLN) (engl. *sentinel lymph node biopsy*) te u slučaju pozitivnih SLN ili klinički pozitivnih LČ aksile, daljnjom diseksijom aksile (26).

Nadalje, kod bolesnika određenih za neoadjuvantno liječenje, status limfnih čvorova primarno se određuje kliničkim pregledom i dijagnostičkim metodama (UZV, CT, MR i mamografija). Od navedenih dijagnostičkih metoda najčešće se koristi ultrazvuk u svrhu klasifikacije limfnih čvorova aksile (27). Dosadašnja istraživanja pokazala su kako slikovne metode u čak 30% slučajeva pogrešno klasificiraju limfne čvorove (28). Specifično, za osjetljivost ultrazvuka u detekciji metastatskih limfnih čvorova, prati se varijacija od 47% - 87%, što svjedoči o značajnom utjecaju ljudskog faktora (29). Nadalje, osjetljivost CT-a u određivanju statusa LČ je oko 72%, dok je osjetljivost MRI 75-80% (ovisno o različitim

istraživanjima) (29). Prema tome, postoji potreba za unaprijeđenje određivanja statusa limfnih čvorova kod bolesnika s rakom dojke.

Zbog velikog rasta u količini i dostupnosti medicinskih podataka otvara se prilika za korištenje novih dijagnostičkih modaliteta. Uporabom podaktovne znanosti (engl. *Data science*) i strojnog učenja (engl. *Machine learning*) možemo suplementirati ili čak zamijeniti dosadašnje metode u svrhu postizanja personaliziranijeg pristupa liječenju. Metode strojnog učenja korištene su i za klasifikaciju metastatskih LČ u bolesnika s rakom dojke te dobrom prediktivnom moći pokazuju potencijal za kliničku uporabu, a što će detaljnije biti opisano u nastavku disertacije.

1.1.8. Liječenje raka dojke

Odabir terapijskog modaliteta ovisi o gradusu tumora, stadiju bolesti i molekularnom podtipu, kako bi postigli što personaliziraniji, sigurniji i učinkovitiji pristup liječenju. Raniji oblici raka dojke liječe se kirurškom ekscizijom (30). Standardni pristupi operaciji dojke su ili potpuna ablacija dojke (mastektomija), nakon koje obično slijedi rekonstrukcija dojke ili pošteđni kirurški zahvat (npr. lumpektomija ili kvadrantektomija). Lumpektomija podrazumijeva eksciziju tumora dojke s rubom okolnog normalnog tkiva. Studije pokazuju da su totalna mastektomija i lumpektomija + zračenje ekvivalentni u pogledu preživljenja bez povrata bolesti (engl. *Disease Free Survival – DFS*) i ukupnog preživljenja (engl. *Overall Survival - OS*) (30).

Nakon operacije može se ordinirati adjuvantna terapija (radioterapija, kemoterapija, hormonska terapija ili kemoimunoterapija). Radioterapija se može primijeniti na područje cijele dojke ili ciljano (nakon pošteđnog kirurškog zahvata), na područje prsnog koša (nakon mastektomije) ili na područje aksile (za regionalnu kontrolu bolesti) (31).

Kemoterapeutici primjenjivani u liječenju raka dojke spadaju u različite skupine citotoksika, od alkilirajućih lijekova, antimetabolita, do inhibitora tubulina (32). Među antraciklinima koriste se doxorubicin, daunorubicin, epirubicin i idarubicin. Mehanizam djelovanja antraciklina sastoji se od interkalacije DNK i posljedične inhibicije sinteze makromolekula (33).

Od taksana koriste se paklitaksel i docetaksel, čiji je mehanizam djelovanja vezivanje za mikrotubule i sprječavanje njihovog razdvajanja, što vodi u blokadu staničnog ciklusa i posljedičnu apoptozu (32). Nadalje, derivati platine vežu se na DNK i time inhibiraju replikaciju te se među njima najčešće koriste cisplatina i karboplatina (34).

Hormonska terapija primjenjuje se ciljano kod bolesnika čiji tumori izražavaju hormonske receptore (ER, PR). Primjer su selektivni modulatori estrogenskih receptora poput tamoksifena i aromatazni inhibitori poput anastrozola, letrozola i eksemestana (35). Također su razvijeni selektivni degradatori estrogenskih receptora poput fulvestranta koji se s većim afinitetom od tamoksifena veže za ER (36).

Zatim, imamo ciljanu terapiju poput trastuzumaba koji inhibira HER-2 receptor te se ordinira bolesnicima s HER2 pozitivnim rakom dojke (37). U takvu skupinu lijekova spada i pertuzumab koji djeluje kao inhibitor dimerizacije HER-2 receptora te na taj način blokira prijenos signala za proliferaciju i stanični rast, kao i konjugat trastuzumab-emptasin koji kombinira citostatik s monoklonskim protutijelom (38, 39). Među ciljanu terapiju pripadaju i CDK 4/6 inhibitori poput abemacicliba koji je odobren u adjuvantom liječenju bolesnika s hormonski pozitivnom bolešću (40). CDK 4/6 inhibitori djeluju tako što blokiraju o ciklinu ovisne kinaze 4 i 6, čime zaustavljaju progresiju stanice u S fazu interfaze, što rezultira apoptozom i staničnom smrću (40).

1.1.9. Neoadjuvantno liječenje

Neoadjuvantno sustavno liječenje (engl. *Neoadjuvant Systemic Therapy - NST*) je predoperativno liječenje tumora kemoterapijom, imunoterapijom (pertuzumab i trastuzumab) ili endokrinom terapijom (41). Ovaj oblik liječenja primjenjuje se zbog svog učinka na kirurško liječenje, tj., zbog smanjenja stupnja proširenosti bolesti kako bi se umjesto mastektomije omogućio manje radikalni zahvat. Osim toga, neoadjuvantna terapija nudi potencijalne mogućnosti za predviđanje odgovora liječenja i relativno brzu procjenu za razvoj i odobrenje lijekova praćenjem koristi od intervencije u ranim fazama bolesti (42, 43).

Prema zadnjim (2022. godina) NCCN (engl. *National Comprehensive Cancer Network*) preporukama, bolesnici s rakom dojke određeni za neoadjuvantno liječenje moraju zadovoljavati sljedeće karakteristike: 1) veličina tumora > 5 cm (bez obzira na imunohistokemijski podtip), 2) tumori ≥ 2 cm koji su HER-2 ili trostruko negativnog podtipa, 3) tumori inflamatornog podtipa (44, 45).

Nadalje, bolesnice s HER-2 pozitivnim podtipom bolesti i pozitivnim aksilarnim limfnim čvorovima mogu također biti liječene neoadjuvantnom terapijom, bez obzira na veličinu tumora (46). Takve bolesnice mogu biti liječene kombinacijom pertuzumaba i trastuzumaba, nakon čega bivaju operirane. Prema rezultatima KATHERINE studije, kod onih bolesnica gdje je prisutna ostatna bolest, naknadno adjuvantno ordiniranje trastuzumab-emasina, dovodi do statistički značajnog produljenja preživljenja (47). Slično se pokazalo i u adjuvantnom pristupu, gdje se pri primjeni pertuzumab-trastuzumaba statistički značajno produljenje preživljenja postiglo samo u skupini bolesnika s pozitivnim pazušnim limfnim čvorovima (48).

To odražava važnost adekvatne klasifikacije limfnih čvorova, kako bi se što točnije identificirale one bolesnice koje bi imale najveću korist od drukčijih terapijskih modaliteta.

1.2. Strojno učenje

1.2.1. Uvod u strojno učenje

Strojnim učenjem smatramo matematičke algoritme/modele koji bez eksplicitnog programiranja mogu naučiti međusobne odnose između različitih varijabli u određenoj skupini podataka (49). Dok za klasični kompjutorski program moramo odrediti sustav pravila i koraka da bi dobili nekakav odgovor, u strojnom učenju mi algoritmu prikažemo ulazne varijable i specifične mete/odgovore te algoritam na opetovani način pronade sustav pravila koja povezuju ulazne varijable s ciljnom varijablom (49). Algoritmom smatramo sustav strojnog učenja prije izlaganja podacima (prije „treniranja“), dok sustav nakon izlaganja podacima (nakon „treniranja“) smatramo modelom.

Strojno učenje možemo podijeliti u tri osnovna tipa: 1) Nadgledano (nadzirano) strojno učenje (engl. *Supervised learning*), 2) Nenadgledano strojno učenje (engl. *Unsupervised learning*), 3) Podržano učenje (engl. *Reinforcement learning*) (50).

Nadgledano strojno učenje označava probleme u kojima imamo set podataka s označenim ciljnim varijablama (51). Na primjer, imamo set podataka gdje su ulazne varijable karakteristike ispitanika, a ciljna varijabla dijagnoza određene bolesti. Tradicionalno, set podataka dijelimo na takozvani „trening“ set (na kojemu se algoritam uči), zatim „validacijski“ set (na kojem optimiziramo hiperparametre i validiramo rezultate), te „test“ set (na kojemu se odvija završna procjena učinkovitosti modela). Jako je važno da su trening i test setovi potpuno odvojeni, tj. da ne dolazi do „curenja“ podataka, što može dovesti do lažno visokih performansi modela. Shodno tomu, test set mora sadržavati primjere podataka kojima model nije bio izložen.

Nadalje, studija prikazana u ovoj disertaciji također koristi Nadgledano učenje, gdje su ulazne varijable karakteristike ispitanika i tumora dojke, a ciljna varijabla dijagnoza metastatskog limfnog čvora. Podatci za algoritme Nadgledanog strojnog učenja najčešće se

spremaju u tabularnom formatu gdje svaki redak označava drugog ispitanika, a stupci označavaju prediktore i ciljnu varijablu.

Nenadgledano strojno učenje označava primjere gdje se algoritimi koriste za grupiranje podataka sa sličnim karakteristikama (engl. *Clustering*) (52). Tako u Nenadgledanom strojnom učenju nema označene ciljne varijable već su dostupne samo ulazne karakteristike na temelju kojih se podatci grupiraju. Ovakav tip strojnog učenja koristi se za segmentaciju bolesnika (53).

Podržano se učenje temelji na principu učenja preko pokušaja i pogreške (engl. *trial and error*), gdje algoritam opetovano, kroz mnoštvo pokušaja, dođe do optimalnog pristupa ili rješenja (54). Podržano učenje vuče motivaciju iz Pavlovljevog uvjetovanja, gdje se nakon željene radnje pruži nagrada, a nakon neželjene radnje kazna. Primjeri za Podržano učenje su robotika, AI sustavi u kompjutorskim igricama te preporučiteljski sustavi (53).

Nadalje, pod pojmom tradicionalnog strojnog učenja podrazumijevaju se algoritimi poput linearne i logističke regresije, stabala odluke, slučajnih šuma, metoda potpornih vektora (engl. *Support vector machine*) te stabala s podizanjem gradijenta (engl. *Gradient boosted trees*) (55). S druge strane, pojam dubokog učenja (engl. *Deep learning*) odnosi se na artificialne neuralne mreže, tj. na modele koji se temelje na neuralnim mrežama (56).

Metode dubokog učenja koriste više procesuirajućih razina (više razina/nizova artificialnih neurona) kako bi naučili prikaz podataka s više nivoa/razina apstrakcije (56). Ovakve su metode dramatično unaprijedile sustave za prepoznavanje govora, prirodno procesuiranje jezika (engl. *Natural language processing*), vizualno prepoznavanje objekata, kao i sustave za otkrivanje lijekova i genomske analize (engl. *Genomics*) (56).

1.2.2. Strojno učenje u medicini i onkologiji

Poradi sve veće količine i dostupnosti medicinskih podataka, stvorila se nova grana u medicini u kojoj se koriste tehnike strojnog učenja za postizanje bolje dijagnostike i personaliziranijeg pristupa liječenju. Stoga strojno učenje, kao proces razvoja sustava koji iz podataka uči prepoznati obrasce i napraviti točna predviđanja budućih događaja, ima velik potencijal za preobrazbu medicine i zdravstva. Modeli strojnog učenja mogli bi podržati složeno donošenje kliničkih odluka, ali također mogu i automatizirati mnoge svakodnevne zadatke koje kliničari smatraju gubitkom vremena. Prediktivni modeli strojnog učenja našli su uporabu u onkologiji, poglavito za dijagnozu bolesti, stupnjevanje bolesti, prognoziranje ishoda i optimizaciju terapije (57).

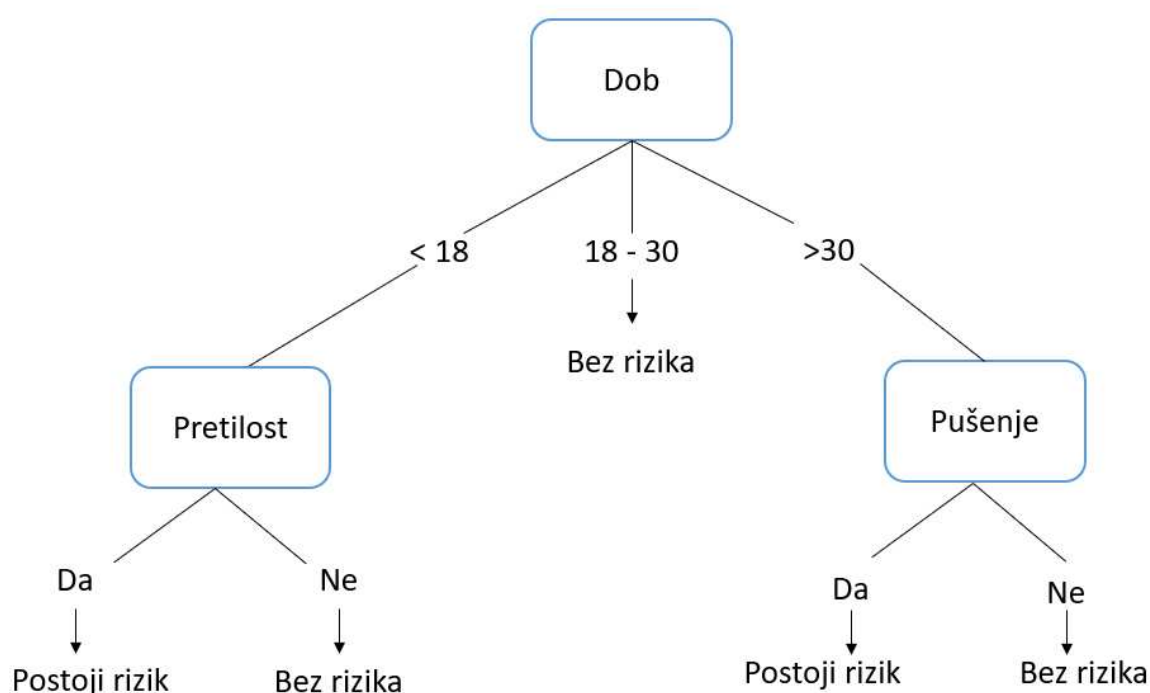
Za rak dojke razvijeno je više modela koji klasificiraju rak dojke temeljem slikovne dijagnostike (engl. *radiomics*), kao i na temelju histopatoloških slika. Na primjer, Altameem i sur. razvili su model na temelju mamografskih slika koji sjedinjuje predikcije (engl. *Ensemble method*) više različitih modela i tako postiže jako visoke razine točnosti (99.32 % točno klasificiranih slika) (58). Slično, razvijen je i model za klasifikaciju histopatoloških slika (DenseNet), koji se temelji na transfernom učenju (engl. *Transfer learning*), gdje se parametri modela koji je treniran na puno više primjera prebace na novi model koji se potom dalje još finalno optimizira (59). DenseNet model pokazuje skoro 100%-tnu točnost u histopatološkoj klasifikaciji raka dojke (99.28%) (59). Jednako tako, slični pristupi u strojnom učenju korišteni su i za klasifikaciju limfnih čvorova u raku dojke, što je i tema ove disertacije.

Nadalje, po pitanju optimizacije i personalizacije terapije razvijeni su modeli poput sustava CURATE.AI koji automatizirano određuje optimalnu dozu lijeka za specifičnu bolest i bolesnika (60).

U sljedećem dijelu slijedi detaljniji opis modela strojnog učenja korištenih u ovom istraživanju s pripadajućim primjerima.

1.2.3. Stablo odluke

Stabla odluke (engl. *decision trees*) predstavljaju jedan od algoritama koji se mogu jednostavno objasniti i razumjeti (61). Stabla odluke koriste „korak po korak“ sustav gdje se svaki korak temelji na odluci ako/osim (engl. „*if/else*“), što zapravo predstavlja i način na koji ljudi donose odluke. Jednostavan primjer algoritma stabla odluke ovdje se prikazuje na primjeru hipotetskog određivanja rizika za kardiovaskularnu bolest na Slici 2.



Slika 2. Stablo odluke za klasifikaciju rizika kardiovaskularne bolesti.

Kao što vidimo na slici jedan stablo odluke zapravo predstavlja izvrnuto stablo, gdje vrh predstavlja „korijen“ stabla (engl. *root node*) tj. prvu varijablu na temelju koje se stablo dalje dijeli, što je i varijabla s najvećom diskriminacijskom moći. Iz korijena stablo se dalje grana prema određenim karakteristikama do sljedećeg čvora. Sljedeći čvor predstavlja novu prediktivnu varijablu, na temelju čijih karakteristika se stablo također dalje grana. Na kraju, kad

imamo čvor iz kojeg se stablo ne grana u nove čvorove već donosi finalnu odluku i to nazivamo „listom“ stabla (engl. *leaf node*).

