

Luka Šimek

e-mail: simek.luka@gmail.com
telefon: +385 92 5028 177
website: <https://tinjano.github.io>
GitHub: <https://github.com/tinjano>



Sažetak

Matematičar po obrazovanju i godinama entuzijazma. Usredotočen na statistiku i širi pojam *data sciencea*. Raznoliki skup teoretskih znanja i praktičnih vještina s jakim temeljima. Sposoban za razne zadatke, od prikupljanja podataka iz otvorenih izvora, preko vizualizacije do inferencijalnog i prediktivnog modeliranja.

Obrazovanje

- 2019.-2022.** Preddiplomski studij matematike, Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet
prosjek ocjena: 4.22
izborni kolegiji: Građa računala, Baze podataka, Markovljevi lanci, Objektno programiranje (C++)
demonstrator iz Matematičke analize (1 i 2) i Programiranja (1 i 2)
- 2015.-2019.** XV. gimnazija, Zagreb
brončana medalja na Srednjoeuropskoj matematičkoj olimpijadi (MEMO) 2017.
drugo mjesto na Državnom natjecanju iz matematike 2016.
sudjelovanje na državnim natjecanjima iz matematike, informatike, logike i fizike

Znanje i vještine

programski jezici	Python web scraping: BeautifulSoup, selectolax, requests, HTTPX, Playwright znanstvene biblioteke: NumPy, SymPy, SciPy, statsmodels, PyMC, pandas, Polars, scikit-learn, TensorFlow, transformers vizualizacija: Matplotlib, seaborn iskustvo u: SQL, R, C, C++
matematičko znanje	sistematsko i intuitivno znanje analize, linearne algebre, diskretne (kombinatorne) matematike, teorije vjerojatnosti, statistike i statističkog (strojnog) učenja
opće informatičko znanje	znanje o algoritmima i strukturama podataka, iskustvo u dizajnu i implementaciji algoritama, osnovno znanje asinkronog programiranja i računalnih mreža
ostalo	JupyterLab, LaTeX, alati Linuxa, git, opće vještine rješavanja problema, precizna i stručna komunikacija na hrvatskom i engleskom jeziku

Odabrani projekti

[essay-quality](#). Inspiriran natjecanjem na Kaggle-u, cilj ovog projekta bio je predvidjeti ocjene eseja na temelju *event logova*. Izračunamo veliki broj prediktora pomoću biblioteke Polars, vizualiziramo njihovu važnost i distribucije te kreiramo dva modela: klasifikator temeljen na multinomijalnoj logističkoj regresiji i *boosted decision tree* regresor. Zatim razmatramo dodatne mogućnosti multifraktalnom analizom (MFDFA) raznih vremenskih nizova. Biblioteke: Polars, scikit-learn, XGBoost

[ad1](#). Vizualizacija reklamiranja na Google-u u Hrvatskoj. Skupljamo podatke koristeći BigQuery i razmatramo razne parametre i trendove, stvarajući na kraju kvalitetne grafove spremne za prezentaciju. Biblioteke: pandas, Matplotlib.

[fastdice](#). Istražujemo donošenje odluka u stohastičkom okruženju na primjeru bacanja kocaka sličnom kao u igri Jamb. Inzistirajući na brzim vektoriziranim implementacijama, kreiramo potklasu tenzora za Monte Carlo simulacije, klasu koja predstavlja Markovljev lanac i na kraju alate za računanje optimalnih strategija u Markovljevim procesima odlučivanja s dovoljno malim prostorom stanja i radnji. Biblioteke: NumPy.