

Inovacije procesa za kružno gospodarstvo primjenom umjetne inteligencije (AI)

Želimir Kurtanjek¹

¹Sveučilište u Zagrebu Prehrambeno-biotehnološki fakultet, Pierottijeva ulica 6, 10000 Zagreb

Sažetak: Razvoj i primjena umjetne inteligencije (AI) je sustavski prisutna u svim oblicima društvenog života, tehnologijama i znanosti. Usprkos zabrinutosti, očekuje se da će primjena AI imati najvažniji učinak u rješavanju globalnih problema u razvoju kružnog gospodarstva. Motivacija ovog rada je pokazati potencijalnu mogućnost AI modela zasnovanih na spoznaji kauzalnih veza na primjerima novih tehnologija ekstrakcije biološki aktivnih molekula, zaštiti okoliša i razvoju novih materijala. Prikazani su rezultati modela umjetne inteligencije primjenom načela strukturnog kauzalnog modela (SCM) Bayes-ove mreže (BN) fuzijom temeljnih znanja i podatkovnih baza. Kauzalna analiza posljedica mogućih intervencije u tehnološkom procesu provodi se transformacijom BN mreže primjenom d-separacije BN mreža. Prikazani su rezultati kauzalnih AI modela za unaprijeđenje "zelene" tehnologije ekstrakcije biomolekula, biorafinerije, razgradnje tekstilnih otpadnih voda i oporabe građevinskog otpada.

Ključne riječi: umjetna inteligencija, Bayes-ova mreža, kauzalnost, molekularni deskriptori

1. Uvod

Razvoj i primjena umjetne inteligencije (AI) je danas prisutna u svim oblicima društvenog života, tehnologijama i znanosti. Nedavna popularizacija i otvorenost pristupu velikom jezičnom modelu, koji je generativno unaprijedno obučena neuronska transformer mreža (LLM GPTChat), izazvala je veliku globalnu pažnju i zabrinutost. Usprkos opravdanoj zabrinutosti, očekuje se da primjena umjetne inteligencije ima najvažniji učinak u rješavanju globalnih problema kao što su razvoj kružnog gospodarstva, smanjenje globalnog zagrijavanja, digitalna agronomija i proizvodnja hrane, te medicina i razvoj suvremenih sustava edukacije [1-2]. Sadašnji stadij AI zasnovan na masivnom sustavu pretraživanja digitalnih baza podataka ima veliki otisak ugljičnog dioksida procijenjen na 4 g CO₂ emisije po pitanju. To je posljedica algoritma LLM zasnovanog na statističkom (agnostičkom) učenju pretraživanjem. Namjera

ovog rada je pokazati potencijalnu mogućnost AI modela zasnovanih na spoznaji kauzalnih veza na primjerima novih tehnologija ekstrakcije biološki aktivnih molekula [3], razvoju biorafinerija [4], zaštiti okoliša [5] i razvoju novih materija [6]. Prikazani su rezultati modela umjetne inteligencije primjenom načela strukturnog kauzalnog modela (*Structural Causal Model*, SCM). Temelj SCM je Bayesova mreža (BN) fuzijom apriori znanja i uvjetovanih razdioba vjerojatnosti povezivanjem endogenih i egzogenih varijabli povezanih usmjerenim rubovima u obliku necikličkog grafa (*Directed Acyclic Graph*, DAG). Kauzalna analiza je posljedica odluke ili intervencije inženjera u tehnološkom procesu, a provodi se transformacijom BN mreže zamjenom slučajnog karaktera uzroka s determinističim modelom i određivanje podskupa mreže za sprječavanje interefencijskih utjecaja. Ovdje su prikazani rezultati kauzalnog AI modela za unapređenje “zelene” tehnologije ekstrakcije biomolekula u farmaceutici i nutricionizmu [3]. Primjenjeni su molekularni deskriptori i stabla odučivanja za funkcionalno povezivanje. Ista metodologija je primjenjena za analizu i potencijalno unapređenje tehnologije biorafinerije i biološkog procesa obrade otpadnih voda tekstilne industrije [4-5]. Također je prikazan kauzalni AI model kakvoće betonskih smjesa [6]. Procjenjeno je da proizvodnja cementa doprinosi 8% globalne emisije CO₂, a istraživanje novih kompozitnih formulacija cementa može značajno doprinjeti razvoju kružnog gospodarstva i smanjenju onečišćenja okoliša.