Cilj algoritma stabla odluke je postići podjelu podataka na temelju ulaznih varijabli/karakteristika, tako da završni čvorovi (listovi) sadrže što homogeniju skupinu primjera. Na našem primjeru predikcije kardiovaskularnog rizika, cilj je pronaći karakteristike i granice unutar istih, pomoću kojih možemo podijeliti podatke na takav način da nakon podjele dob > 30 i pušenje= Da, imamo skupinu bolesnika gdje svi imaju visok kardiovaskularni rizik (homogena skupina). Ponavljanjem procesa kroz moguće kombinacije prediktora i različitih podjela, stablo odluke gradira prediktore po važnosti i pronalazi one podjele koje najviše smanjuju entropiju, tj. dovode do što homogenijih finalnih skupina. Teoretski, stabla odluke mogu biti toliko duboka da točno nauče prikaz ulaznih podataka i time dobiju završne skupine s malim brojem primjera (61). Takav oblik stabla, dok savršeno zna ulazne podatke, jako slabo generalizira na novim primjerima jer je previše naučio detalje „trening“ seta podataka koji ne moraju biti prisutni u općoj populaciji. To je primjer „prenaučivosti“ modela (engl. *overfitting*), kada modeli jako dobro nauče odnose ulaznih varijabli, ali slabo generaliziraju na novim primjerima (61).

1.2.4. K-najbliži susjedi („*k-nearest neighbors*“)

Algoritam k-najbližih susjeda, poznat i kao KNN ili k-NN, neparametarski je klasifikator s nadziranom učenjem koji koristi izračun udaljenosti između točaka za izradu finalnog rezultata (62). Iako se može koristiti za probleme regresije ili klasifikacije, obično se koristi kao algoritam klasifikacije, polazeći od pretpostavke da se slične točke mogu pronaći jedna blizu druge (62).

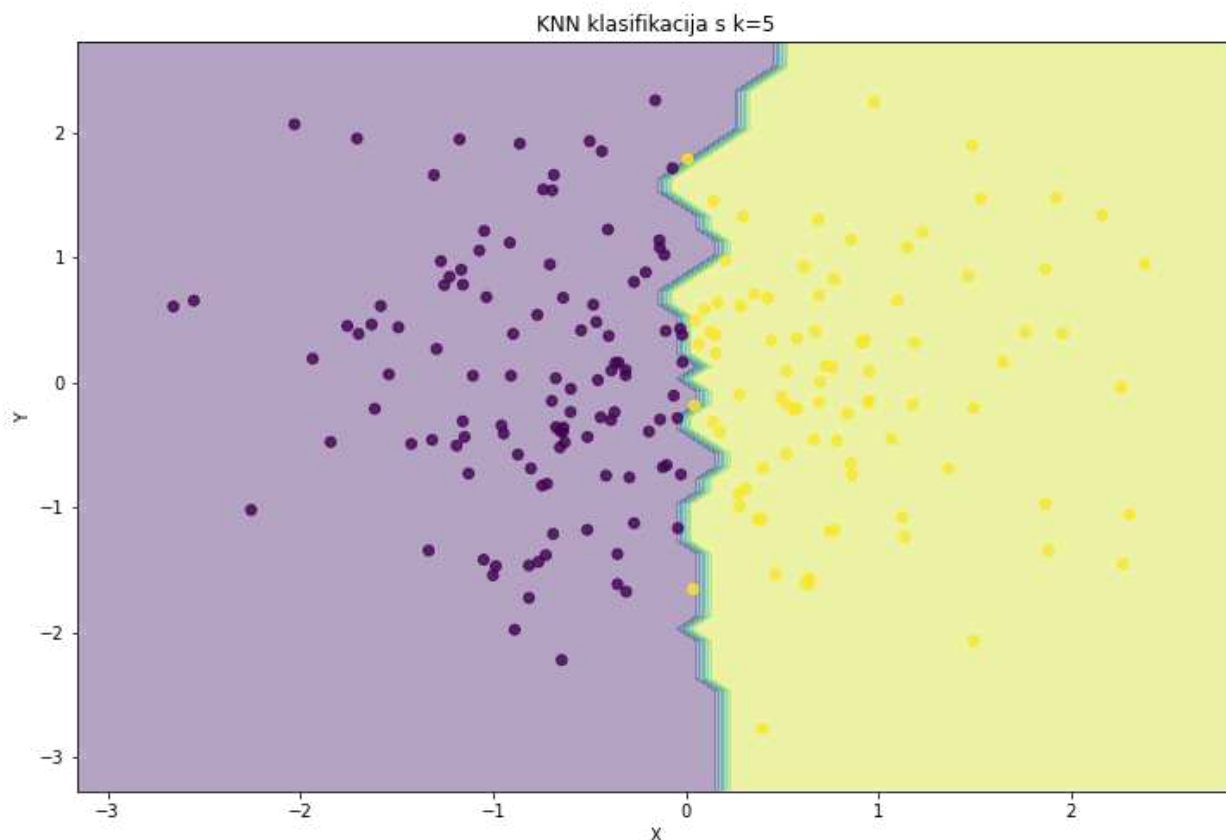
Za probleme klasifikacije dodjeljuje se oznaka klase na temelju većine glasova, tj. koristi se oznaka koja je najčešće zastupljena oko dane podatkovne točke (Slika 3.). Time,

algoritam k-najbližeg susjeda nema proces „učenja“ podataka, već samo svrstava novi podatak među podatke iz trening seta i procjenjuje „k“ najbližih primjera kako bi donio odluku kojoj klasi novi primjer pripada (62).

Prilikom računanja udaljenosti između pojedinih točaka najčešće se koriste Euklideanska udaljenost i Manhattan udaljenost (63). Euklideanska udaljenost zasniva se na sljedećoj formuli:

$$d_{A,B} = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}$$

Manhatanska udaljenost temelji se na istom principu, ali se umjesto kvadrata razlika uzima njihova apsolutna vrijednost (63).

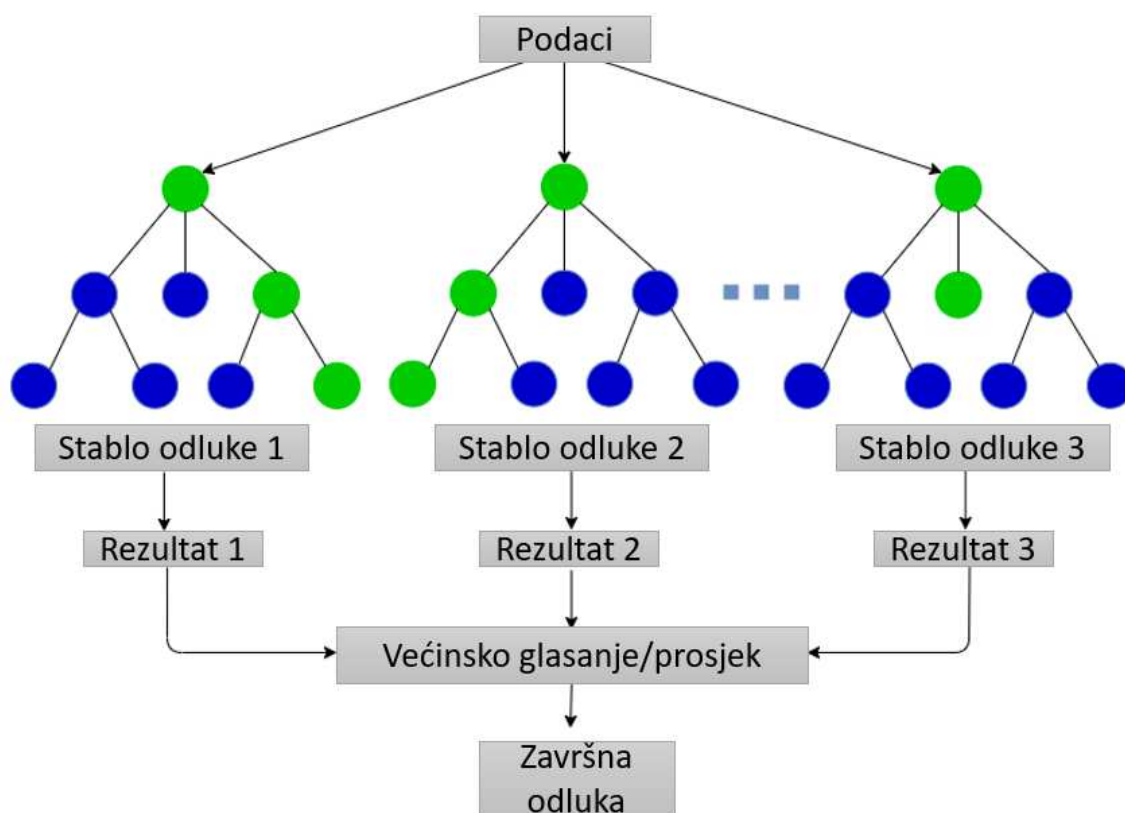


Slika 3. Primjer rezultata algoritma k-najbliži susjed s k=5

Preko računanja navedenih udaljenosti, algoritam svrstava novu točku među prethodne te prema tome kojoj klasi susjedi većinski pripadaju, klasificira i novu točku.

1.2.5. Slučajne šume

Slučajne šume, kao što se i iz imena može vidjeti, temelje se na skupini algoritama stabala odluke (Slika 4.) (64). Glavna ideja slučajnih šuma je da se uzme više stabala odluke te se svako stablo trenira na nasumično odabranom podsetu inicijalnih podataka. Time se dobiju različita stabla, među kojima svako stablo na različit način nauči svojstva inicijalnih podataka. Naposljetku, grupiranjem odluka pojedinih stabala, točnije jednostavnim glasanjem kod klasifikacijskih problema, dobiva se završna odluka (Slika 4.). Shodno tomu, kod regresijskih problema završna odluka dobiva se na temelju aritmetičke sredine pojedinih odluka (Slika 4.).



Slika 4. Primjer procesa odluke u Slučajnoj šumi

Zbog toga što slučajne šume sjedinjuju rezultate dobivene od više različitih stabala odluke (koja su trenirana na različitim uzorcima ulaznih podataka), dobiva se algoritam koji je značajno otporniji na prenaučenosť od pojedinog stabla odluke (64). Istraživanja performansi različitih modela na tabličnim podacima pokazala su nadmoć Slučajnih šuma i stabala s podizanjem gradijenta nad artifičijalnim neuronskim mrežama (65, 66).

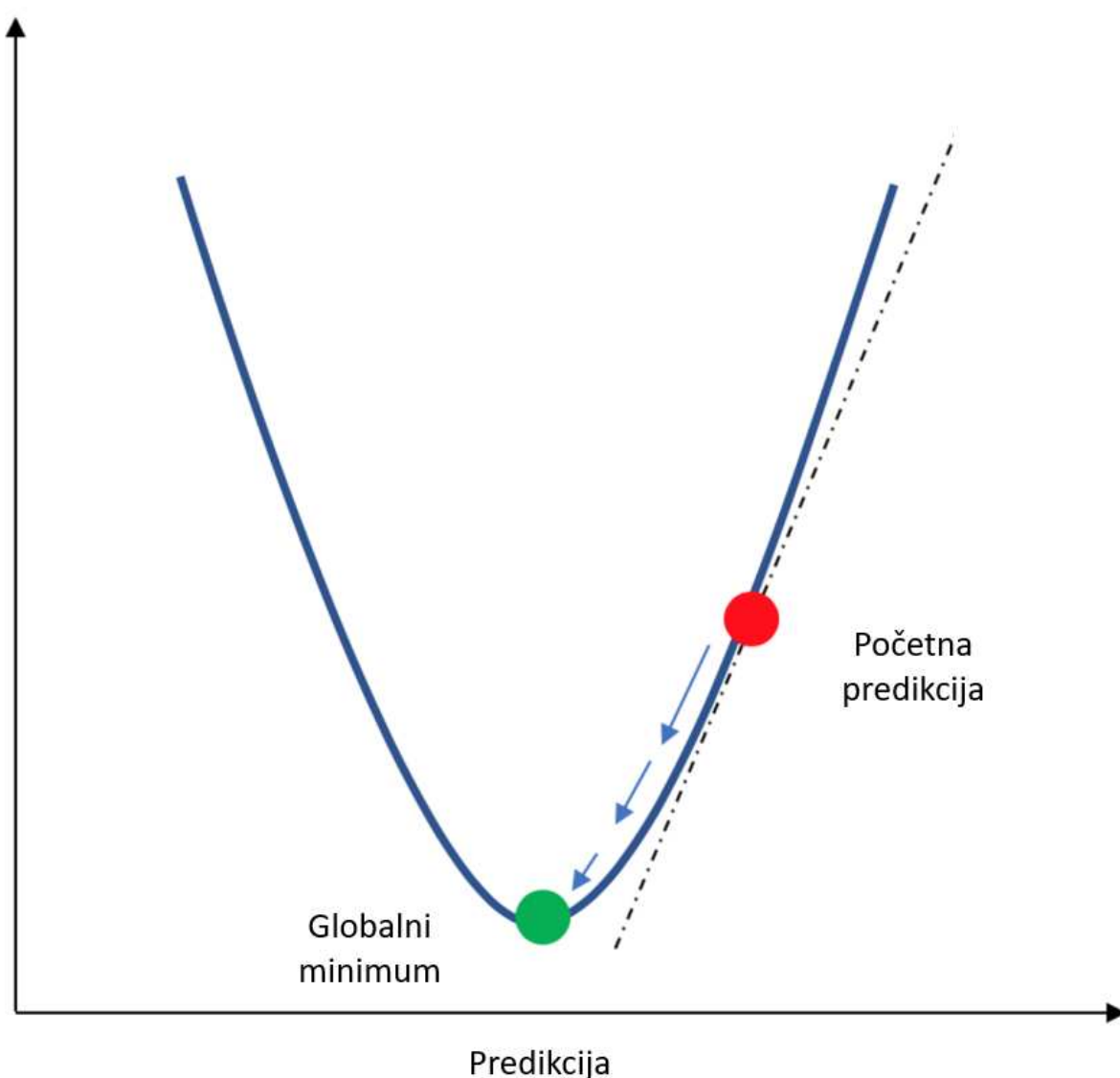
1.2.6. Stabla s podizanjem gradijenta („Gradient Boosted Trees“)

Stabla s podizanjem gradijenta predstavljaju drugu skupnu metodu (uz Slučajne šume) koja se temelji na agregaciji rezultata više pojedinih stabla (64). Ideja podizanja gradijenta ne odnosi se samo na stabla odluke, već se može primijeniti na bilo koju vrstu algoritma s uvjetom da se algoritam u pitanju smatra „slabim učenicom“ (relativno mala sposobnost prediktivne moći). Prema tome, podizanje gradijenta se može primijeniti i na potkapacitiranim linearnim regresijama, logističkim regresijama, algoritmima potpornih vektora itd.

Podizanje gradijenta zasniva se na tomu da se završna prediktivna funkcija $F(x)$ dobije iz agregacije rezultata pojedinih slabih algoritama $f_i(x)$ (67). Nadalje kod podizanja gradijenta, potkapacitirani algoritmi poredani su u nizu te svaki sljedeći algoritam uči na pogreškama prethodnih. Prvi korak u primjeni algoritama podizanja gradijenta je odabiranje funkcije gubitka s preduvjetom da je ista diferencijabilna (67). Funkcija mora biti diferencijabilna kako bi primjenom derivacija (gradijenta) uspjeli pronaći smjer od minimuma funkcije. Što je funkcija gubitka veća, performanse modela su lošije, i obrnuto, što je funkcija gubitka manja, bolja je prediktivna učinkovitost modela.

Stoga, cilj je pronaći parametre modela s kojima se dobiva globalni minimum funkcije gubitka. Takva verzija modela je optimalna za specifične podatke na kojima je model treniran i za pripadajuću funkciju gubitka.

Shodno tome, osnovna premisa algoritama podizanja gradijenta je da svaki sljedeći algoritam (u našem slučaju svako sljedeće potkapacitirano stablo odluke) napravi jedan korak prema minimumu funkcije gubitka (Slika 5.), što je i osnova algoritma spuštanja gradijenta (engl. *Gradient descent*) koji je ključan za artifičijalne neuronske mreže (68) (Slika 5.).



Slika 5. Određivanje minimuma funkcije gubitka

Primjeri stabala s podizanjem gradijenta su XGBoost (engl. *eXtreme Gradient Boosting*), LightGBM (engl. *Light Gradient-Boosting Machine*) i CatBoost (engl. *Categorical Boosting*). Uz Slučajne šume, i stabla s podizanjem gradijenta su algoritmi koji postižu najbolje rezultate na tabularnim podacima, kao što je pokazano u istraživanju od Shwarz-Ziva i suradnika (65).

1.2.7. TabNet (artificijalna neuronska mreža za tabularne podatke)

TabNet je artificijalna neuronska mreža specifične arhitekture skrojene za primjenu na tabličnim podacima (69). Artificijalne neuronske mreže (engl. Artificial Neural Network, ANN) vrsta su algoritma strojnog učenja inspiriranog strukturom i funkcijom bioloških neurona u mozgu (70). ANN sastoje se od više slojeva međusobno povezanih čvorova ili "neurona", od kojih svaki izvodi jednostavnu matematičku operaciju na svojim ulazima i šalje vrijednost sljedećem sloju (70). Ulazni sloj ANN-a prima podatke za analizu, a izlazni sloj proizvodi konačni rezultat analize. Između ulaznog i izlaznog sloja nalazi se jedan ili više skrivenih slojeva koji izvode međuzračune kako bi izdvojili i transformirali značajke ulaznih podataka (70).

Tijekom treninga, težine veza između neurona prilagođavaju se na temelju pogreške između predviđenog izlaza i stvarnog izlaza. Ovaj proces, poznat kao povratno širenje, koristi gradijentno spuštanje za ažuriranje težina na način koji minimizira pogrešku (70).

Unatoč svojim uspjesima, ANN imaju neka ograničenja, kao što je potreba za velikom količinom označenih podataka za obuku, potencijal za prenaučenosť i nedostatak interpretabilnosti njihovog internog rada (56). Unatoč tome, ANN su i dalje moćan alat za rješavanje složenih problema u različitim domenama.

