

Naslov HR: Prediktivno modeliranje godišnjih stopa nataliteta i mortaliteta u Republici Hrvatskoj na temelju demografskih i ekonomskih varijabli

Naslov EN: Predictive modeling of the annual natality and mortality rates in the Republic of Croatia based on demographic and economic variables

Proučiti izabranu literaturu o primjenama suvremenih metoda strojnog učenja u predviđanju demografskih varijabli kao što su natalitet, fertilitet, mortalitet, broj stanovništva itd. Na temelju javno dostupnih izvora, prikupiti podatkovni skup koji sadrži stope nataliteta i mortaliteta u Republici Hrvatskoj po godinama, kao i podatke o nizu potencijalno relevantnih demografskih i ekonomskih prediktorskih varijabli. Definirati arhitekturu i metodologiju treninga za nekoliko izabranih modela strojnog učenja u svrhu predviđanja stopa nataliteta i mortaliteta na temelju predloženih prediktorskih varijabli. Eksploratornim i statističkim analizama utvrditi koje prediktorske varijable predstavljaju najjačnije pojedinačne indikatore stopa nataliteta i mortaliteta. Provesti usporedno vrednovanje naučenih modela primjenom uobičajenih evaluacijskih metrika za regresijske modele strojnog učenja. Analizom interpretabilnosti modela odrediti koji prediktori utječu a koji ne utječu na stope nataliteta i mortaliteta te usporediti dobivene zaključke s rezultatima eksploratornih i statističkih analiza.

Relevantne napomene i informacije:

- Spajanje i slaganje (obrada) podataka iz različitih izvora
- Na temelju literature odrediti kako podatke prirediti za model.
- Napraviti prikaz analize podataka i dobivenih predikcija.
- Metoda za interpretabilnost modela je primjerice SHAP

Potencijalno relevantna literatura:

- Boonen, T. J., & Li, H. (2017). Modeling and forecasting mortality with economic growth: A multipopulation approach. *Demography*, 54(5), 1921-1946.
- Turner, N., & Robbins, K. (2023). Association Between County-Level Natality and Income in the US, 2000-2020. *JAMA pediatrics*, 177(2), 198-202.
- Leasure, D. R., Jochem, W. C., Weber, E. M., Seaman, V., & Tatem, A. J. (2020). National population mapping from sparse survey data: A hierarchical Bayesian modeling framework to account for uncertainty. *Proceedings of the National Academy of Sciences*, 117(39), 24173-24179.

- Şahinarslan, F. V., Tekin, A. T., & Çebi, F. (2021). Application of machine learning algorithms for population forecasting. *International Journal of Data Science*, 6(4), 257-270.