2. Kauzalni modeli umjetne inteligencije

Temelj razvoja kauzalnih modela $P(Y|X, \theta)$ umjetne inteligencije je načelo Bayes-ove statistike kojom se najizgledniji model (posteriori) procjenjuje prethodnim informacijama i podacima (prior), j.d. (1). Model je Bayes-ova mreža kauzalnih zavisnosti egzogenih faktora (ulazne veličine), endogenih faktora (veličina stanja X), svrhovitosti modela (izlazna veličina Y) i parametara (θ).

$$P(\theta|Y, X) = (P(Y, X|\theta) \cdot P(\theta)) / P(Y, X) \quad (1)$$

Grafički prikaz modela je mreža kauzalnih relacija $G = \{V, E\}$ prikazana kao usmjereni aciklički graf (DAG) određen čvorovima $V = \{X, Y\}$ i usmjerenim neposrednim kauzalnim poveznicama $\{E\}$. DAG graf je Markov model sa svojstvom jednostavne dekompozicije gustoće vjerojatnosti kao produkt uvjetovanih vjerojatnosti čvorova X_i i pripadajućih neposrednih uzroka $par(X_i)$.

$$P(X) = \prod_{i=1}^N P(X_i | par(X_i)) \quad (2)$$

Mjera vjerodostojnosti je Bayesov kriterij informacije, BIC , funkcija je strukture grafa G , broja parametara modela K , podataka N i vjerodostojnost modela podataka L :

$$BIC(X|G) = K \cdot \ln(N) - 2 \cdot \ln(L(X|G)) \quad (3)$$

Određivanje strukture mreže G kauzalnih veza je ključni korak inferencije najvjerojatnijeg modela. Najčešće se koristi statistički PC algoritam uvjetnih nezavisnih veza ili numerički postupak minimiziranja BIC informacijskog kriterija:

$$\hat{G} = \underset{i}{\min} [BIC(X|G_i)] \quad (4)$$

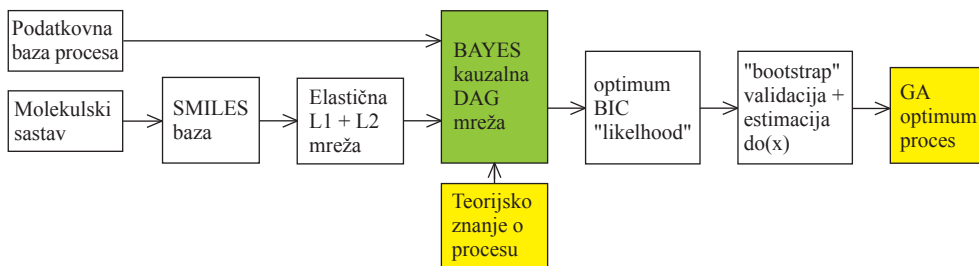
Za određivanje neposrednog ili posrednog kauzalnog učinaka pojedinog egzogenog ili endogenog faktora potrebno je odrediti podskup Z za odvajanje (d -separacija) utjecaja interferirajućih faktora. Određivanjem Z skupa procjenjuje se gustoća vjerojatnosti P posljedica djelovanja $do(X = x)$ promjenom uzroka X na Y ,

$$P(Y|do(X)) = \sum_{z \in Z} P(Y|X=x, Z=z)P(Z=z) = P(Y|X, Z) \quad (5)$$

Model gustoće vjerojatnosti posljedica djelovanja promjene $do(X = x)$ moguće je odrediti strojnim učenjem s Bayes stohastičkom neuronskom mrežom $BNN(Y, X)$. Kao kvantitativna mjera, ATE $y(x)$, uzročno posljedične veze uzima se derivacija očekivane vrijednosti E :

$$ATE[y(x)] = \frac{d}{dx} E_z [(Y|do(X))] \quad (6)$$

Na Slici 1 prikazan je dijagram toka razvoja kauzalnog modela umjetne inteligencije, validacije modela i optimiranje kemijskih i biotehnoloških procesa primjenom genetičkog algoritma. Osnova modela su integracija baze eksperimentalnih podataka i temeljnih inženjerskog znanja (fizike, kemije, biologije), određivanje mreže kauzalnih relacija (DAG graf) i postupaka dubokog učenja Bayes-ovom neuronskom mrežom (BNN). Razvoj modela započinje povezivanjem s iscrpnom višedimenzionalnom bazom procesnih podataka i bazom molekulskih deskriptora. Kemijske formule se transformiraju u velike SMILES (2D ili 3D) baze numeričkih podataka. Svaka molekula je određena velikim brojem N međusobno povezanih numeričkih deskriptora, $N > 5000$ za 2D strukturne značajke. Velika dimenzija prostora pretraživanja molekulskih deskriptora se modelom ciljane regularizacije (određena svrhovitost sustava) s elastičnom mrežom (kombinacijom L1 i L2) svede na desetak ključnih deskriptora.