Specifičnost TabNeta je u tome što koristi mehanizam „pozornosti“ (engl. *Attention*), čime se na svakom koraku odabiru najvažniji prediktori te se omogućava povećana

intepretabilnost (jer imamo uvid koji su prediktori najviše utjecali na završnu odluku modela) (69).

1.2.8. Validacija modela strojnog učenja

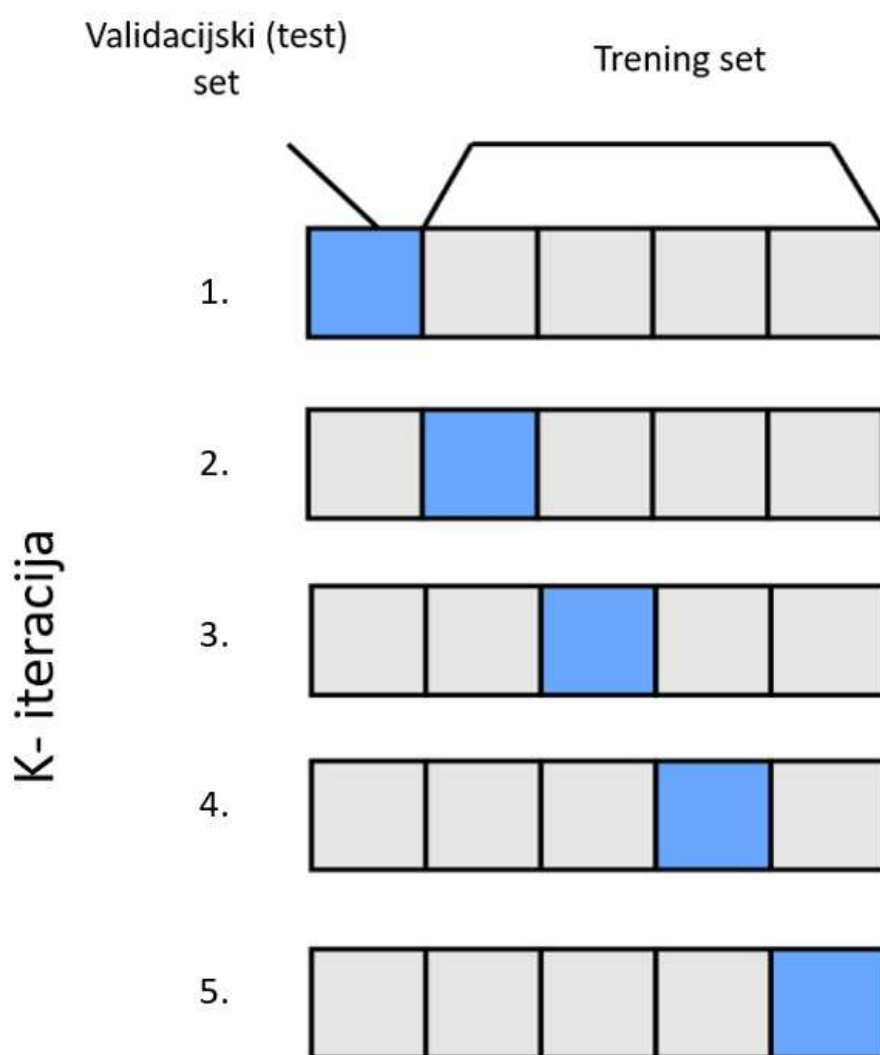
Nakon što model strojnog učenja nauči reprezentaciju određenog skupa podataka, cilj je napraviti ispravnu validaciju modela kako bi mogli procijeniti performanse tijekom primjene na stvarnom problemu (71). Na medicinskom primjeru, želimo procijeniti kakve će biti performanse modela kad se bude koristio u kliničkoj praksi (71).

Stoga, želimo procijeniti nesigurnost u predikciji ili dobiti mjere disperzije poput 95% intervala pouzdanosti kako bi znali u kojem rasponu se nalazi prava učinkovitost modela. Korištenje nasumične podjele podataka na samo jedan „trening“ i „test“ set vodi prema potencijalnoj prenaučivosti modela i preoptimističnim rezultatima (71, 72).

Metodama poput križne validacije i „bootstrappinga“ možemo na adekvatan način procijeniti realnu učinkovitost našeg modela, kao i odrediti mjere disperzije (72). Nadalje, klasične metode računanja intervala pouzdanosti, poput *aritmetička sredina* $\pm z * standardna\ pogreška$, također mogu dovesti do pretjerane pouzdanosti u rezultat (uski interval pouzdanosti), jer je broj uzoraka u nazivniku formule za standardnu pogrešku. Stoga, veći broj uzoraka dovodi do užih intervala pouzdanosti, što se upravo i događa na primjerima iz podaktovne znanosti koja je obično karakterizirana velikim brojem podataka. Također, klasična formula za interval pouzdanosti pretpostavlja normalnu raspodjelu podataka, što često nije slučaj. Shodno tomu, ne preporuča se koristiti klasičnu metodu izračuna intervala pouzdanosti, već se preporuča korištenje križne validacije i „bootstrappinga“ (73).

Križna validacija odnosi se na metodu gdje se od cjelokupnog seta podataka izradi K različitih podjela (najčešće 10 ili 5) u „trening“ i „test“ skupine (72). Na primjeru peterogrupne križne validacije (engl. *5-fold cross validation*), napravimo 5 različitih kombinacija

cjelokupnog seta podataka, gdje u svakoj kombinaciji različitih 20% od ukupnog seta sačinjava „test“ set. Ponavljanjem križne validacije N puta, možemo izračunati srednje vrijednosti ciljanih mjera kao i mjere disprezije te time dobiti vrijedni uvid u istinsku učinkovitost modela (Slika 6.).



Slika 6. Peterogrupna križna validacija

„*Bootstrapping*“ je generalna tehnika za računanje statističkih mjera koja se može koristiti bez obzira na tip distribucije podataka. Zasniva se na principu ponavljanja uzorkovanja iz originalnog seta podataka. Naime, iz originalnog seta podataka nasumično se izvuče jedan primjer koji se potom vrati (kako bi mogao biti ponovno izvučen) te se iteracija ponavlja. Time dobivamo nove skupove podataka koji su slični (ali ne i isti) originalom skupu. Uz pomoć trening ili test setova dobivenih bootstrapom, ispitamo performanse modela i računamo srednje vrijednosti te mjere disperzije.

Naposljetku, dok validaciju na lokalnim podacima nazivamo internom, postoji i primjer eksterne validacije, kad se model strojnog učenja validira na primjerima drugog medicinskog centra (čiji podaci nisu korišteni u dosadašnjim trening i test setovima) (71).

1.2.9. Objašnjivo strojno učenje

Objašnjivo strojno učenje (engl. *Explainable machine learning*) podrazumijeva metode kojima se pokušava pojasniti način na koji algoritam/model dolazi do konačnih odluka. Dok su algoritimi jednostavnih struktura poput stabala odluke intuitivni i može ih se intepretirati te razlučiti kako su došli do rezultata, kompleksnije strukture poput umjetnih neuronskih mreža i stabala s pojačanjem gradijenta, značajno kompliciraju proces i smatraju se crnim kutijama (engl. *black box*). S ciljem da se i procesi unutar takvih algoritama razjasne, razvijene su metode model-agnostičkog objašnjenja, točnije, metode koje ne ovise o samoj strukturi modela prilikom itepretacije procesa donošenja odluka.

Takva metoda je i SHAP (engl. *Shapley Additive Explanations*) koja se zasniva na primjeni Shapley vrijednosti (73). Shapley vrijednosti vuku korijen iz teorije igara (engl. *Game theory*), te im je cilj izmjeriti doprinos svakog pojedinog igrača u igri. Lloyd Shapley je 1953. godine predložio sustav koji proizlazi iz konteksta u kojem "n" igrača zajednički sudjeluju dobivajući nagradu "p" koja je namijenjena pravednoj raspodjeli na svakog od "n" igrača prema

individualnom doprinosu, takav doprinos je upravo Shapleyeva vrijednost (74). Ova teorija tek je nedavno primijenjena za interpretaciju algoritama zbog velike komputacijske cijene koja je potrebna za izračun svih mogućih permutacija i dobivanja srednje vrijednosti individualnog doprinosa.

Na primjeru gdje imamo četiri prediktora, da bi se izračunao doprinos jednog od prediktora pri dobivanju finalnog rezultata, izračunava se razlika između rezultata koji se generira kada je član prisutan u odnosu na dobiveni rezultat koji se generira kada je taj član odsutan. Time se dobiva marginalni doprinos tog prediktora za specifičnu koaliciju.

Proces se ponavlja za svaku podskupinu (ili koaliciju) koja se može generirati u kojoj je prisutan prediktor za kojeg tražimo izračun doprinosa. Srednja vrijednost dobivenih razlika (srednji granični doprinos) je Shapleyeva vrijednost.

1.3. Strojno učenje u klasifikaciji limfnih čvorova u raku dojke

Dosad korištene tehnike strojnog učenja s ciljem klasifikacije limfnih čvorova kod bolesnika s rakom dojke možemo podijeliti u tri glavne skupine: 1) modeli koji koriste radiološke slike/parametere (engl. *radiomics*), 2) modeli koje koriste kliničko-patološke podatke, 3) kombinirani pristup („*radiomics*“ + kliničko-patološki podaci).

1.3.1. Modeli koji koriste radiološke slike/parametre (Radiomika)

Radiomika je pojam koji označava korištenje modela strojnog učenja treniranih na podacima dobivenim iz analize radioloških slika. Za treniranje modela mogu se koristiti parametri koji su ručno ili automatizirano ekstrahirani iz radioloških slika, ili se mogu koristiti radiološke slike same po sebi (kod modela dubokog učenja). Prethodni pristupi zasnivali su se na manualnoj segmentaciji željenih dijelova originalne radiološke slike i treniranju tradicionalnih modela poput potpornih vektora i logističke regresije. Metode dubokog učenja

ne zahtijevaju prethodnu segmentaciju ili ekstrakciju značajki, već sam algoritam izvlači važne značajke iz ulaznih podataka (poput slike) (56).

Shodno tomu, algoritmi dubokog učenja proizašli su kao najbolja opcija za izradu prediktivnih modela na radiološkim slikama.

Metodama poput konvolucijskih neuronskih mreža (engl. *Convolutional neural network*- *CNN* prenesenog učenja (engl. *Transfer learning*) postižu se zavidni rezultati u detekciji tumora i metastaza. Preneseno učenje odnosi se na metodu gdje se parametri jednog modela prenose na novi model. Često se primjenjuje kod slikovnih modela koristeći pred-trenirane arhitekture poput ResNet (engl. *Residual Network*) koja je istrenirana na velikoj količini slikovnih primjera (75). Time se osnovne značajke slika (poput rubova, linija itd..) ne moraju nanovo učiti već bivaju prenesene iz već razvijenog modela.

Na primjer, Wang i sur. su razvili model prenesenog učenja istreniran na slikama magnetne rezonancije (T1, T2 i difuzijskim sekvencama) koji postiže AUC od 0.997 prilikom evaluacije na test setu (76).

Razvijeni su i modeli s visokom prediktivnom moći koristeći ultrazvučne slike raka dojke i limfnih čvorova, poput konvolucijske neuronske mreže razvijene od strane Ozakia i sur. (AUC= 0.966) (77).

Bez obzira na jako dobre rezultate, obje navedene studije validirale su modele na samo jednom test setu, te stoga postoji velika šansa za prenaučenosť i pretjerano pozitivne rezultate.

1.3.2. Modeli na kliničko-patološkim podacima

U ovom pristupu koriste se modeli strojnog učenje istrenirani samo na kliničko-patološkim podacima. Studije poput one od Dihge i sur. te Takade i sur. pokazale su vijabilnost treniranja modela na kliničko-patološkim podacima u svrhu klasifikacije limfnih čvorova kod bolesnika s rakom dojke (AUC 0.74 [0.72-0.76] i 0.77 [0.69-0.86] (78, 79). Obično se koriste

veličina tumora, ER status, PR status, HER-2 status, Ki-67, histološki tip, dob bolesnika i status limfnih čvorova (ciljna varijabla) kao ulazni podatci. Na temelju tih karakteristika model nauči odnose između prediktora i ciljne varijable. Ovakav tip podataka se na strukturiran način unosi u tablicu tako da se svrstava u tabularne podatke. Već spomenuta istraživanja pokazala su nadmoć algoritama poput slučajnih šuma i stabala pojačanih gradijentom na tabularnim podacima, tako da i na kliničko-patološkim podacima predstavljaju najbolji odabir (65).

Zanimljivo u istraživanju od Dihge i sur., razvijena je artificijalna neuralna mreža istrenirana na navedenim kliničko-patološkim podacima, među kojima je korišten i status limfovaskularne invazije, što često i nije dio standardne obrade bolesnika s rakom dojke (78). Obzirom da je u našem istraživanju cilj napraviti model koji koristi podatke dobivene standardnom obradom bolesnika s rakom dojke, model neće koristiti limfovaskularnu invaziju kao ulazni parametar.

1.3.3. Kombinirani pristup

Treći pristup obuhvaća kombinaciju radioloških i kliničkopatoloških karakteristika za treniranje modela strojnog učenja. U istraživanju od Zhenga i sur. razvijen je model koristeći ultrazvučne slike i kliničkopatološke karakteristike ispitanika (80). Podatci su skupljeni od ukupno 1342 bolesnika te su na metodama dubokog učenja razvijena tri modela (kliničkopatološki, ultrazvučni i kombinirani). Model koji je kombinirao oba modaliteta podataka postigao je značajno bolje performanse od obaju pojedinih modela (AUC 0.902 za kombinirani vs 0.735 za kliničkopatološki i 0.796 za ultrazvučni) (80).

Zatim imamo studiju Chena i sur., koji su primijenili logističku regresiju na kombinaciji MRI i transkriptomskih podataka (81). Tako istrenirana logistička regresija postiže AUC od 0.82 u klasifikaciji metastatskih limfnih čvorova kod raka dojke (81).

Prednost kombiniranih metoda je to što korištenje različitih modaliteta podataka smanjuje potencijal za prenaučenosť modela i time povećava sposobnost generalizacije. Tako

dobivamo robustniji model koji postiže bolje rezultate prilikom primjene u svakodnevnom radu. Osim toga, kombiniranjem različitih modaliteta podataka dobiva se i bolja/točnija slika o bolesti i bolesniku te se otkrivaju potencijalno novi odnosi između prediktora i ciljne varijable.

2. CILJEVI I HIPOTEZE ISTRAŽIVANJA

Glavni ciljevi istraživanja su:

1. Optimizirati i validirati objašnjive modele strojnog učenja za predikciju statusa limfnih čvorova kod bolesnika s rakom dojke koji ispunjavaju kriterije za NST.
2. Odrediti dijagnostičke performanse modela strojnog učenja za predikciju statusa limfnih čvorova kod bolesnika s rakom dojke koji ispunjavaju kriterije za NST.
3. Odrediti dijagnostičke performanse modela strojnog učenja za predikciju statusa limfnih čvorova za ukupnu populaciju u studiji.

Sporedni ciljevi istraživanja su:

1. Analiza prediktorskih varijabli primjenom metode SHAP s ciljem utvrđivanja važnosti prediktora.
2. Ispitati objašnjivost modela primjenom metode SHAP.

Hipoteze istraživanja su:

1. Kompleksniji modeli (poput Slučajnih šuma i XGBoosta) postići će bolje performanse od logističke regresije i univarijantne logističke regresije u klasifikaciji metastatskih limfnih čvorova u bolesnica oboljelih od raka dojke.
2. Veličina tumora i indeks proliferacije stanica Ki67 bit će najvažniji prediktori prema Shapley vrijednostima.

3. ISPITANICI I POSTUPCI

3.1. Ispitanici

Istraživanje je odobrilo Etičko povjerenstvo KBC-a Split (2181-147). U istraživanje su uključeni anonimizirani podatci dobiveni iz prospektivno održavane baze podataka, od ukupno 13580 bolesnika liječenih od raka dojke. Uzorak čine ispitanici liječeni od raka dojke u 25 hrvatskih bolnica u periodu od 01/2017 do 01/2022. Podatci su kroz navedeni period prikupljeni iz bolničkih informacijskih sustava (BIS), tako što su pretraživani svi bolesnici s MKB šifrom 50 (za rak dojke).

Svi patohistološki podatci su u standardnom formatu po ASCO/CAP smjernicama koje koriste sve hrvatske bolnice. Uključeni su svi ispitanici s potpunim podacima (PHD, status LČ, ER, PR, HER2, Ki67, veličina tumora, dob). Izbačen je dio ispitanika s nepotpunim podacima (nedostatak neke od ulaznih varijabli), nakon čega je preostalo 9705 ispitanika. Među tim ispitanicima njih je 1324 inicijalno primilo NST zbog čega su također izbačeni iz analize. Naposljetku, nakon primjene kriterija za NST (trostruko negativni, HER2+ te tumori >5cm veličine), identificirano je 719 ispitanika za završnu analizu, optimizaciju i validaciju prediktivnih modela. Također, izrađeni su i modeli na cjelokupnoj populaciji s potpunim podacima (n= 8381).

3.2. Postupci

3.2.1. *Eksploratorna analiza podataka i čišćenje podataka*

Inicijalna baza podataka (n= 13580) je ispravljena od pogrešno unesenih vrijednosti, izvršena je standardizacija formata različitih varijabli, te su izbačeni duplikati, kao i ispitanici s nepotpunim podacima (nakon procesuiranja n= 9705). Potom su izbačeni i ispitanici koji su inicijalno primili neoadjuvantnu terapiju, poradi utjecaja iste na smanjenje stupnja bolesti i potencijalnu negativizaciju pozitivnih limfnih čvorova (n= 8381).

Potom su identificirani ispitanici koji ispunjavaju kriterije za neoadjuvantno liječenje (prema NCCN kriterijma), čime je dobiveno 719 ispitanika za primarnu analizu (Slika 7.).

