



PRIOPĆENJE ZA MEDIJE:

Petra Buljević Zdjelarević, Ured za odnose s javnošću

Institut Ruđer Bošković

T. +385 (1) 457-1269, (99) 267-95-14

E. info@irb.hr W. www.irb.hr

Zagreb, 1. veljače 2017.

Ruđerovci su razvili novu računalnu metodu za brzo i samostalno čitanje znanstvene literature

Ovim istraživanjem znanstvenici su usavršili računalne statističke tehnike koje mogu u samo nekoliko minuta 'pročitati' i 'razumjeti' tisuće tekstova u dovoljnoj mjeri da iz njih prepoznaju značajke živih organizama, za koje bi inače istraživačima trebale godine iščitavanja literature.

Znanstvenici Zavoda za elektroniku Instituta Ruđer Bošković (IRB), Maria Brbić, dr. sc. Fran Supek i suradnici, objavili su znanstveni rad u kojem su opisali tehnologije umjetne inteligencije koje samostalno 'čitaju' znanstvenu literaturu iz područja mikrobiologije. Rad je objavljen u uglednom časopisu Nucleic Acids Research. Riječ je o časopisu koji spada među šest posto najboljih časopisa u kategoriji, a koji pruža informacije o najnovijim dostignućima u molekularnoj biologiji.

"Naš tim je stvorio algoritme koji uče kako raspoznavati značajke različitih vrsta bakterija kroz analize tekstova Wikipedije, studentskih radova, te stručnih izvora. Važnost ovakvih algoritama je velika zato jer je obim znanstvene literature, ali i sveukupnog sadržaja na internetu, sve veći i veći, i istraživači teško mogu pratiti sve nove informacije koje se pojavljuju." – objašnjava dr. sc **Fran Supek**.

Primjerice, razvijeni računalni algoritmi mogu iz pojavljivanja ključnih riječi pouzdano zaključiti preživljava li bakterija visoke temperature ili zračenje, stvara li otporne spore, živi li u određenom okolišu (more, rijeke, tlo) te izaziva li bolesti biljaka, životinja ili čovjeka. Dodatno, ruđerovci su razvili i novi način analize teksta 'bez nadzora', gdje umjetna inteligencija kombinira razne tekstove kako bi opisala nove, smislene značajke koje prije nisu bile sistematično popisane. Primjerice, sistematizirana je grupa bakterija koja uzrokuje kvarenje piva ili onih koje žive u pročišćivačima otpadnih voda.

Osim 'čitanja' teksta, **istraživači IRB-a su naučili umjetnu inteligenciju da na sličan način 'čita' i genetski zapis različitih bakterija**. Tako i na temelju genoma (skupa svih gena) jednog organizma može predvidjeti kakve značajke će taj mikrob imati.

Ovaj rad otkriva da i poredak gena na kromosomima, koji se znatno razlikuje između živih organizama, vrlo dobro odražava mnoge njihove značajke. Primjerice, mikrobi koji stvaraju



spore i time preživljavaju teške uvjete čak i stotinama godina, pokazuju vrlo karakteristična 'genska susjedstva' koja nisu prisutna u mikrobima koji ne mogu stvoriti spore.

Dakle, algoritmi umjetne inteligencije 'čitaju' i tekst znanstvenih radova, i genski zapis, a znanstvenici IRB-a su istražili i način da se ta dva izvora i povežu. Tako su razvijene statističke metode koje asociraju pojedine gene sa raznim značajkama, i time je pronađeno mnogo gena koji potencijalno igraju ulogu npr. u stvaranju spora, u otpornosti na kemikalije ili u izazivanju bolesti raznih organa kod čovjeka.

"Ovim istraživanjem usavršili smo računalne statističke tehnike koje mogu u nekoliko minuta 'pročitati' i 'razumjeti' tisuće tekstova u dovoljnoj mjeri da iz njih prepoznaju značajke živih organizama. Sve zajedno, umjetna inteligencija je razmatrala sve kombinacije 424 raznovrsne značajke i 3046 bakterijskih vrsta. To je preko milijun kombinacija! Za provjeru svake od njih bi čovjeku-istraživaču trebale godine iščitavanja literature. Naši algoritmi će sa lakoćom obrađivati i nadolazeću literaturu u budućnosti, i automatski je povezati sa genskim zapisom organizama." – zaključuje dr. Supek.

Rad je nastao u okviru 'Future and Emerging Technologies' (FET) projekta Europske unije pod nazivom MAESTRA (<http://maestra-project.eu/>), te uključuje suradnju sa grupom dr. Anite Kriško sa Mediteranskog Instituta za istraživanje života u Splitu.

Rezultati računalnih predviđanja su slobodno dostupni preko Web baze podataka <http://protraits.irb.hr/>, koju je izradio Matija Piškorec s Instituta Ruđer Bošković.

Rad pod naslovom 'The landscape of microbial phenotypic traits and associated genes' je objavljen u uglednom časopisu Nucleic Acids Research koji izdaje Oxford University Press (faktor odjeka 9.2; [doi:10.1093/nar/gkw964](https://doi.org/10.1093/nar/gkw964)). Za ovo istraživanje, znanstvena novakinja IRB-a Maria Brbić nagrađena je godišnjom nagradom Zaklade Branimir Jernej (<http://www.zaklada-branimir-jerneji.hr/>).

KORISNE POVEZNICE:

RAD: <https://academic.oup.com/nar/article/44/21/10074/2290929/The-landscape-of-microbial-phenotypic-traits-and>