Slika 1: Dijagram toka razvoja kauzalnog AI modela industrijskih kemijsko inženjerskih i biotehnoloških proizvodnih sustava

Mjerene procesne varijable i ključni molekularni deskriptori određuju čvorove Bayes-ove DAG kauzalne mreže. Teorijsko znanje o procesu apriori definira poznate neposredne kauzalne veze između čvorova DAG mreže. Nove kauzalne veze su slučajne veličine koje se određuju PC algoritmom i/ili optimiranjem Bayes-ov informacijskog kriterija BIC izraženim kompleksnošću mreže i pripadajućom vjerodostojnosti modela podacima (*“likelihood”*). Eksperimentalni podaci i apriori poznate kauzalne veze ne određuju jednoznačno najvjerodostojniji model i potrebno je primijeniti ekstenzivan postupak validacije. Primjenjuje se postupak višestrukog ponavljanja uzorkovanja (*“bootstrap”*) i statističkim odlučivanjem o signifikantnosti pojedinih uzročno posljedičnih veza. Za procjenu kauzalnih posljedica, $do(x)$, potencijalnih intervencija egzogenih faktora provodi se postupak eliminacije d -separacijom DAG mreže uklanjanjem veza između endogenih faktora koji djeluju interferirajuće istovremeno na uzrok i posljedicu. Kauzalne funkcionalne veze $Y(X)$ su stohastičke i najčešće nelinearne a gustoće vjerojatnosti $P(Y|do(x))$ modeliraju se postupkom strojnog učenja s Bayes-ovom neuronskom mrežom. Za usporedbu kauzalnih učinaka više faktora primjenjuje se linearna parametrizacija prosječnog učinka („ATE average treatment effect“). Najvažnija primjena modela kauzalne umjetne inteligencije tehničkih sustava je u mogućnosti genetičkog algoritma GA optimiranja nelinearnih kauzalnih veza i kompleksnih ograničenja.

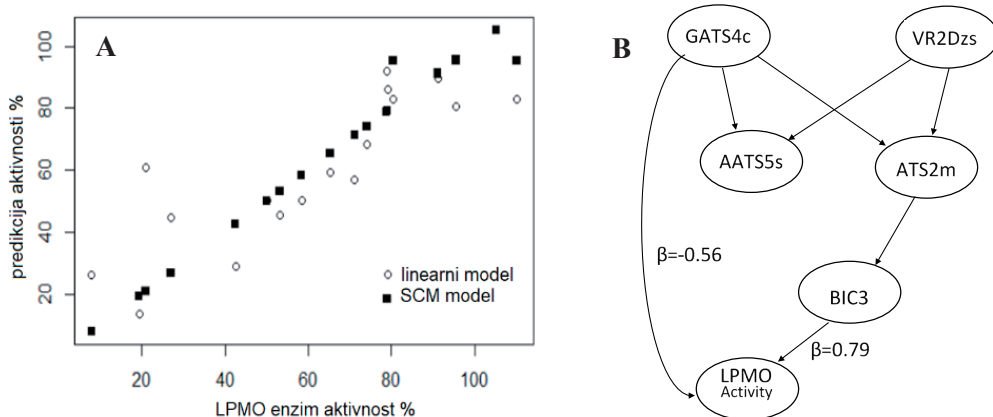
3. Rezultati i diskusija

Zahvaljujući sustavskom konceptu kauzalne umjetne inteligencije moguće je navesti veliki broj ključnih primjera potencijalnih razvoja novih industrijskih tehnologija u sustavu cirkularnih gospodarskih sustava i dekarbonizacije prirodnog okoliša. Među najvažnijim primjerima su razvoj visokotemperaturnih supravodljivih materijala i fizijskih generatora energije. U ovom prikazu pokazani su rezultati mogućnosti kauzalnih modela umjetne inteligencije za unaprijeđenje katalitičkog procesa enzimske obrade lignoceluloznih materijala, zelene ekstrakcije biomolekula eutektičkim smje-

sama, primjena enzima litičke polisaharidne monoooksigenaze (LPMO) u razgradnji boja tekstilnih otpadnih voda, i primjeni prirodnih agregata i građevinskih otpadnih materijala. Za modeliranje su primijenjeni računalni paketi u otvorenom pristupu [7-8].