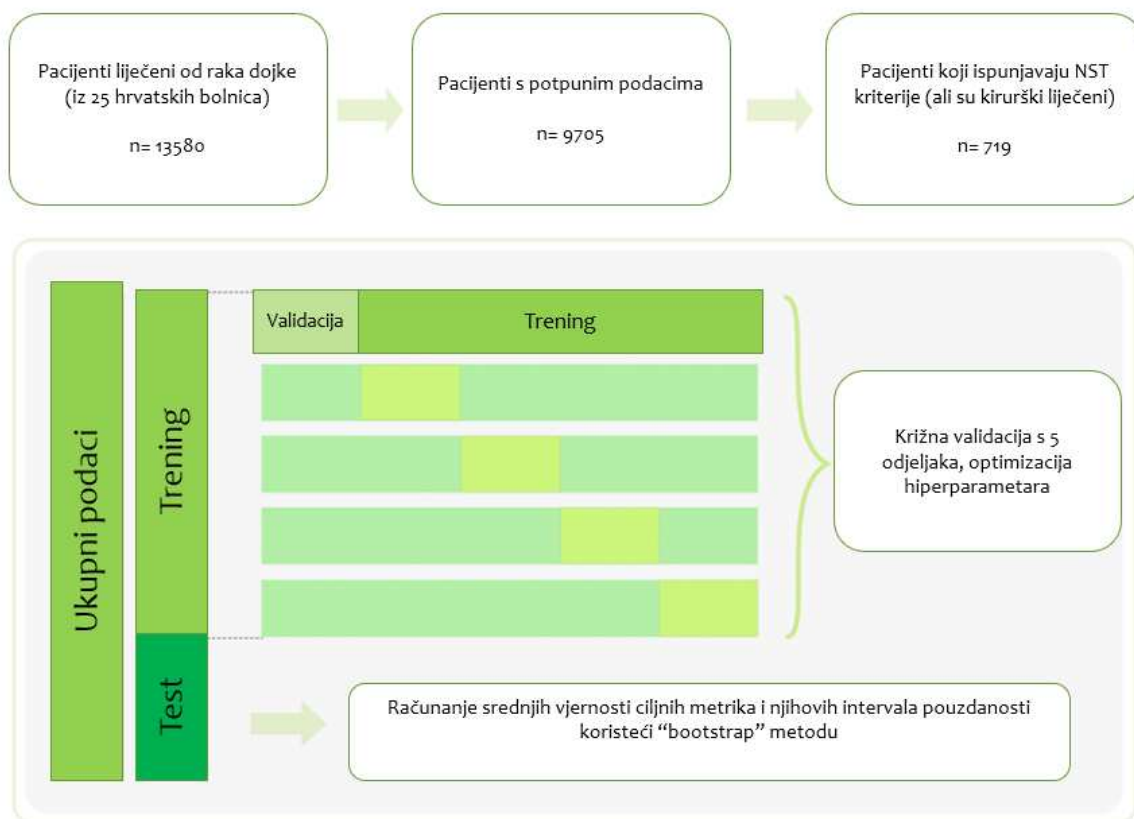
Kategorijske varijable su nadalje procesuirane metodom („dummy encoding“) gdje svaka klasa u varijabli postaje nova zasebna varijabla, te se izradi novi stupac s vrijednostima 0/1 (ovisno o pripadnosti određenoj klasi). Također, ciljna varijabla je procesuirana na sličan način, gdje 1 označava prisutnost metastatskog limfnog čvora.

Provedena je deskriptivna statistička analiza karakteristika ispitanika s pozitivnim i negativnim limfnim čvorovima. Studentov t-test je korišten za analizu numeričkih varijabli, dok je χ^2 test korišten za analizu kategorijskih varijabli. Razina statističke značajnosti je postavljena na $P < 0.05$. Izračuni su izvedeni s programskim jezikom Python (verzija 3.9.5, Python Software Foundation) korištenjem biblioteke „*scipy*“.

3.2.2. Modeliranje i optimizacija

Prije modeliranja, podatci su nasumično podijeljeni u trening (80 %) i test set (20%) za NST grupu, te trening (90 %) i test set (10 %) za ukupnu populaciju. Zbog značajno većeg ukupnog broja ispitanika ($n = 8381$), korišten je manji postotak za test set u toj analizi.

Nakon podjele podataka i provjere točnosti formata ulaznih varijabli, pristupilo se modeliranju. Primjenjeni su sljedeći modeli: 1) univarijatna logistička regresija (trenirana samo na veličini tumora), 2) logistička regresija, 3) KNN, 4) Slučajne šume, 5) XGBoost, 6) TabNet. Hodogram postupka može se vidjeti na Slici 7.



Slika 7. Hodogram postupaka

Slučajne šume i XBoost su odabrani poradi dokazanih performansi na tabularnim podacima, dok su logistička regresija i KNN odabrani kao bazični algoritam za usporedbu (66). Nadalje, TabNet je odabran za ispitivanje nadmoći algoritama utemeljenih na stablima naspram umjetnim neuronskim mrežama. Provedena je i analiza XGBoost modela na ukupnim podacima kojima nedostaju određene varijable ($n= 10540$), jer XGBoost ima ugrađenu sposobnost rješavanja nedostajućih varijabli.

Iz trening seta provedena je stratificirana peterogrupna križna validacija (stratifikacija je napravljena prema ciljnoj varijabli), nakon čega je primijenjena tehnika pretrage kombinacija hiperparametara „*Grid-search*“, čime je izvršena optimizacija za logističku regresiju, KNN, Slučajne šume i XGBoost. Primjenom ove tehnike testirane su multiple kombinacije hiperparametara, te su identificirani hiperparametri koji dovode do najboljih performansi na

validacijskim setovima podataka. Također XGBoost optimiziran je i dodatnom tehnikom iz „optuna“ biblioteke. Optuna je moćna Python biblioteka dizajnirana za pojednostavljenje procesa optimizacije hiperparametara (82). Koristeći učinkovit algoritam rezanja, automatski istražuje i sužava najbolje kombinacije hiperparametara za modele strojnog učenja.

Za logističku regresiju optimizirani su sljedeći hiperparametri: 1) „*solver*“ (različiti algoritmi za optimizaciju), 2) regularizacija (za kontrolu prenaučenosti L1 vs L2 regularizacija), 3) snaga regularizacije.

Zatim za KNN optimiziran je „*k*“, tj. broj najbližih susjeda koje uzimamo u obzir prilikom klasifikacije.

Potom, za slučajne šume optimizirani su: 1) maksimalna dubina stabala, 2) minimalni broj uzoraka potreban za razdvajanje unutarnjeg čvora, 3) minimalan broj uzoraka u završnom čvoru („*leaf node*“) i 4) broj stabala (estimatora).

Za XGBoost optimizirani su: 1) maksimalna dubina stabala, 2) minimalni uzorak u jednom čvoru prije razdvajanja, 3) minimalan broj uzoraka u završnom čvoru („*leaf node*“), 4) minimalna težina parametra potrebna za stvaranje novog čvora („*min_child_weight*“), 5) gamma (pseudo-regularizacijski parametar), 6) alpha (snaga L1-regularizacije na razini lista), 7) „*subsample*“ (nasumično odabran udio podataka za treniranje), 8) nasumični udio prediktora prilikom izrade stabala („*colsample_by_tree*“), 9) nasumični udio prediktora na svakoj razini stabla („*colsample_by_level*“) i 10) broj stabala.

Modeliranje je provedeno korištenjem Python Programming Language (verzija 3.9.5, Python Software Foundation, Wilmington, DE, USA) korištenjem biblioteka “numpy”, “pandas”, “scikit-learn” i “xgboost”, te s R programskim jezikom (R Core Team, 2022, Beč, Austrija) koristeći pakete "tidymodels", "ranger", "xgboost", "pROC".

3.2.3. Performanse modela

Performanse modela ispitane su na test setu uporabom „*bootstrap*“ tehnike, kojom je na 2000 test setova evaluiran model te su izračunate srednje vrijednosti ciljnih metrika, kao i njihovi 95 % intervali pouzdanosti.

Ucrtane su ROC-krivulje te je izračunata površina ispod krivulje („*area under the curve* – *AUC*“) koristeći sljedeću formulu: $AUC = \int_0^1 TPR \, dFPR$, gdje je $TPR = \frac{TP}{TP+FN} = \text{osjetljivost}$, a $FPR = \frac{FP}{TN+FP} = 1 - \text{specifičnost}$ (TPR= *True positive rate*, FPR= *False positive rate*, TP= *True positive*, FP= *False positive*, TN= *True negative*).

Također su izračunati F1-mjera („*F1-score*“, harmonična aritmetička sredina između osjetljivosti i pozitivne prediktivne vrijednosti), preciznost (pozitivna prediktivna vrijednost), negativna prediktivna vrijednost, osjetljivost i specifičnost. Model s najvećim AUC učinkom izabran je za daljnju analizu. Optimalne odsječne točke za osjetljivost i specifičnost procijenjene su prema optimizaciji F1-mjere.

3.2.4. Analiza važnosti prediktora i objašnjivost modela

SHAP metoda korištena je za analizu važnosti različitih prediktora te za analizu objašnjivosti odabranog modela. Analiza je provedena na modelima koji su postigli najveći AUC (za NST skupinu i za ukupnu populaciju). Prikazano je koliko pojedine varijable pridonose odluci modela. Nadalje, prikazan je pozitivan ili negativan utjecaj pojedinih varijabli, kao i inter-varijablina razlika. Završno, prikazan je zbroj Shapley vrijednosti i time skupna važnost svake pojedine varijable. Izračuni su također izvedeni s programskim jezikom Python korištenjem biblioteke „*shap*“.

4. REZULTATI

4.1. Karakteristike ispitanika

Ukupno je 5845 (69,7%) bolesnika identificirano s negativnim limfnim čvorovima, dok je 2536 (30,3%) bolesnika identificirano s pozitivnim limfnim čvorovima (Tablica 2). Statistički značajne razlike između slučaja (pozitivni LČ) i kontrola (negativni LČ) uočene su u veličini tumora, PR i HER-2 receptorima, kao i u indeksu proliferacije Ki-67 (Tablica 2).

Tablica 2. Karakteristike ispitanika (ukupna populacija)

	Ispitanici s pozitivnim LČ (n= 2536)	Ispitanici s negativnim LČ (n= 5845)	Ukupno (n= 8381)	P- vrijednost
Dob	63,6 (13.5)	62,3 (12.1)	62,7 (12,6)	0,535*
Veličina tumora (cm)	2,7 (1,9)	1,7 (1,1)	2,01 (1,5)	< 0,001*
Gradus tumora				< 0.001†
1	376 (14,8 %)	1638 (28 %)	2014 (24 %)	
2	1460 (57,6 %)	3148 (53,8 %)	4608 (54,9 %)	
3	700 (27,6 %)	1059 (18,2 %)	1759 (20,9 %)	
ER	80,7 (33,7)	83,1 (32,2)	82,4 (32,7)	0,064*
PR	50,8 (39,5)	54,3 (39,4)	53,2 (39,4)	< 0,001*
HER-2				< 0,001*
0	1092 (43,1 %)	2776 (47,5 %)	3868 (46,2 %)	
1	815 (32,1 %)	1953 (33,4 %)	2768 (33 %)	
2	344 (1,3 %)	629 (10,8 %)	914 (10,9 %)	
3	285 (11,2 %)	487 (8,3 %)	831 (9,9 %)	
Ki-67	29,7 (18,7)	25,1 (18,4)	26,5 (18,7)	< 0,001*
Histološki tip				< 0,001†
NOS-invazivni	2055 (81 %)	4586 (78,5 %)	6641 (79,2 %)	
Lobularni invazivni	324 (12,8 %)	693 (11,9 %)	1017 (12,1 %)	
Karcinom s medularnim karakteristikama	24 (0,9 %)	47 (0,8 %)	71 (0,8 %)	
Drugi (rjeđi oblici)	133 (5,2 %)	519 (8,9 %)	652 (7,8 %)	
Imunofenotip				< 0,001†
Luminalni A	429 (16,9 %)	1628 (27,8 %)	2057 (24,6 %)	
Luminalni B	1517 (59,8 %)	3154 (53,9 %)	4671 (55,7 %)	
Luminalni B-her2	310 (12,3 %)	508 (8,8 %)	818 (9,8 %)	
Her2 pozitivni	120 (4,7 %)	148 (2,6 %)	268 (3,2 %)	
Trostruko negativni	160 (6,3 %)	407 (6,9 %)	567 (6,7 %)	

Podatci su prikazani kao srednja vrijednost (standardna devijacija) ili kao cijeli broj (postotak). LČ- limfni čvorovi, ER- estrogeneski receptori, PR- progesteronski receptori, HER2- humani receptor za epidermalni čimbenik rasta, Ki-67- indeks proliferacije stanica, NOS- „not otherwise specified“. * t-test za nezavisne uzorke, † hi-kvadrat test.

U skupini bolesnika koji ispunjavaju NST kriterije bilo je 426 (55%) bolesnika s metastazama u limfnim čvorovima i 350 (45%) bolesnika bez metastaza. Budući da su metastaze u limfnim čvorovima prisutne u 55% ciljne populacije, možemo vidjeti da je naša ciljna varijabla prilično uravnotežena (55% naspram 45%), što se razlikuje od ukupne populacije gdje je omjer u korist skupine bez metastaza (34 % u odnosu na 66%).

Štoviše, za NST kriterijsku skupinu, značajna razlika je pokazana u veličini tumora, PR i Ki-67, ali također i u dobi bolesnika i ER, dok nije bilo značajne razlike za HER-2 (Tablica 3).

Tablica 3. Karakteristike ispitanika (NST skupina)

	Ispitanici s pozitivnim LČ (n= 392)	Ispitanici s negativnim LČ (n= 327)	Ukupno (n= 719)	P- vrijednost
Dob	66,9 (14,9)	64,2 (14,1)	65,7 (14,6)	0,016*
Veličina tumora	5,7 (3,02)	3,9 (2,5)	4,9 (2,9)	< 0,001*
Gradus tumora				0,027†
1	14 (3,6 %)	8 (2,4 %)	22 (3 %)	
2	133 (33,9 %)	97 (29,6 %)	230 (32 %)	
3	245 (62,5 %)	222 (68 %)	467 (65 %)	
ER	40,01 (45,9)	13,7 (32,6)	28,1 (42,5)	< 0,001*
PR	21,8 (35,04)	7,51 (22,5)	15,2 (30,6)	< 0,001*
HER-2				0,060*
0	180 (46 %)	173 (53 %)	353 (49,1 %)	
1	79 (20,1 %)	55 (16,8 %)	134 (18,7 %)	
2	35 (8,9 %)	42 (12,8 %)	77 (10,7 %)	
3	98 (25 %)	57 (17,4 %)	155 (21,5 %)	
Ki-67	43,4 (23,3)	50,6 (25,7)	46,7 (24,7)	< 0,001*
Histološki tip				< 0,001†
NOS-invazivni	282 (71,9 %)	252 (77,2 %)	534 (74,3 %)	
Lobularni invazivni	63 (16,1 %)	20 (6,1 %)	83 (11,5 %)	
Karcinom s medularnim karakteristikama	15 (3,8 %)	12 (3,6 %)	27 (3,7 %)	
Drugi (rjeđi oblici)	32 (8,2 %)	43 (13,1 %)	75 (10,4 %)	
Imunofenotip				< 0,001†
Luminalni A	26 (6,6 %)	7 (2,3 %)	33 (4,6 %)	
Luminalni B	111 (28,4 %)	35 (10,7 %)	146 (20,3 %)	
Luminalni B-her2	35 (8,9 %)	8 (2,4 %)	43 (5,9 %)	
Her-2 pozitivni	83 (21,2 %)	67 (20,4 %)	150 (20,8 %)	
Trostruko negativni	137 (34,9 %)	210 (64,2 %)	347 (48,4 %)	

Podatci su prikazani kao srednja vrijednost (standardna devijacija) ili kao cijeli broj (postotak). LČ- limfni čvorovi, ER- estrogneski receptori, PR- progesteronski receptori, HER2- humani receptor za epidermalni čimbenik rasta, Ki-67- indeks proliferacije stanica, NOS- „*not otherwise specified*“. * t-test za nezavisne uzorke, † hi-kvadrat test.

4.2. Performanse modela strojnog učenja

4.2.1. Performanse za NST- skupinu

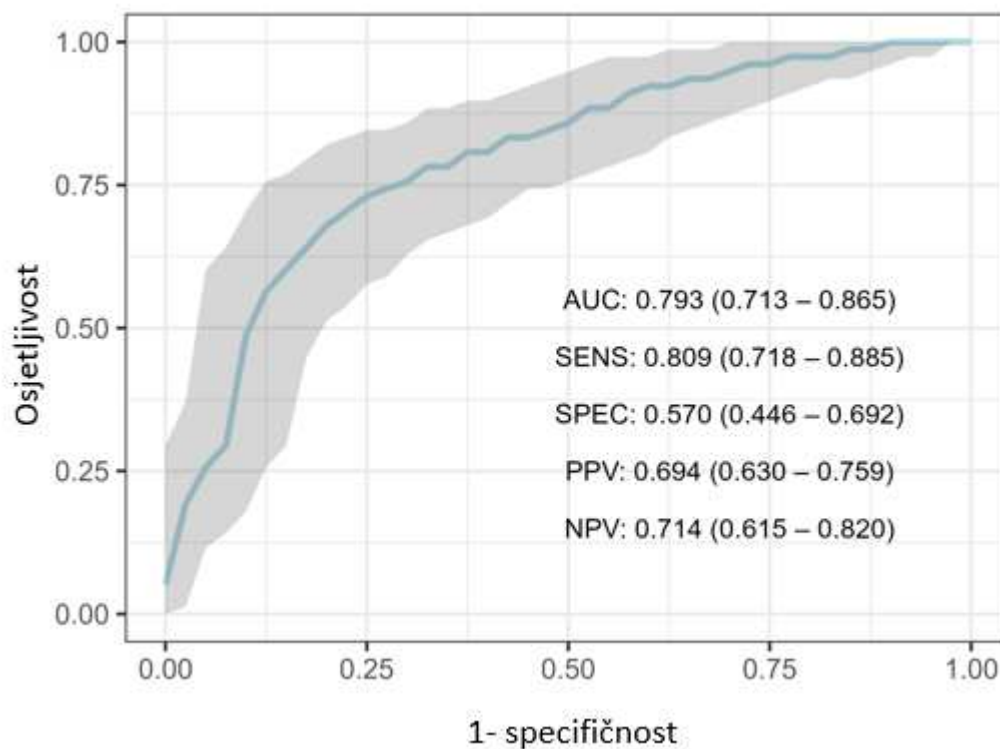
Ukupno je 6 modela istrenirano na podacima iz NST-skupine te su se slučajne šume pokazale kao model s najboljim performansama. Prije optimizacije model slučajnih šuma postigao je AUC od 0,76 (na test setu), dok je nakon optimizacije hiperparametara učinkovitost porasla na 0,793 (srednji AUC, 95% IP: 0,713-0,865) (Slika 8., Tablica 4.).

