

RUDARENJE PODATAKA

Povijesno gledano, pojmu pronalaženja korisnih uzoraka u podacima dani su različiti nazivi uključujući rudarenje podataka (*data mining*), izdvajanje znanja, otkrivanje informacija i dr. Terminom rudarenje podataka najčešće se koriste statističari, analitičari podataka i stručnjaci za upravljanje informacijskim sustavima. Sintagma otkrivanje znanja u bazama podataka (*knowledge discovery in databases – KDD*) predložena je 1989. *KDD* se odnosi na sveukupni proces otkrivanja korisnog znanja iz podataka, a rudarenje podataka je jedan korak u tom procesu. Rudarenje podataka je primjena specifičnih algoritama za izvlačenje uzoraka iz podataka. Dodatni koraci u procesu otkrivanja znanja – priprema, odabir i čišćenje podataka, ugradnja odgovarajućeg prethodnog znanja i pravilno tumačenje rezultata rudarenja neophodni su da bi se osiguralo korisno znanje iz podataka (Fayyad i dr. 1996).

Hand i dr. (2001) definiraju rudarenje podataka kao znanost izdvajanja korisnih informacija iz velikih skupova podataka ili baza podataka. To je nova disciplina koja objedinjuje statistiku, strojno učenje, upravljanje podacima i bazama podataka, prepoznavanje uzoraka, umjetnu inteligenciju i druga područja.

Rudarenje podataka ima primjenu u različitim područjima pa i u raznim granama geodezije. Kartograflanje zemljišnog pokrova primjenom satelitskih snimaka mukotrpan je zadatak radi li se ručno. Automatske metode su brže i preciznije, a nedostatak im je nemogućnost zanemarivanja neželjenih objekata i teškoće u radu sa snimkama visoke rezolucije. Navedene se teškoće mogu prevladati rudarenjem podataka (Cucu-Dumitrescu i dr. 2010). Andrienko i dr. (2001) pokazali su da interaktivna vizualizacija i metode rudarenja podataka mogu djelovati kao komplementarni instrumenti analize podataka. Njihova integracija podržava iterativni proces istraživačke analize podataka (*exploratory data analysis*).

Literatura

- Andrienko, N., Andrienko, G., Savinov, A., Voss, H., Wettschereck, D. (2001): *Exploratory Analysis of Spatial Data Using Interactive Maps and Data Mining*, Cartography and Geographic Information Science, 3, 151–165, <https://www.tandfonline.com/toc/tcag20/28/3?nav=tocList>, (28. 2. 2018.).
- Cucu-Dumitrescu, C., Serban, F., Buican, M., Datcu, M. (2010): *Support for Automation of Cartography Based on Earth Observation Images – A Data Mining Approach*, Proceedings of ESA Living Planet Symposium, 28 June – 2 July, 2010, Bergen, Norway, <http://adsabs.harvard.edu/abs/2010ESASP686E.476C>, (28. 2. 2018.).
- Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996): *From Data Mining to Knowledge Discovery in Databases*, AI Magazine, Vol. 17, 3, 37–54.
- Hand, D., Mannila, H., Smyth, P. (2001): *Principles of data mining*, The MIT Press, Cambridge, Massachusetts, London.

Nedjeljko Frančula