3.1 Biorafinerija

Jedna od potencijalnih novih tehnologija je primjena enzima litičke polisaharidne mono-oksigenaze (LPMO) u pripremi lignoceluloznih sirovina za proizvodnju energenata i aktivnih tvari u biorafinerijama. Razvijen je strukturni kauzalni model (SCM) umjetne inteligencije za predikciju aktivnosti enzima zavisno o molekulskim deskriptorima pojedinih supstrata [4]. Za pojedine supstrate određene su SMILES formule i pripadajućih 5400 molekulskih deskriptora. Primijenjena je elastična mreža L1 + L2 za regularizaciju prostora deskriptora. Utvrđeno je da se 95 % aktivnosti LPMO enzima može predvidjeti samo s 5 osnovnih deskriptora (Slika 2). Važnost nelinearnih interakcija deskriptora prikazana je usporedbom točnosti predikcije enzimske aktivnosti strukturnim SCM modelom ansamblom stabala odlučivanja i višestrukim linearnim modelom. Prikaz na Slici 2A jasno pokazuje važnost nelinearnog kauzalnog učinka.

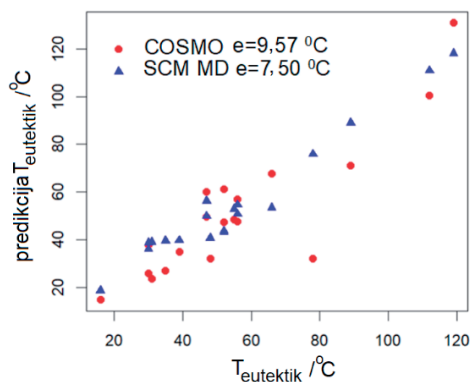


Slika 2: (A) Usporedba predikcije aktivnosti LPMO enzima primjenom linearnog višestrukog modela i strukturnog kauzalnog modela (SCM), (B) prikaz usmjerenog acikličkog grafa (DAG) s ključnim molekularnim deskriptorima [4]

Kauzalna povezanost učinaka pojedinih deskriptora može se odrediti iz DAG prikaza na Slici 2B. Najvažniji pozitivni prosječni učinak ($ATE = 0,79$) je deskriptora BIC3 (Bond informacijski indeks 3 stupnja). Najveći inhibirajući (negativni $ATE = -0,56$) ima GATS4c (2D Geary koeficijent 4. stupnja zadržke s Gasteiger težinskim koeficijentima naboja).

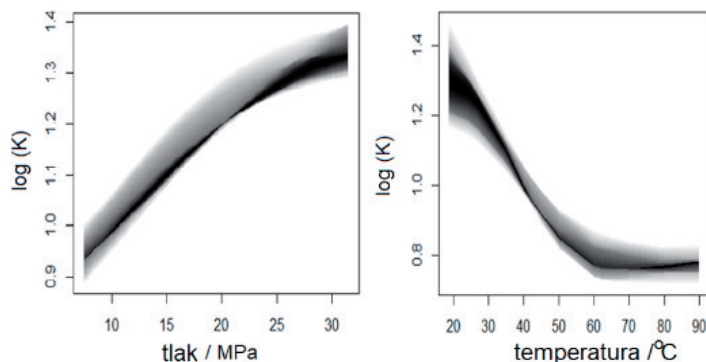
3.2 Zelena ekstrakcija NADES i superkritičnim CO₂-H₂O sustavima