Za bazični prag odluke od 0,5, osjetljivost je 0,809 (95% IP: 0,718-0,885), specifičnost 0,570 (95% IP: 0,446-0,692), negativna prediktivna vrijednost 0,714 (95% IP: 0,615-0,820) te preciznost (pozitivna prediktivna vrijednost) 0,694 (95% CI: 0,630-0,759) (Slika 8.).

XGBoost, kao drugi model utemeljen na stablima odluke, postigao je drugi najbolji rezultat s AUC 0,783 (95% IP: 0,703-0,858) (Tablica 4.).

Univarijatna logistička regresija (trenirana samo na veličini tumora) postigla je srednji AUC od 0,688 (95% IP: 0,626-0,745), dok je logistička regresija postigla srednji AUC od 0,763 (95% IP: 0,683-0,838) (Tablica 4.).

KNN algoritam postigao je srednji AUC od 0,701 (95% IP: 0,622-0,782) (Tablica 4.). Naposljetku, umjetna neuronska mreža TabNet postigla je srednji AUC od 0,740 (95% IP: 0,652-0,812) (Tablica 4.).



AUC- površina ispod krivulje, SENS- osjetljivost, SPEC- specifičnost, PPV- pozitivna prediktivna vrijednost, NPV- negativna prediktivna vrijednost.

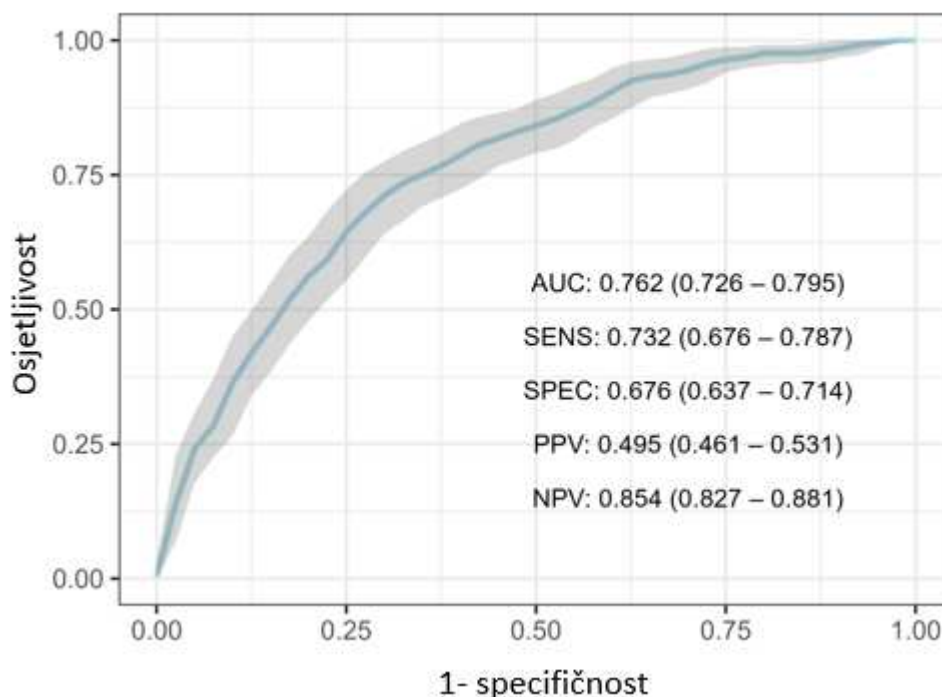
Slika 8. ROC-AUC krivulja za model Slučajne šume (NST-skupina)

Tablica 4. Performanse modela za NST-skupinu

Model	Srednji AUC (95% IP)
Slučajna šuma	0,793 (0,713-0,865)
XGBoost	0,783 (0,703-0,858)
Logistička regresija	0,763 (0,683-0,838)
TabNet	0,740 (0,652-0,812)
KNN	0,701 (0,622-0,782)
Univarijatna logistička regresija	0,688 (0,626-0,745)

4.2.2. Performanse za ukupnu populaciju

Na ukupnoj ispitivanoj populaciji XGBoost se pokazao kao najbolji model sa srednjom vrijednosti AUC-a od 0.762 (95% IP: 0.726-0.794), dok su slučajne šume drugi najbolji model s minimalnom razlikom (AUC 0.760 (95% IP: 0.710-0.780)) (Slika 9., Tablica 5.). Poput Slučajne šume i XGBoosta, također su i logističke regresije postigle slabiji rezultat na ukupnoj populaciji, gdje je univarijatna logistička regresija postigla srednji AUC od 0.589 (95% IP: 0.577-0.614), dok je logistička regresija postigla 0.741 (95% IP: 0.706-0.775) (Tablica 5.). Naposljetku, KNN algoritam je postigao srednji AUC od 0.612 (95% IP: 0.572-0.652), dok je TabNet postigao AUC od 0.740 (95% IP: 0.708-0.776) (Tablica 5.).



AUC- površina ispod krivulje, SENS- osjetljivost, SPEC- specifičnost, PPV- pozitivna prediktivna vrijednost, NPV- negativna prediktivna vrijednost.

Slika 9. ROC-AUC krivulja za XGBoost model (ukupna populacija)

Tablica 5. Performanse modela za ukupnu populaciju

Model	Srednji AUC (95% IP)
XGBoost	0,762 (0,726-0,795)
Slučajna šuma	0,760 (0,710-0,780)
Logistička regresija	0,741 (0,718-0,776)
TabNet	0,740 (0,708-0,776)
KNN	0,612 (0,572-0,652)
Univarijatna logistička regresija	0,589 (0,577-0,614)

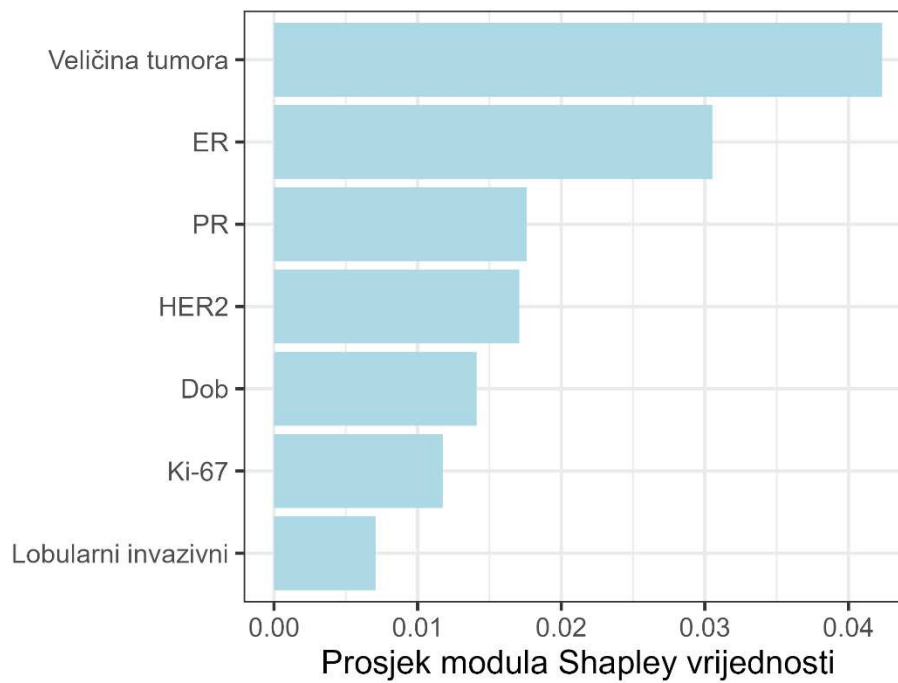
Što se tiče ostalih metrika pri bazičnom pragu odluke, XGBoost je postigao osjetljivost od 0,344 (95% IP: 0,289-0,403), specifičnost 0,903 (95% IP: 0,877-0,926), F1-mjeru 0,448 (95% IP: 0,389-0,507), pozitivnu prediktivnu vrijednost od 0,607 (95% IP: 0,539-0,680) te negativnu prediktivnu vrijednost 0,761 (95% IP: 0,744-0,778) (Slika 9.).

Kako bi ispravili klasne nepravilnosti (69.7% negativni LČ vs 30.30% pozitivni LČ) promijenjen je osnovni prag tako što se maksimizirala F1-mjera. Smanjivanje praga odluke na 0,28 povećalo je F1-mjeru na 0,581 (95% IP: 0,545-0,618), osjetljivost na 0,732 (95% IP: 0,676-0,787) i negativnu prediktivnu vrijednost na 0,854 (95% IP: 0,827-0,881). Suprotno tomu, specifičnost se smanjila na 0,676 (95% IP: 0,637-0,714), te pozitivna prediktivna vrijednost na 0,495 (95% IP: 0,461-0,531).

Kod evaluacije XGBoost modela na skupu podataka kojemu neke varijable nedostaju postignut je nešto lošiji rezultat s srednjim AUC od 0,731 (95% IP: 0,634-0,771).

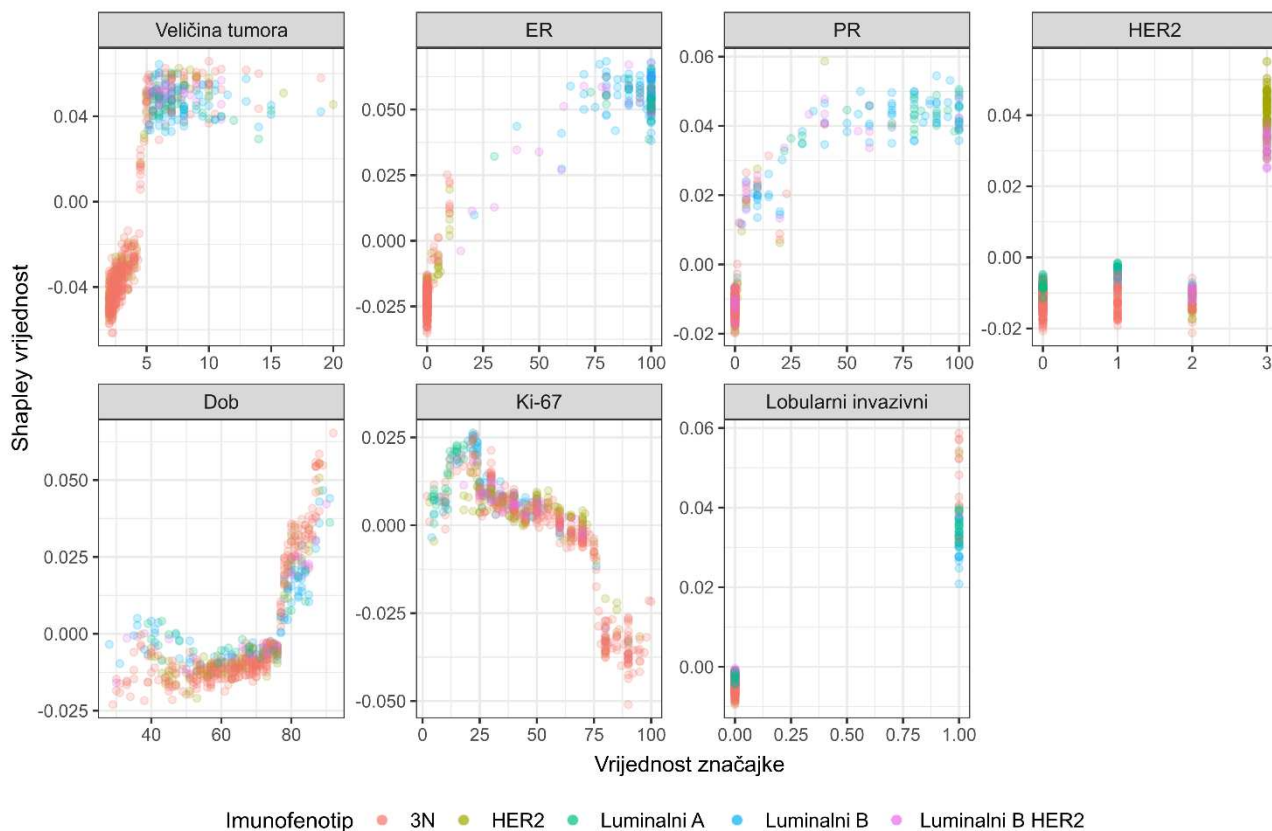
4.2.3. Analiza važnosti prediktora

Nakon izračuna Shapely vrijednosti za NST skupinu, veličina tumora bila je najvažnija značajka, a po važnosti slijede ER, PR i HER2 status (Slika 10.).



Slika 10. Najvažniji prediktori po Shapley vrijednostima za NST skupinu.

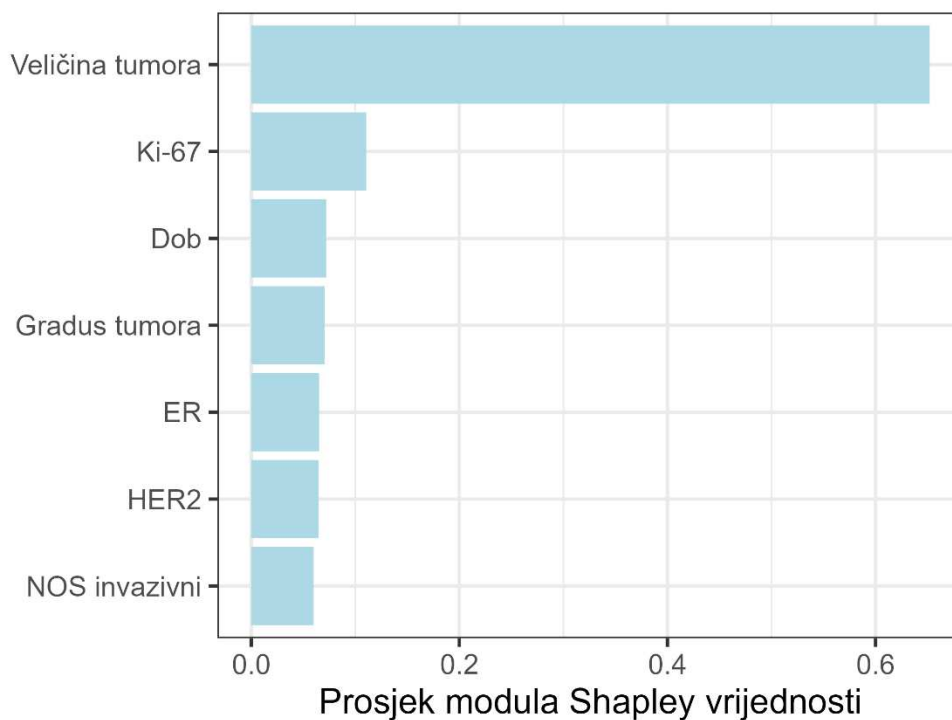
Metastaze u limfne čvorove pokazale su linearnu ovisnost o veličini tumora do 5 cm, nakon čega se postiže plato (Slika 11.). Što se tiče dobi, postoji jasan i nagli porast povezanosti nakon 75. godine (Slika 11.). S višim vrijednostima ER i PR veća je povezanost sa zahvaćenošću pazušnih limfnih čvorova, dok je negativnost hormonskih receptora povezana s odsutnošću metastaza (Slika 11.). Postoji značajna korelacija između statusa hormonskih receptora i veličine tumora. Naime, većina tumora s visokom izražajnošću ER i PR bili su imunofenotipa luminalni A i luminalni B, a morali su biti > 5 cm u veličini kako bi se svrstali u NST skupinu.