Primjena niskotemperaturnih i netoksičnih ekstrakcijskih tehnologija ima ključnu važnost za razvoj održivih i kružnih procesa u prehrambenoj i farmaceutskoj industriji. Zbog svojstava kao što su nezapaljivost, biorazgradivost i nehlapljivost, veliku pozornost kao zelena otapala privlače prirodna eutektična otapala (*Natural Deep Eutectic Solvents*, NADES). Uglavnom su to smjese nabijenog akceptora i nabijenog donora vodika koji su međusobno povezani vodikovim vezama. Matematičko modeliranje je vrlo složeno i primjenjuju se načela kvantne mehanike, kao što je dostupno u komercijalnom računalnom programskom paketu COSMO. Cilj ovog istraživanja bilo je primijeniti molekulske deskriptore i strukturni model kauzalne umjetne inteligencije (SCM MD) za predikciju temperature eutektika NADES sustava i optimirati proces zelene ekstrakcije superkritičnim ugljičnim dioksidom [3]. Model je razvijen na podacima 37 industrijski važnih molekula. Primjenom modela, elastičnom mrežom (L1 + L2) određena su 4 najvažnija molekulska deskriptora (SC.5, chi klaster 5-tog stupnja; nH.1, broj vodikovih atoma; BIC4, Bond informacijski indeks 4-tog stupnja; CIC4, komplementarni informacijski indeks 4-tog stupnja simetrije susjedstva) za 95% točnosti predikcije. DAG kauzalna struktura modela određena je statističkim PC algoritmom, a učenje funkcionalnih zavisnosti strojnim učenjem s ansamblom stabala odlučivanja. Usporedbom točnosti predikcije (Slika 3) s COSMO računalnim paketom jasno pokazuje točniju predikciju i praktičnu prednost SCM DM modela kauzalne umjetne inteligencije.



Slika 3: Usporedba predikcije temperature eutektika kvantno-mehaničkim modelom COSMO i strukturno kauzalnim modelom umjetne inteligencije temeljem molekularskih deskriptora SCM MD [3]

Za razliku od NADES sustava koji su predmet znanstvenog istraživanja, primjena ekstrakcije superkritičnim CO₂-H₂O ima primjenu u industrijskim procesima. Cilj ovog istraživanja bio je optimirati procesne parametre, tlak i temperaturu, primjenom SCM DM modeliranja.



Slika 4: Predikcije Bayes-ovim neuronskim mrežama modela (BNN) razdioba uvjetnih gustoća vjerojatnosti kauzalnih učinaka $P(Y|do(X=x))$ tlaka (A) i temperature (B) na koeficijent particije K [3]

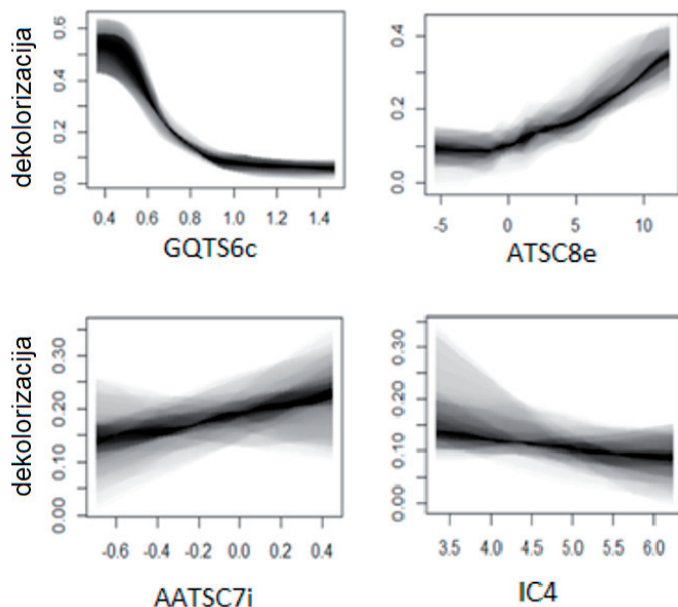
Model je razvijen na podatkovnoj literaturnoj bazi s 471 uzorkom koeficijenta particije objavljenih u zadnjih 30 godina u području temperatura 10-90 °C i tlaka 7-35 MPa. Za 95% točnosti predikcije logaritma koeficijenta particije, $\log(K)$, određeni su molekularni deskriptori (MolLogP, Moriguchi indeks; MPC, minimalni parcijalni naboj; PEOE_VSA10, VSA indeks naboja; MaxESTale značajka). Kauzalna povezanost SCM molekularskih deskriptora i procesnih parametara određena je PC algoritmom.

Temeljem SCM DM modela primjenom Bayes-ovih neuronskih mreža određene su $P(\log(K)|T, p)$ uvjetne raspodjele gustoće vjerojatnosti zavisnosti particije o procesnim parametrima. Predikcije su prikazane na Slici 4 i rezultate treba kontekstualno vezati za molekule iz primijenjene baze podataka. Za optimiranje tehnološkog procesa, osim prikazanog parametra termodinamičke ravnoteže, potrebno je uzeti u obzir kinetiku prijenosa tvari u procesu ekstrakcije.

3.3 Dekolorizacija tekstilnih otpadnih voda

Onečišćenje okoliša otpadnim vodama tekstilne industrije predstavlja globalni problem održivosti upotrebe postojećih sirovina. U ovom radu je istražena mogućnost upotrebe modela umjetne inteligencije za predikciju razgradnje (dekolorizacije) primjenom LPMO enzima [5]. Laboratorijska eksperimentalna istraživanja provedena su na skupu od 59 najvažnijih bojila. Praćenje procesa dekolorizacije provedeno je optički mjerenjem apsorpcije na valnoj duljini od 280 nm. Intenzitet dekolorizacije određen je linearnom regresijom apsorpcije i vremena. Ključni molekularni deskriptori određeni su elastičnom mrežom LASSO algoritma. Za 95% točnosti predikcije intenziteta dekolorizacije određen je skup od 6 ključni deskriptora: GATS6c-2D Geary koeficijent s zadržkom 6 i Gasteiger težinama naboja, ATSC6e-2D i ATSC8e-2D Broto-Moreau koeficijentima autokorelacije sa zadržkama od 6 i 8 i težinskim

koeficijentima elektronegativnosti, AATSC3v – srednji centralni Moreau-Broto koeficijent autokorelacije s zadržkom 3 i Van der Waals volumnim težinskim koeficijentima, AATSC7i srednji centralni Moreau-Broto autokorelacije sa zadržkom 7, i IC4-4 informacijskim koeficijentom okružja. Na Slici 5 prikazane su gustoće vjerojatnosti kauzalne zavisnosti intenziteta dekolorizacije o pojedinim ključnim molekulskim deskriptorima. Poznavanje kauzalnosti potencijalno omogućuje sintezu novih molekula s boljim učinkom razgradnje LPMO enzimom.

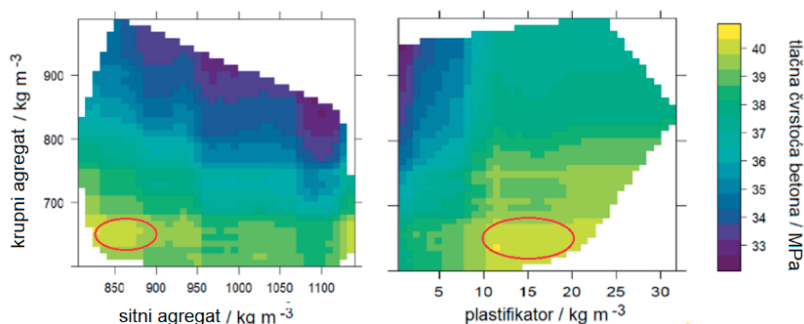


Slika 5: Predikcije Bayes-ovim neuronskim mrežama (BNN) modela razdioba uvjetnih gustoća vjerojatnosti $P(Y|do(X=x))$ kauzalnih učinaka glavnih molekulskih deskriptora na dekolorizaciju otpadnih voda tekstilne industrije [5]

3.4 Oporaba građevinskog otpada

Proizvodnja cementa je industrija s velikim ugljičnim otiskom na globalni prirodni okoliš. Intenzitet emisije ugljika smanjuje se inovacijom novih kompozitnih materijala, upotrebom prirodnih agregata i oporabljenim građevinskim otpadom u betonskoj smjesi. U radu je istražena mogućnost primjene genetičkog algoritma i modela umjetne inteligencije za predikciju tlačne čvrstoće betona zavisno od sastava betonske mješavine [6]. Razvijen je kauzalni AI model temeljem baze podataka s $n = 1030$ uzoraka, i $p = 9$ varijabli sastava (cement, šljunak, lebdeći pepeo, voda, plastifikator, grubi agregat, sitni agregat, vrijeme, i tlačna čvrstoća betona). Razvijeni model je usmjereni aciklički graf (DAG) određen heurističkim postupkom optimiranja Bayesovog informacijskog kriterija (BIC). AI model dobiven strojnim učenjem omogućuje

predikciju tlačne čvrstoće betona s prosječnom apsolutnom pogreškom 3 MPa (4,3 %) u odnosu na pogrešku od 10 MPa za višestruki linearni model. Značajno unapređenje točnosti AI modela ukazuje na važnost nelinearne kauzalne interakcije između komponenata betonske mješavine. Na Slici 6 prikazani su parcijalni kauzalni učinci sitnog i krupnog agregata te plastifikatora na kvalitetu betona (tlačnu čvrstoću).