Slika 11. Odnos najvažnijih prediktora i metastaza u LČ za NST skupinu

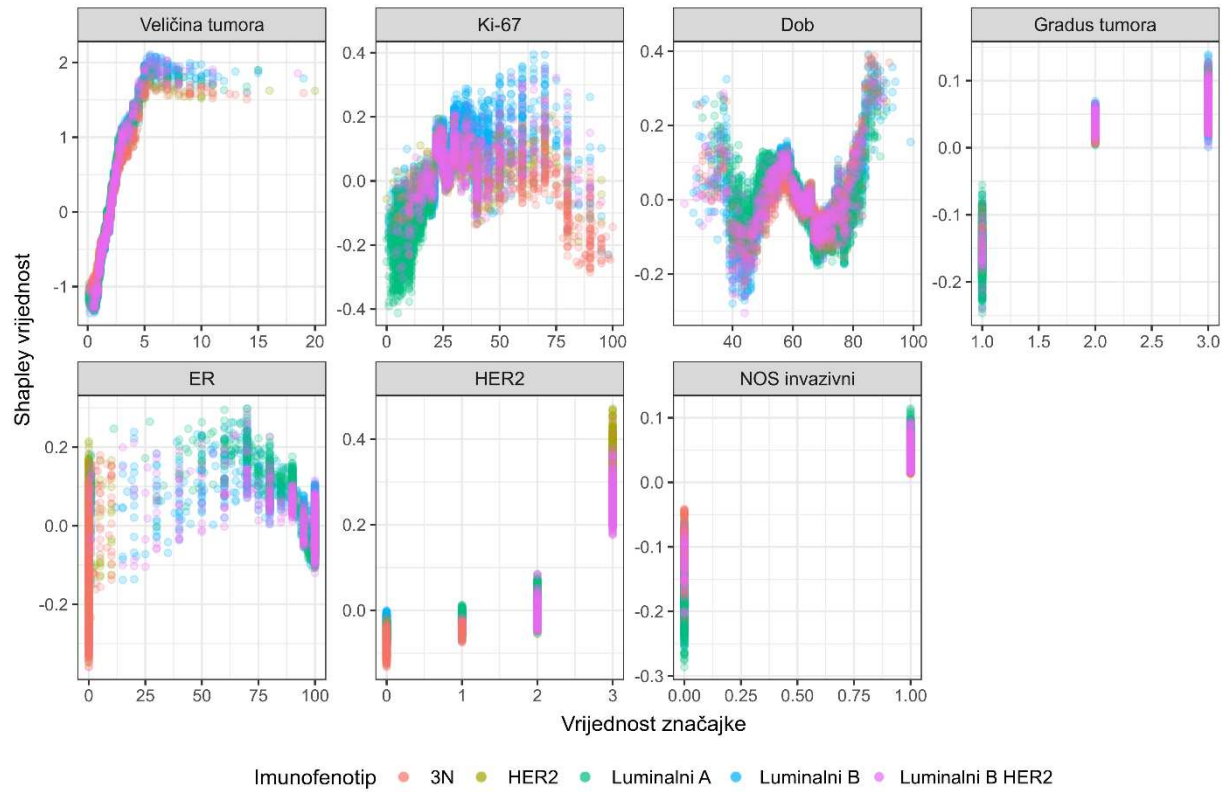
Za HER-2, model je povezao HER-2 pozitivnost s većom vjerojatnošću za metastaze u LČ (Slika 10. i Slika 11.). Zanimljivo, Ki-67 pokazuje povećanje Shapley vrijednosti od 0 do 25%, nakon čega se postupno smanjuje, s naglim padom vrijednosti na oko 75%. Međutim, nakon detaljnijeg uvida, možemo vidjeti da su te vrijednosti pretežno povezane s trostruko negativnim imunofenotipom (Slika 11.). Konačno, model je povezao lobularni invazivni histološki tip s većom vjerojatnošću metastaza u limfnim čvorovima (Slika 10.).

Dok je veličina tumora također bila najvažnija značajka kada su Shapleyjeve vrijednosti izračunate na ukupnoj populaciji studije, druga najvažnija značajka bila je Ki-67, a zatim dob te gradus tumora (Slika 12.).



Slika 12. Najvažniji prediktora po Shapley vrijednostima za ukupnu populaciju.

Za veličinu tumora, trend iz skupine NST kriterija potvrđen je na ukupnoj ispitivanoj populaciji (rastuća povezanost, s platoon nakon 5 cm) (Slika 13). Isto vrijedi i za HER-2 status, gdje je HER-2 pozitivnost povezana s metastazama u limfnim čvorovima (Slika 13). Gradus tumora također pokazuje jasan trend povećanja (Slika 13.). Za Ki-67 postoji primjetan porast Shapley vrijednosti pri vrijednostima indeksa od 25%, uz smanjenje od oko 75%. Međutim, Shapleyjeve vrijednosti pokazuju visoku razinu disperzije (Slika 13.), što ukazuje na ovisnost o vrijednostima drugih varijabli. Dob pokazuje zanimljivu nelinearnu ovisnost, gdje su bolesnici mlađi od 40 i stariji od 75 povezani s većom vjerojatnošću metastaza (Slika 13.).



Slika 13. Odnos najvažnijih prediktora i metastaza u LČ za ukupnu populaciju

5. RASPRAVA

Kako bi se poboljšali ishodi liječenja, sve veći broj bolesnica s ranim rakom dojke liječi se NST-om. Nadalje, kako bi se postigla optimalna strategija liječenja za svakog bolesnika i implementirala precizna onkologija, zbog neadekvante izvedbe postojećih dijagnostičkih alata, sve je veća potreba za dodatnim metodama procjene statusa metastaza u limfnim čvorovima. Zatim, za određivanje statusa aksilarnih limfnih čvorova u bolesnica s rakom dojke koje su inicijalno liječene kirurški, preporučuje se disekcija limfnih čvorova čuvara (SLN). Zabilježeno je da disekcija limfnih čvorova čuvara ima udio lažno negativnih rezultata od 7.5%, ali može doseći i 27.3%, što rezultira neradikalnom aksilarnom operacijom i posljedično neoptimalnim strategijama adjuvantne terapije temeljenim na pogrešnom klasificiranju (83, 84). Stoga, osim bolesnika koji primaju NST, bolesnici koji su inicijalno liječeni kirurški također mogu imati koristi od dodatnih neinvazivnih metoda za određivanje statusa limfnih čvorova.

U ovoj studiji uvježbano je, optimizirano i validirano više modela strojnog učenja koji mogu učinkovito pomoći u predviđanju metastaza u pazušnim limfnim čvorovima raka dojke. Slučajna šuma, metoda koja se temelji na ansamblu, proizvela je najvišu srednju vrijednost AUC-a (0,793 za NST skupinu, 0,760 za ukupnu populaciju studije). Budući da se algoritam Slučajne šume sastoji od više Stabala odluke i kasnije kombinira njihova predviđanja, on smanjuje rizik od prenaučivosti i tako pruža robusniji model (85). XGBoost, još jedan robusni skupni model temeljen na stablima odluke, dao je nešto bolji rezultat od Slučajne šume na ukupnoj populaciji studije. To je u skladu s prethodnim studijama koje pokazuju da XGBoost i Slučajna šuma općenito postižu usporedive rezultate, ali na različitim primjerima, jedan će nadmašiti drugoga (86). Naposljetku, oba su modela postigla bolje performanse u usporedbi s osnovnim univarijantnim modelom koji je koristio samo veličinu tumora kao prediktor.

Kada se razmatra važnost značajki prema Shapleyjevim vrijednostima, najvažnije značajke u skupini koja ispunjava NST kriterije bile su veličina tumora, ER, PR i HER2 status. Zahvaćenost čvorova rasla je s veličinom tumora do 5 cm, pa stoga naši nalazi potvrđuju nalaze

od Sopik i Naroda, gdje je također postignut plato u nodalnim metastazama nakon približno 5 cm veličine tumora (87). Za NST skupinu, Shapleyjeve vrijednosti pokazuju rastući trend s povećanjem razina ER i PR. Iako je to na prvi pogled iznenađujuće, ovo je otkriće vjerojatno artefakt kriterija odabira. Naime, temeljem kriterija odabira, tumori s visokim ER i/ili PR vrijednostima, koji odgovaraju imunofenotipovima luminalni A i B, uključeni su samo ako je veličina tumora veća od 5 cm. To dovodi do umjetne korelacije između visokih vrijednosti ER i vjerojatnosti metastaze u limfnim čvorovima, koja se zapravo temelji na utjecaju veličine tumora. Posljedično, rezultati koji se tiču učinka ER i PR nisu bili potvrđeni kada je ispitana važnost obilježja u ukupnoj populaciji studije. Literatura je također proturječna u pogledu uloge ER u predviđanju zahvaćenosti čvorova. Na primjer, u studiji koju su proveli Alsumai i sur., pozitivan ER status bio je značajan prediktor metastaza u čvorovima, dok u drugoj studiji ER i PR nisu imali značajan učinak na presadnice u pazušnim limfnim čvorovima (88, 89).

Nadalje, Shapleyjeve vrijednosti za dob bolesnika pokazale su da zahvaćenost čvorova raste nakon 75. godine života. Slične rezultate izvijestili su Wildiers i sur., gdje je zahvaćenost čvorova porasla nakon 70. godine (90).

Dok je veličina tumora također bila najvažnija značajka i u ukupnoj studijskoj populaciji, pokazalo se da je druga najvažnija značajka bila Ki67, a slijedi je gradus tumora. Zanimljivo, kada su Shapleyjeve vrijednosti za dob ispitanika izračunate na ukupnoj populaciji studije, primijećen je nelinearni trend, gdje su i ispitanici mlađi od 40 i stariji od 75 povezani s presadnicama u limfnim čvorovima. Veća pojavnost presadnica u limfnim čvorovima u bolesnika mlađih od 40 godina također je prijavljena u prethodnim studijama, gdje je najčešće ranija pojavnost tumora povezana sa genskom etiologijom i agresivnijim oblicima tumora (91, 92). Što se tiče analize Shapley vrijednosti Ki67, više vrijednosti Ki67 općenito su povezane s većom vjerojatnošću presadnica. Zanimljivo je da su za vrijednosti Ki67 > 75%, Shapleyjeve vrijednosti uglavnom negativne. Međutim, kada se uzme u obzir imunofenotip, ovaj obrat

trenda vidljiv je samo za trostruko negativne karcinome. Prethodne studije također su pokazale da je $Ki67 > 20\%$ u pozitivnoj korelaciji s metastazama u limfnim čvorovima, iako su koristile unaprijed definirane kategorije od $< 20\%$ i $> 20\%$ i nisu uzele u obzir razlike u imunofenotipu (93, 94). Stoga su potencijalno propustili složeniji odnos koji je objašnjen kroz Shapleyjeve vrijednosti.

Isto tako, veći gradus tumora obično je povezan s nodalnim presadnicama, što se i pokazalo Shapley vrijednostima gdje se prati jasan trend porasta vjerojatnosti za presadnice u pazušnim LČ s porastom gradusa (91). Model također povezuje pozitivnost HER-2 sa zahvaćenošću čvorova. HER-2 je također bio značajan prediktivni čimbenik za zahvaćenost aksilarnih čvorova (s koeficijentom regresije od 0.30) u studiji gdje su istraživači razvili Lasso regresijski model za predviđanje presadnica u limfnim čvorovima raka dojke (95). Na temelju navedenog, možemo zaključiti da se proces donošenja odluka našeg modela može klinički objasniti jer je naučio odnose koji su potvrđeni u prijašnjim studijama.

Osim prethodnih zapažanja, vrijedno je napomenuti da je Shapleyeva analiza neovisno identificirala najčešće korištene granične točke za ER, PR, Ki-67 i HER2 pozitivnost.

Trenutno smo svjedoci velikog zanimanja za studije temeljene na radiomici koje kombiniraju radiološke nalaze i metode dubokog učenja za predviđanje statusa limfnih čvorova raka dojke (96-99). Zanimljivo istraživanje Zhenga i sur. spaja radiomiku dubokog učenja i ultrazvučne elastografije s kliničko-patološkim podacima, te polučuje izvrsne rezultate s AUC od 0,902 (95% IP: 0,843-0,961) (80). Također su primijenili model koji je treniran samo na kliničko-patološkim značajkama podataka, što je proizvelo slabiji AUC u usporedbi s našim (0,72 [95% IP: 0,63-0,82] u odnosu na 0,79 [95% IP 0,72-0,87]) (80). Ova relativno lošija izvedba artificijalnih neuronskih mreža (ANN) mogla bi se potencijalno objasniti generalno lošijom izvedbom neuronskih mreža na tabličnim podacima (65). Pokazalo se da modeli temeljeni na stablima (Slučajna šuma i XGBoost) nadmašuju metode dubokog učenja na

skupovima podataka s do 10 000 primjera (66). Lošija izvedba ANN potvrdila se i u našem slučaju, gdje je TabNet postigao nešto lošije rezultate od algoritama utemeljenih na stablima (0,74 za skupinu s NST kriterijima i za ukupnu populaciju vs 0,79 i 0,76 za Slučajnu šumu i XGBoost).

Slično tome, istraživanje od Dihge i sur. koje je procjenjivalo ANN obučene na kliničko-patološkim značajkama za predviđanje zahvaćenosti limfnih čvorova raka dojke postiglo je AUC od 0,74 (95% IP: 0,72-0,76) (78). Međutim, jedna od njihovih značajki s najvećom prediktivnom moći bila je limfovaskularna invazija, značajka koja se ne može uvijek dobiti „core“ biopsijom i stoga se ne koristi u našem modelu (100).

Lošija izvedba od strane KNN algoritma može se objasniti takozvanim "prokletstvom dimenzionalnosti" („*Curse of dimensionality*“) (101). Naime, u podaktovnj znanosti dimenzionalnošću smatramo broj različitih varijabli korištenih u modelu. Obzirom da KNN koristi izračun udaljenosti između različitih točaka unutar podataka, veći broj dimenzija eksponencijalno povećava ukupni prostor u kojem se sve točke podataka nalaze (102). Shodno tomu, podatci s većim broj prediktora (poput prediktora korištenih u ovoj studiji) nisu idealni za uporabu KNN algoritma.

Slučajna šuma također je bila algoritam s najboljom izvedbom u studiji Tsenga i sur., s performansama sličnima našim (srednja vrijednost AUC od 0,75) (103). Još jedan model temeljen na stablu (XGBoost) pokazao se najboljim modelom u studiji Li i sur., gdje su autori koristili značajke tumorskih gena za predviđanje metastatskog statusa kod raka dojke (104). Njihov optimizirani model postigao je AUC od 0,82 (SD \pm 0,15) (104). Isto tako, u studiji koju su proveli Meng i sur., model temeljen na Lasso regresiji postigao je AUC od 0,77 (95% IP 0,69-0,86) za predviđanje statusa metastaza u limfnim čvorovima koji ne pripadaju SLN (95). Štoviše, starija studija koja je kombinirala kliničko-patološke nalaze s dijagnostičkim nalazima mamografije i ultrazvuka postigla je AUC od 0,77 (95% IP: 0,689-0,856) u predviđanju

pazušnih limfnih čvorova raka dojke s Izmjeničnim stablom odluke (ADTree) (79). Stoga su naši rezultati usporedivi s rezultatima drugih studija koje su predviđale status pazušnih limfnih čvorova raka dojke, kao i opći status presadnica (78-80, 95, 104).

Budući da je status limfnih čvorova bitan čimbenik koji utječe na optimalan izbor liječenja i prognozu, vjerujem da bi se slični modeli mogli optimizirati i evaluirati na lokalnoj razini kako bi pomogli kliničarima u multidisciplinarnom radu, osobito kada se radi o bolesnicima koji bi primili NST. Štoviše, drugi korisni podatci o tumoru i/ili bolesniku koji se mogu dobiti, poput ekspresije gena i serumskih biomarkera, mogli bi dovesti do boljih performansi modela. Sukladno tome, buduća istraživanja mogu procijeniti povećava li dodavanje genetskih podataka i podataka o biomarkerima točnost modela strojnog učenja.

Istraživanje na temelju kojeg je izrađena ova disertacija, sadrži nekoliko ograničenja. Prvo, provedena je samo na hrvatskoj populaciji bolesnika s rakom dojke. Stoga bi se slični modeli trebali potvrditi na drugim populacijama, kako bi se omogućila bolja generalizacija. Još jedno ograničenje ove studije je njezina retrospektivna priroda, iako podatci potječu iz prospektivno održavane baze podataka. Možda najvažnije ograničenje studije bio je relativno velik broj bolesnika koji su bili isključeni iz primarne analize zbog nepotpunih podataka. Naravno, nemamo dokaza da ti podatci nedostaju potpuno slučajno („*missing completely at random*“ -*MCAR*). Nadalje, moguće manje ograničenje generalizacije rezultata ove analize moglo je biti uzrokovano činjenicom da je dio podataka prikupljen tijekom karantene povezane s pandemijom bolesti COVID-19.

Zaključno, ovo istraživanje nudi novost predstavljanjem objašnjivog okvira strojnog učenja s klinički relevantnim procesom donošenja odluka. Daljnja snaga ove studije je u tome što pruža jedinstvenu perspektivu u kojoj je dobiven multicentrični skup podataka, a iz subjekata koji su inicijalno kirurški liječeni, ekstrahiran je dodatni podskup primjenom NCCN kriterija za NST (45). Na ovaj način, model je prilagođen za bolesnike prikladne za NST, koji

mogu izvući najveću korist od takve neinvazivne metode za određivanje statusa metastaza u pazušnim limfnim čvorovima.

6. ZAKLJUČCI

1. Objašnjivi modeli strojnog učenja temeljeni na ansamblu stabala odluke postižu dobre rezultate u klasifikaciji pazušnih limfnih čvorova u bolesnika s rakom dojke.
2. Slučajne šume i XGBoost postižu značajno bolje performanse od unvarijantne logističke regresije i KNN algoritma.
3. Prema Shapley vrijednostima veličina tumora je najvažniji prediktor za određivanje statusa pazušnih limfnih čvorova.
4. Ovisnost veličine tumora i metastaza u pazušne limfne čvorove linearno raste do veličine od 5 cm, a nakon čega se postiže plato.
5. Prema Shapley vrijednostima Ki67 je drugi prediktor po važnosti za određivanje statusa pazušnih limfnih čvorova, dok je gradus tumora treći po važnosti.
6. Ovisnost Ki67 i metastaza u pazušne limfne čvorove raste do razine od 75%, nakon čega se prati pad povezanosti poglavito za trostruko negativne tumore.
7. Shapley vrijednosti identificirale su dob < 40 god. i > 75 god. kao čimbenik povezan s metastazama u pazušne limfne čvorove.
8. S porastom gradusa tumora stupnjevito raste i vjerojatnost metastaza u pazušne limfne čvorove.
9. HER-2 pozitivnost povezana je s metastazama u pazušne limfne čvorove.
10. Shapleyeva analiza je neovisno identificirala najčešće korištene granične točke za ER, PR, Ki67 i HER2 pozitivnost.

Pokazana je učinkovitost objašnjivih metoda strojnog učenja temeljenih na stablima odluke obučeni na kliničko-patološkim značajkama dobivenih tijekom redovitih predoperativnih/pre-NST postupaka u predviđanju metastaza raka dojke u pazušnim limfnim čvorovima. Takvi modeli mogu dovesti do točnije dijagnoze i boljeg odabira liječenja, posebno za bolesnike određene za NST, gdje su radiološki i klinički nalazi često jedini način procjene limfnih čvorova. Potencijalni napredak u dijagnozi na temelju modela strojnog učenja za neke

bi bolesnike rezultirao NST-om i, posljedično, potencijalno drukčijim adjuvantnim pristupom s tretmanima koji nisu križno rezistentni te boljim ishodima za bolesnike. Od budućih studija očekuje se dodavanje genetskih podataka i podataka o biomarkerima te naknadna validacija u multinacionalnim i multicentričnim studijama.

7. SAŽETAK

Uvod: Određivanje statusa pazušnih limfnih čvorova iznimno je važno za prognozu bolesti i pravilan odabir liječenja kod bolesnika oboljelih od raka dojke. Kod bolesnika određenih za neoadjuvantno sustavno liječenje (NST) klinički pregled i radiološke metode primarni su način određivanja statusa pazušnih limfnih čvorova. Takve metode dovode do čak 30% pogrešno dijagnosticiranih slučajeva, čime se ukazuje na potrebu za dodatnim metodama. Zbog sve veće dostupnosti podataka u onkologiji, metode strojnog učenja idealan su oblik za poboljšanje dijagnostike pazušnih limfnih čvorova kod raka dojke. Cilj ove disertacije je treniranje i validacija modela strojnog učenja (Slučajna šuma, XGBoost, TabNet Linearna regresija i KNN) na kliničko-patološkim podacima, kojima je svrha točna klasifikacija metastatskih pazušnih limfnih čvorova.

Ispitanici i metode: U ovo istraživanje uključeni su anonimizirani podaci od ukupno 13580 bolesnika liječenih od raka dojke. Uzorak čine ispitanici liječeni od raka dojke u 25 hrvatskih bolnica u razdoblju od siječnja 2017. godine do siječnja 2022. godine. Kroz navedeno razdoblje podatci su prikupljeni iz bolničkih informacijskih sustava (BIS), tako što su pretraživani svi bolesnici s MKB šifrom 50 (za rak dojke). Uključeni su svi ispitanici s potpunim podacima (PHD, status LČ, ER, PR, HER2, Ki67, veličina tumora, dob). Isključen je iz istraživanja dio ispitanika s nepotpunim podacima (nedostatak neke od ulaznih varijabli), nakon čega je preostalo 9705 ispitanika. Među tim ispitanicima, njih 1324 je inicijalno primilo NST zbog čega su također isključeni iz analize. Naposljetku, nakon primjene kriterija za NST, identificirano je 719 ispitanika za završnu analizu, optimizaciju i validaciju prediktivnih modela. Također, izrađeni su i modeli na cjelokupnoj populaciji s potpunim podacima (n= 8381). Optimizacija hiperparametara modela izvršena je tehnikom peterogrupne križne validacije. Validacija modela izvršena je tehnikom „*bootstrap*“ na „*test*“ setu podataka. Analiza važnosti prediktora i procesa donošenja odluke modela izvršena je SHAP tehnikom.

Rezultati: U skupini koja ispunja NST kriterije najbolje performanse je pokazao model Slučajna šuma sa srednjim AUC od 0.793 (95% IP: 0.713-0.865), dok je XGBoost imao najbolju performansu na ukupnoj populaciji sa srednjim AUC od 0.762 (95% IP: 0.726-0.794). Analiza Shapley vrijednosti je pokazala kako je najvažniji prediktor veličina tumora, praćena Ki-67 indeksom, dobi ispitanika i gradusom tumora. Povezanost metastaza u pazušne limfne čvorove i veličine tumora linearno raste do veličine od 5 cm, nakon koje se postiže plato. Povezanost Ki67 s metastazama u pazušne LČ raste do Ki-67 od 75%, nakon čega se prati pad, ali prvenstveno za trostruko negativni tip raka dojke. Slučajna šuma i XGBoost postižu značajan napredak u performansama naspram bazičnog modela univarijantne logističke regresije (trenirane samo na veličini tumora).

Zaključak: Ovo istraživanje pokazalo je kako objašnjive metode strojnog učenja temeljene na stablu odluke, trenirane na kliničko-patološkim značajkama dobivenim tijekom redovitih predoperativnih/pre-NST postupaka postižu dobre rezultate u predviđanju metastaza raka dojke u aksilarnim limfnim čvorovima. Takvi modeli mogu dovesti do točnije dijagnoze i boljeg odabira liječenja, posebno za bolesnike određene za NST, gdje su radiološki i klinički nalazi često jedini način procjene limfnih čvorova. Potencijalni napredak u dijagnostici na temelju modela strojnog učenja za neke bi bolesnike rezultirao NST-om i, posljedično, drugačijim adjuvantnim liječenjem s tretmanima koji nisu križno rezistentni i boljim ishodima za bolesnike. Od budućih studija očekuje se dodavanje genetskih podataka i podataka o biomarkerima te naknadna validacija u multinacionalnim/multicentričnim studijama.

8. SUMMARY

TITLE: MACHINE LEARNING MODELS FOR PREDICTING LYMPH NODE METASTASES IN BREAST CANCER PATIENTS

Introduction: Determining the status of axillary lymph nodes is extremely important for prognosis and correct therapy selection in breast cancer patients. In patients designated for neoadjuvant systemic therapy (NST), clinical examination and radiological methods are the primary means of determining the status of axillary lymph nodes. Such methods lead to as many as 30% of misdiagnosed cases, which indicates the need for supplementary methods. Due to the increasing availability of data in oncology, machine learning methods are an ideal candidate for improving the diagnosis of axillary lymph nodes in breast cancer. The goal of this dissertation is the training and validation of machine learning models (Random Forest, XGBoost, TabNet Linear Regression and KNN) on clinico-pathological data, with the purpose of accurate axillary lymph node classification.

Subjects and Methods: Anonymized data from a total of 13580 patients treated for breast cancer were included in this research. The sample consists of subjects treated for breast cancer in 25 Croatian hospitals in the period from January 2017, to January 2022. Data were collected from hospital information systems during the mentioned period, by searching all patients with ICD code 50 (for breast cancer). All subjects with complete data (PHD, lymph node status, ER, PR, HER2, Ki67, tumor size, age) were included. Subjects with incomplete data (missing some of the input variables) was excluded, after which 9705 subjects remained. Among these, 1324 of them initially received NST, which is why they were also excluded from the analysis. Finally, after applying the criteria for NST, 719 patients were identified for final analysis, optimization and validation of the predictive models. Also, models were trained and validated on the entire population with complete data (n= 8381). Hyperparameter optimization was performed using the 5-fold cross validation technique. Validation of the model was performed using the

"bootstrap" technique on the "test" data set. The feature importance analysis and the analysis of the model's decision-making process was performed using the SHAP technique.

Results: In the group meeting the NST criteria, the best performance achieved by the Random forest model with a mean AUC of 0.793 (95% CI: 0.713-0.865), while XGBoost had the best performance on the total population with a mean AUC of 0.762 (95% CI: 0.726-0.794). Shapley value analysis showed that the most important predictor is tumor size, followed by Ki67 index, patient age and tumor grade. The association between axillary lymph node metastases and tumor size grows linearly up to a size of 5 cm, after which a plateau is reached. The association of Ki67 with LN metastasis increases up to Ki67 of 75%, after which a decline is observed, but mostly for the triple-negative type of breast cancer. Random forest and XGBoost achieve significant performance gains over the baseline univariate logistic regression model (trained only on tumor size).

Conclusion: Explainable tree-based machine learning methods trained on patient and tumor features obtained during regular pre-operative/pre-NST procedures achieve a good performance in predicting breast cancer axillary lymph node metastasis. Such models can lead to more accurate diagnosis and better treatment selection, especially for NST patients, where radiological and clinical findings are often the only way of axillary lymph node assessment. Potential upstage of diagnosis based on machine learning models for some patients would result in NST and, consecutively, potentially more adjuvant therapy with non-cross resistant treatments and better patient outcomes. The addition of genetic and biomarker data and subsequent validation in multinational/multicenter studies is expected from future studies.

9. LITERATURA

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin.* 2022;72:7-33.
2. Giaquinto AN, Sung H, Miller KD, Kramer JL, Newman LA, Minihan A, i sur. Breast Cancer Statistics, 2022. *CA Cancer J Clin.* 2022;72:524-41.
3. Hrvatski zavod za javno zdravstvo [Internet]. Citirano 30.3.2023. Dostupno na: <https://www.hzjz.hr/sluzba-epidemiologija-prevencija-nezaraznih-bolesti/publikacije-odjel-za-maligne-bolesti/>
4. Dorling L, Carvalho S, Allen J, González-Neira A, Luccarini C, Wahlström C, i sur. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med.* 2021;384:428-39.
5. Kelsey JL. A review of the epidemiology of human breast cancer. *Epidemiol Rev.* 1979;1:74-109.
6. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. *Epidemiol Rev.* 1993;15:36-47.
7. Van den Brandt PA, Spiegelman D, Yaun S-S, Adami H-O, Beeson L, Folsom AR, i sur. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol.* 2000;152:514-27.
8. Ciszewski T, Jopacka-Szatan K. Breast cancer risk factors. *Prz Menopauzalny.* 2015;14:196-202.
9. Feng Y, Spezia M, Huang S, Yuan C, Zeng Z, Zhang L, i sur. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes & Diseases.* 2018;5:77-106.
10. Dupont WD, Parl FF, Hartmann WH, Brinton LA, Winfield AC, Worrell JA, i sur. Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. *Cancer.* 1993;71:1258-65.
11. Rosen EM, Pishvaian MJ. Targeting the BRCA1/2 tumor suppressors. *Curr Drug Targets.* 2014;15:17-31.
12. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A.* 1971;68:820-3.
13. Fodde R, Smits R. Cancer biology. A matter of dosage. *Science.* 2002;298:761-3.
14. Matoba S, Kang J-G, Patino WD, Wragg A, Boehm M, Gavrilova O, i sur. p53 regulates mitochondrial respiration. *Science.* 2006;312:1650-3.
15. Liberti MV, Locasale JW. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci.* 2016;41:211-8.

16. Harris JR, Lippman ME, Osborne CK, Morrow M. *Diseases of the Breast*: Lippincott Williams & Wilkins; 2012.
17. Tan PH, Ellis I, Allison K, Brogi E, Fox SB, Lakhani S, i sur. The 2019 World Health Organization classification of tumours of the breast. *Histopathology*. 2020;77:181-5.
18. Board WCoTE, Cancer IAfRo. *WHO Classification of Breast Tumours: WHO Classification of Tumours, Volume 2*: World Health Organization; 2019.
19. Cserni G. Histological type and typing of breast carcinomas and the WHO classification changes over time. *Pathologica*. 2020;112:25-41.
20. Sawaki M, Shien T, Iwata H. TNM classification of malignant tumors (Breast Cancer Study Group). *Jpn J Clin Oncol*. 2018;49:228-31.
21. Umemura S, Kurosumi M, Moriya T, Oyama T, Arihiro K, Yamashita H, i sur. Immunohistochemical evaluation for hormone receptors in breast cancer: a practically useful evaluation system and handling protocol. *Breast Cancer*. 2006;13:232.
22. Fulawka L, Blaszczyk J, Tabakov M, Halon A. Assessment of Ki-67 proliferation index with deep learning in DCIS (ductal carcinoma in situ). *Sci Rep*. 2022;12:3166.
23. Yanagawa M, Ikemot K, Kawauchi S, Furuya T, Yamamoto S, Oka M, i sur. Luminal A and luminal B (HER2 negative) subtypes of breast cancer consist of a mixture of tumors with different genotype. *BMC Res Notes*. 2012;5:1756-0500.
24. Takahashi H, Oshi M, Asaoka M, Yan L, Endo I, Takabe K. Molecular Biological Features of Nottingham Histological Grade 3 Breast Cancers. *Ann Surg Oncol*. 2020;27:4475-85.
25. He Z, Chen Z, Tan M, Elingarami S, Liu Y, Li T, i sur. A review on methods for diagnosis of breast cancer cells and tissues. *Cell Prolif*. 2020;53:12.
26. Pusina S. Diagnostic accuracy of sentinel lymph node biopsy in axillary lymph nodes at the early stages of breast cancer. *Med Arch*. 2013;67:252-5.
27. Diessner J, Anders L, Herbert S, Kiesel M, Bley T, Schlaiss T, i sur. Evaluation of different imaging modalities for axillary lymph node staging in breast cancer patients to provide a personalized and optimized therapy algorithm. *J Cancer Res Clin Oncol*. 2022.
28. Choi HY, Park M, Seo M, Song E, Shin SY, Sohn Y-M. Preoperative axillary lymph node evaluation in breast cancer: current issues and literature review. *Ultrasound Q*. 2017;33:6-14.
29. Marino MA, Avendano D, Zapata P, Riedl CC, Pinker K. Lymph Node Imaging in Patients with Primary Breast Cancer: Concurrent Diagnostic Tools. *Oncologist*. 2020;25:e231-e42.
30. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, i sur. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and

lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med.* 2002;347:1233-41.

31. Boyages J. Radiation therapy and early breast cancer: current controversies. *Med J Aust.* 2017;207:216-22.

32. Nabholz J-M, Gligorov J. The role of taxanes in the treatment of breast cancer. *Expert Opin Pharmacother.* 2005;6:1073-94.

33. Gewirtz D. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem Pharmacol.* 1999;57:727-41.

34. Tian H, Ma D, Tan X, Yan W, Wu X, He C, et al. Platinum and Taxane Based Adjuvant and Neoadjuvant Chemotherapy in Early Triple-Negative Breast Cancer: A Narrative Review. *Front Pharmacol.* 2021;12.

35. Group EBCTC. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *The Lancet.* 2011;378:771-84.

36. Lee CI, Goodwin A, Wilcken N. Fulvestrant for hormone-sensitive metastatic breast cancer. *Cochrane Database of Systematic Reviews.* 2017(1).

37. Boekhout AH, Beijnen JH, Schellens JH. Trastuzumab. *Oncologist.* 2011;16:800-10.

38. Ishii K, Morii N, Yamashiro H. Pertuzumab in the treatment of HER2-positive breast cancer: an evidence-based review of its safety, efficacy, and place in therapy. *Core Evid.* 2019;14:51-70.

39. LoRusso PM, Weiss D, Guardino E, Girish S, Sliwkowski MX. Trastuzumab Emtansine: A Unique Antibody-Drug Conjugate in Development for Human Epidermal Growth Factor Receptor 2-Positive Cancer. *Clin Cancer Res.* 2011;17:6437-47.

40. Martin JM, Goldstein LJ. Profile of abemaciclib and its potential in the treatment of breast cancer. *Onco Targets Ther.* 2018;11:5253-9.

41. Selli C, Sims AH. Neoadjuvant Therapy for Breast Cancer as a Model for Translational Research. *Breast Cancer.* 2019;13

42. Sims AH, Bartlett JM. Approaches towards expression profiling the response to treatment. *Breast Cancer Res.* 2008;10:115.

43. Bardia A, Baselga J. Neoadjuvant therapy as a platform for drug development and approval in breast cancer. *Clin Cancer Res.* 2013;19:6360-70.

44. Korde LA, Somerfield MR, Carey LA, Crews JR, Denduluri N, Hwang ES, i sur. Neoadjuvant chemotherapy, endocrine therapy, and targeted therapy for breast cancer: ASCO guideline. *J Clin Oncol.* 2021;39:1485-505.

-
45. Gradishar WJ, Moran MS, Abraham J, Aft R, Agnese D, Allison KH, i sur. Breast cancer, version 3.2022, NCCN clinical practice guidelines in oncology. *JNCCN*. 2022;20:691-722.
 46. Gianni L, Pienkowski T, Im YH, Tseng LM, Liu MC, Lluch A, i sur. 5-year analysis of neoadjuvant pertuzumab and trastuzumab in patients with locally advanced, inflammatory, or early-stage HER2-positive breast cancer (NeoSphere): a multicentre, open-label, phase 2 randomised trial. *Lancet Oncol*. 2016;17:791-800.
 47. von Minckwitz G, Huang C-S, Mano MS, Loibl S, Mamounas EP, Untch M, i sur. Trastuzumab Emtansine for Residual Invasive HER2-Positive Breast Cancer. *N Engl J Med*. 2018;380:617-28.
 48. von Minckwitz G, Procter M, de Azambuja E, Zardavas D, Benyunes M, Viale G, i sur. Adjuvant Pertuzumab and Trastuzumab in Early HER2-Positive Breast Cancer. *N Engl J Med*. 2017;377:122-31.
 49. Jovel J, Greiner R. An Introduction to Machine Learning Approaches for Biomedical Research. *Front Med*. 2021;8.
 50. Sah S. Machine Learning: A Review of Learning Types. *ArXiv*. 2020.
 51. Singh A, Thakur N, Sharma A, editors. A review of supervised machine learning algorithms. 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom); 2016 16-18 March 2016.
 52. Dridi S. Unsupervised Learning - A Systematic Literature Review. *ArXiv*. 2021.
 53. Loftus TJ, Shickel B, Balch JA, Tighe PJ, Abbott KL, Fazzino B, i sur. Phenotype clustering in health care: A narrative review for clinicians. *Front Artif Intell*. 2022;5.
 54. Jia J, Wang W, editors. Review of reinforcement learning research. 2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC); 2020 16-18 Oct. 2020.
 55. Khan S, Sajjad M, Hussain T, Ullah A, Imran AS. A Review on Traditional Machine Learning and Deep Learning Models for WBCs Classification in Blood Smear Images. *IEEE Access*. 2021;9:10657-73.
 56. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;523:436-44.
 57. Londhe VY, Bhasin B. Artificial intelligence and its potential in oncology. *Drug Discov Today*. 2019;24:228-32.
 58. Altameem A, Mahanty C, Poonia RC, Saudagar AKJ, Kumar R. Breast Cancer Detection in Mammography Images Using Deep Convolutional Neural Networks and Fuzzy Ensemble Modeling Techniques. *Diagnostics*. 2022;12:1812.
 59. Wakili MA, Shehu HA, Sharif MH, Sharif MHU, Umar A, Kusetogullari H, et al. Classification of Breast Cancer Histopathological Images Using DenseNet and Transfer Learning. *Comput Intell Neurosci*. 2022;10.