Slika 6: Predikcije kauzalnih učinaka udjela agregata (sitnog, krupnog) i plastifikatora na tlačnu čvrstoću betona. Optimalna područja sastava naznačena su crvenom granicom [6]

Primjena kauzalnog AI modela je naročito važna za optimiranje procesa s gledišta kružne ekonomije smanjenjem otiska ugljikom, uporabe građevinskog otpada, te upotrebe prirodnih agregata. Prednosti upotreba modela umjetne inteligencije AI i genetičkog algoritma GA naročito su važne zbog potencijalne raznolikosti kompozitnih mješavina, nelinearnosti interakcija sastava, složenosti gospodarskih ograničenja i promjenljivih tržišnih uvjeta.

4. Zaključak

Razvoj kružnog gospodarstva temelji se na sustavskom pristupu povezivanja postojećih i novih tehnoloških procesa, zaštite okoliša, gospodarske održivosti i unapređenja ljudskih života. Heterogenost i globalna širina ciljeva zahtijevaju informacijsko povezivanje velikog broja podataka iz raznorodnih izvora i primjenu umjetne opće inteligencije (*Artificial General Intelligence*, AGI) za analizu, nadzor i upravljanje sustava. Ključnu ulogu umjetne inteligencije u razvoju kružnog gospodarstva ističu velike globalne razvojne konzultantske kompanije kao što je to primjerice McKinsey [9]. U ovom radu dan je opis povezivanja i primjene algoritma kauzalnog zaključivanja i algoritama strojnog učenja. Prikazani modeli su iz područja kemijskog inženjersva i informacije o molekulskim značajkama povezane su s bazom molekulskih deskriptora. Navedena raznorodnost primjera, fuzija podataka i temeljnih inženjer-

skih znanja naglašava primjenu koncepta kauzalnosti umjetne inteligencije u razvoju inovacija i primjenu.

5. Literatura

- [1] Stanford University, Measuring trends in Artificial Intelligence, <https://aiindex.stanford.edu/report/>
- [2] OECD (2023), Artificial Intelligence in Science: Challenges, Opportunities, and the Future of Research, OECD Publishing, Paris, doi: 10.1787/a8d820bd-en
- [3] Kurtanjek Ž.: Molecule Structure Causal Modelling (SCM) of Choline Chloride Based Eutectic Solvents, *Chemical and Biochemical Engineering Quarterly*, **36** (2022) 4, 223-230, doi: 10.15255/CABEQ.2022.2104
- [4] Rezić T., Vrsalović Prosečki A., Kurtanjek Ž.: New approach to the evaluation of lignocellulose derived by-products impact on lytic-polysaccharide monooxygenase activity by using molecular descriptor structural causality model, *Bioresource Technology*, **342** (2021) 125990, doi: 10.1016/j.biortech.2021.125990
- [5] Rezić I., Kracher D., Oros D., Mujadžić S., Andelini M., Kurtanjek Ž., Ludwig R., Rezić, T.: Application of Causality Modelling for Prediction of Molecular Properties for Textile Dyes Degradation by LPMO, *Molecules*, **27** (2022) 6390, doi: 10.3390/molecules27196390
- [6] Kurtanjek Ž.: Kauzalni AI model kvalitete betonske mješavine, *Kemija u industriji* (2024) u tisku
- [7] dowhy 0.11.1 <https://pypi.org/project/dowhy/>
- [8] R Core Team (2023). *_R: A Language and Environment for Statistical Computing_*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>
- [9] McKinsey, Artificial intelligence and the circular economy: AI as a tool to accelerate the transition, <https://www.mckinsey.com/capabilities/sustainability/our-insights/artificiainelligence-and-the-circular-economy-ai-as-a-tool-to-accelerate-the-transition>