-
60. Blasiak A, Khong J, Kee T. CURATE.AI: Optimizing Personalized Medicine with Artificial Intelligence. *SLAS Technology*. 2020;25:95-105.
 61. Fürnkranz J. Decision Tree. In: Sammut C, Webb GI, editors. *Encyclopedia of Machine Learning*. Boston, MA: Springer US; 2010. 263-7.
 62. Zhang Z. Introduction to machine learning: k-nearest neighbors. *Ann Transl Med*. 2016;4:37.
 63. Sharma S, Kumar S. Comparative analysis of Manhattan Andeuclidean distance metrics using a* algorithm. *J Eng Appl Sci*. 2016;01:196-8.
 64. Gareth James DWTHRT. *An introduction to statistical learning : with applications in R*: New York : Springer, [2013]
 65. Shwartz-Ziv R, Armon A. Tabular data: Deep learning is not all you need. *Inf Fusion*. 2022;81:84-90.
 66. Grinsztajn L, Oyallon E, Varoquaux G. Why do tree-based models still outperform deep learning on tabular data? *ArXiv*. 2022.
 67. Friedman J. Greedy Function Approximation: A Gradient Boosting Machine. *Ann Stat*. 2000;29.
 68. Mason L, Baxter J, Bartlett P, Frean M. *Boosting Algorithms as Gradient Descent in Function Space*. 1999.
 69. Arik SÖ, Pfister T, editors. *Tabnet: Attentive interpretable tabular learning*. Proceedings of the AAAI Conference on Artificial Intelligence; 2021.
 70. Wang S-C. Artificial Neural Network. In: Wang S-C, editor. *Interdisciplinary Computing in Java Programming*. Boston, MA: Springer US; 2003. p. 81-100.
 71. Maleki F, Muthukrishnan N, Ovens K, Md C, Forghani R. Machine Learning Algorithm Validation. *Neuroimaging Clinics of North America*. 2020;30:433-45.
 72. Raschka S. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. *ArXiv*. 2018;abs/1811.12808.
 73. Sundararajan M, Najmi A. The Many Shapley Values for Model Explanation. In: Hal D, III, Aarti S, editors. *Proceedings of the 37th International Conference on Machine Learning; Proceedings of Machine Learning Research: PMLR*; 2020. p. 9269--78.
 74. Shapley LS. A value for n-person games. *ArXiv*. 1953.
 75. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015:770-8.

-
76. Wang Z, Sun H, Li J, Chen J, Meng F, Li H, i sur. Preoperative Prediction of Axillary Lymph Node Metastasis in Breast Cancer Using CNN Based on Multiparametric MRI. *J Magn Reson Imaging*. 2022;56:700-9.
77. Ozaki J, Fujioka T, Yamaga E, Hayashi A, Kujiraoka Y, Imokawa T, i sur. Deep learning method with a convolutional neural network for image classification of normal and metastatic axillary lymph nodes on breast ultrasonography. *Jpn J Radiol*. 2022;40:814-22.
78. Dihge L, Ohlsson M, Edén P, Bendahl P-O, Rydén L. Artificial neural network models to predict nodal status in clinically node-negative breast cancer. *BMC Cancer*. 2019;19:610.
79. Takada M, Sugimoto M, Naito Y, Moon HG, Han W, Noh DY, i sur. Prediction of axillary lymph node metastasis in primary breast cancer patients using a decision tree-based model. *BMC Med Inform Decis Mak*. 2012;12:1472-6947.
80. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, i sur. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun*. 2020;11:1236.
81. Chen H, Lan X, Yu T, Li L, Tang S, Liu S, i sur. Development and validation of a radiogenomics model to predict axillary lymph node metastasis in breast cancer integrating MRI with transcriptome data: A multicohort study. *Front Oncol*. 2022;12.
82. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework. 2019. 2623-31 p.
83. Pesek S, Ashikaga T, Krag LE, Krag D. The false-negative rate of sentinel node biopsy in patients with breast cancer: a meta-analysis. *World J Surg*. 2012;36:2239-51.
84. Li H, Jun Z, Zhi-Cheng G, Xiang Q. Factors that affect the false negative rate of sentinel lymph node mapping with methylene blue dye alone in breast cancer. *J Int Med Res*. 2019;47:4841-53.
85. Breiman L. Random forests. *Machine learning*. 2001;45:5-32.
86. Bentéjac C, Csörgo A, Martínez-Muñoz G. A comparative analysis of xgboost. *ArXiv*. 2019.
87. Sopik V, Narod SA. The relationship between tumour size, nodal status and distant metastases: on the origins of breast cancer. *Breast Cancer Res Treat*. 2018;170:647-56.
88. Alsumai TS, Alhazzaa N, Alshamrani A, Assiri S, Alhefdhi A. Factors predicting positive sentinel lymph node biopsy in clinically node-negative breast cancer. *Breast Cancer*. 2022:323-34.
89. Yoshihara E, Smeets A, Laenen A, Reynders A, Soens J, Van Ongeval C, i sur. Predictors of axillary lymph node metastases in early breast cancer and their applicability in clinical practice. *The Breast*. 2013;22:357-61.

-
90. Wildiers H, Van Calster B, van de Poll-Franse LV, Hendrickx W, Røislien J, Smeets A, i sur. Relationship between age and axillary lymph node involvement in women with breast cancer. *J Clin Oncol.* 2009;27:2931-7.
 91. Rivadeneira DE, Simmons RM, Christos PJ, Hanna K, Daly JM, Osborne MP. Predictive factors associated with axillary lymph node metastases in T1a and T1b breast carcinomas: analysis in more than 900 patients. *J Am Coll Surg.* 2000;191:1-6.
 92. Gajdos C, Tartter PI, Bleiweiss IJ. Lymphatic invasion, tumor size, and age are independent predictors of axillary lymph node metastases in women with T1 breast cancers. *Ann Surg.* 1999;230:692.
 93. Yin Y, Zeng K, Wu M, Ding Y, Zhao M, Chen Q. The levels of Ki-67 positive are positively associated with lymph node metastasis in invasive ductal breast cancer. *Cell Biochem Biophys.* 2014;70:1145-51.
 94. Jiang Y, Xu H, Zhang H, Ou X, Xu Z, Ai L, i sur. Nomogram for prediction of level 2 axillary lymph node metastasis in proven level 1 node-positive breast cancer patients. *Oncotarget.* 2017;8:72389.
 95. Meng L, Zheng T, Wang Y, Li Z, Xiao Q, He J, i sur. Development of a prediction model based on LASSO regression to evaluate the risk of non-sentinel lymph node metastasis in Chinese breast cancer patients with 1–2 positive sentinel lymph nodes. *Sci Rep.* 2021;11:19972.
 96. Yu Y, Tan Y, Xie C, Hu Q, Ouyang J, Chen Y, i sur. Development and validation of a preoperative magnetic resonance imaging radiomics–based signature to predict axillary lymph node metastasis and disease-free survival in patients with early-stage breast cancer. *JAMA Netw Open.* 2020;3
 97. Pesapane F, Rotili A, Agazzi GM, Botta F, Raimondi S, Penco S, i sur. Recent radiomics advancements in breast cancer: lessons and pitfalls for the next future. *Curr Oncol.* 2021;28:2351-72.
 98. Guo X, Liu Z, Sun C, Zhang L, Wang Y, Li Z, i sur. Deep learning radiomics of ultrasonography: identifying the risk of axillary non-sentinel lymph node involvement in primary breast cancer. *EBioMedicine.* 2020;60:103018.
 99. Sannasi Chakravarthy S, Rajaguru H. Automatic Detection and Classification of Mammograms Using Improved Extreme Learning Machine with Deep Learning. *IRBM.* Article; 2021.
 100. Harris GC, Denley HE, Pinder SE, Lee AH, Ellis IO, Elston CW, i sur. Correlation of histologic prognostic factors in core biopsies and therapeutic excisions of invasive breast carcinoma. *Am J Surg Pathol.* 2003;27:11-5.
 101. Hao N. Curse of Dimensionality. *Wiley StatsRef: Statistics Reference Online.* p. 1-3.

102. Kouiroukidis N, Evangelidis G, editors. The Effects of Dimensionality Curse in High Dimensional kNN Search. 2011 15th Panhellenic Conference on Informatics; 2011 30 Sept.-2 Oct. 2011.
103. Tseng Y-J, Huang C-E, Wen C-N, Lai P-Y, Wu M-H, Sun Y-C, i sur. Predicting breast cancer metastasis by using serum biomarkers and clinicopathological data with machine learning technologies. *Int J Med Inf.* 2019;128:79-86.
104. Li Q, Yang H, Wang P, Liu X, Lv K, Ye M. XGBoost-based and tumor-immune characterized gene signature for the prediction of metastatic status in breast cancer. *J Transl Med.* 2022;20:177.

10. ŽIVOTOPIS

OPĆI PODATCI:

Ime i prezime: Josip Vrdoljak

Datum rođenja: 28. rujna 1994.

Adresa stanovanja: Pazdigradska 46, 21 000 Split

Kontakt: +385 99 21 51 251

E-mail: josip.vrdoljak@mefst.com

OBRAZOVANJE I OSPOSOBLJAVANJE:

2019. doktor medicine, Medicinski fakultet u Splitu

2020. – danas, polaznik poslijediplomskog doktorskog studija “Klinička medicina utemeljena na dokazima”, Medicinski fakultet u Splitu

2022. – danas., MSc. iz umjetne inteligencije (International University of Applied Sciences, Berlin, Njemačka)

RADNO ISKUSTVO

2019. – 2020., sekundarijat, KBC Split

2019. – 2022., vanjski suradnik, Katedra za patofiziologiju, Medicinski fakultet Sveučilišta u Splitu

2020. – 2022., specijalizant, Odjel za endokrinologiju, KBC Split

2022 – danas, asistent, Katedra za patofiziologiju, Medicinski fakultet Sveučilišta u Splitu

NASTAVNA AKTIVNOST

2019.- danas, predmet Patofiziologija (Studij medicine, Studij dentalne medicine, Studij farmacije, Studij medicine na engleskom jeziku, Zdravstveni studiji), Medicinski fakultet Sveučilišta u Splitu

ČLANSTVA U ORGANIZACIJAMA I POVJERENSTVIMA:

2019. - danas, član Hrvatske liječničke komore

RAČUNALNE VJEŠTINE

2022. Certifikat iz podaktovne znanosti (DataCamp)

2023. AI in Healthcare Nanodegree, Udacity

NAGRADE I PRIZNANJA

2018. Dekanova nagrada

STRANI JEZICI

engleski jezik (napredno)

POPIS PUBLIKACIJA

1. Vrdoljak J, Vilović M, Živković PM, Tadin Hadjina I, Rušić D, Bukić J i sur. Mediterranean Diet Adherence and Dietary Attitudes in Patients with Inflammatory Bowel Disease. *Nutrients*. 2020; 12:3429.
2. Vrdoljak J, Krešo A, Kumrić M, Martinović D, Cvitković I, Grahovac M, Vickov J, Bukić J, Božić J. The Role of AI in Breast Cancer Lymph Node Classification: A Comprehensive Review. *Cancers*. 2023; 15:2400.
3. Vrdoljak J, Kumric M, Vilovic M, Martinovic D, Tomic IJ, Krnic M i sur. Effects of Olive Oil and Its Components on Intestinal Inflammation and Inflammatory Bowel Disease. *Nutrients*. 2022; 14:757.

4. Martinovic D, Tokic D, Martinovic L, Vilovic M, Vrdoljak J, Kumric M i sur. Adherence to Mediterranean Diet and Tendency to Orthorexia Nervosa in Professional Athletes. *Nutrients*. 2022; 14:237.
5. Grahovac M, Kumric M, Vilovic M, Martinovic D, Kreso A i sur. Adherence to Mediterranean diet and advanced glycation endproducts in patients with diabetes. *World J Diabetes*. 2021; 12:1942-1956.
6. Martinovic D, Tokic D, Martinovic L, Kumric M, Vilovic M, Rusic D i sur. Adherence to the Mediterranean Diet and Its Association with the Level of Physical Activity in Fitness Center Users: Croatian-Based Study. *Nutrients*. 2021; 13:4038.
7. Vrdoljak J, Kumric M, Ticinovic Kurir T, Males I, Martinovic D i sur. Effects of Wine Components in Inflammatory Bowel Diseases. *Molecules*. 2021; 26:5891.
8. Vrdoljak J, Kumric M, Vilovic M, Martinovic D, Rogosic V, Borovac JA i sur. Can Fasting Curb the Metabolic Syndrome Epidemic? *Nutrients*. 2022; 14:456.
9. Martinovic D, Tokic D, Martinovic L, Kumric M, Vrdoljak J, Bozic J. Orthorexia nervosa and its association with narcissism in fitness center users. *Eat Weight Disord*. 2022; 27:2155–163.
10. Vrdoljak J, Boban T, Petrić Miše B, Boraska Jelavić T, Bajić Ž, Tomić S, Vrdoljak E, Efficacy and safety of TC dose-dense chemotherapy as first-line treatment of epithelial ovarian cancer: a single-institution retrospective cohort study, *Japanese Journal of Clinical Oncology*. 2019; 49: 347–53,
11. Skrabic R, Kumric M, Vrdoljak J, Rusic D, Skrabic I, Vilovic M, Martinovic D, Duplancic V, Ticinovic Kurir T, Bozic J. SGLT2 Inhibitors in Chronic Kidney Disease: From Mechanisms to Clinical Practice. *Biomedicines*. 2022; 10:2458.
12. Vrdoljak J, Boban Z, Barić D, Šegvić D, Kumrić M, Avirović M, Perić Balja M, Periša MM, Tomasović Č, Tomić S, Vrdoljak E, Božić J. Applying Explainable Machine

-
- Learning Models for Detection of Breast Cancer Lymph Node Metastasis in Patients Eligible for Neoadjuvant Treatment. *Cancers*. 2023; 15:634.
13. Tokic D, Mikacic M, Kumric M, Ticinovic Kurir T, Rancic I, Martinovic D, Bukic J, Vrdoljak J, Lizatovic IK, Stipic SS, Supe Domic D, Bozic J. Association between Brain Injury Markers and Testosterone in Critically-Ill COVID-19 Male Patients. *Microorganisms*. 2022; 10:2095.
14. Urlic H, Kumric M, Vrdoljak J, Martinovic D, Dujic G, Vilovic M, Ticinovic Kurir T, Bozic J. Role of Echocardiography in Diabetic Cardiomyopathy: From Mechanisms to Clinical Practice. *Journal of Cardiovascular Development and Disease*. 2023; 10:46.
15. Kumric M, Bozic J, Dujic G, Vrdoljak J, Dujic Z. Chronic Effects of Effective Oral Cannabidiol Delivery on 24-h Ambulatory Blood Pressure and Vascular Outcomes in Treated and Untreated Hypertension (HYPER-H21-4): Study Protocol for a Randomized, Placebo-Controlled, and Crossover Study. *Journal of Personalized Medicine*. 2022; 12:1037.
16. Kljakovic-Gaspic T, Tokic D, Martinovic D, Kumric M, Supe-Domic D, Stojanovic Stipic S, Delic N, Vrdoljak J, Vilovic M, Ticinovic Kurir T, Bozic J. Prognostic Value of Catestatin in Severe COVID-19: An ICU-Based Study. *Journal of Clinical Medicine*. 2022; 11:4496.
17. Kumric M, Zivkovic PM, Ticinovic Kurir T, Vrdoljak J, Vilovic M, Martinovic D, Bratanic A, Lizatovic IK, Bozic J. Role of B-Cell Activating Factor (BAFF) in Inflammatory Bowel Disease. *Diagnostics*. 2022; 12:45.
18. Kumric M, Vrdoljak J, Dujic G, Supe-Domic D, Ticinovic Kurir T, Dujic Z, Bozic J. Serum Catestatin Levels Correlate with Ambulatory Blood Pressure and Indices of Arterial Stiffness in Patients with Primary Hypertension. *Biomolecules*. 2022; 12:1204.

11. DODATAK

Priloženi kod i aplikacija modela mogu se pronaći na sljedećem linku:

- <https://github.com/vrda23/Applying-Explainable-Machine-Learning-Models-for-Detection-of-Breast-Cancer-Lymph-Node-Metastasis>

Također prilažem screenshot Streamlit aplikacije preko koje se model može koristiti u istraživačke i edukacijske svrhe:

The screenshot shows the user interface of the 'Lymph Node Status Classifier' application. It features a 'Submit' button at the top left. Below it are several input fields and sliders for patient data: 'Age' (slider at 50), 'Tumour size (cm)' (input field at 2.90), 'Tumour grade' (dropdown menu at 3), 'ER status (%)' (slider at 50), 'PR status (%)' (slider at 50), 'HER-2 status' (dropdown menu at 0), 'Ki67 (%)' (slider at 20), and 'Histological type' (dropdown menu at 'Ca with medullary features').

Lymph Node Status Classifier

(for patients eligible for neoadjuvant treatment)

Random forest model optimized and validated as described in <https://pubmed.ncbi.nlm.nih.gov/36765592/>

NOTE: App is for experimental and educational purposes ONLY!

Enter patient details in the sidebar to predict lymph node status.

🔗 Prediction

Lymph node status: Positive

🔗 Prediction Probability

Probability of positive lymph node status: 65.36%

Made with Streamlit