

# Életkor prediktálása gépi tanulással

**Témavezető:** Kerepesi Csaba

(MTA SZTAKI, tudományos munkatárs, kerepesi@sztaki.hu)

Érdekes kérdés, milyen pontosan lehet egy egyén életkorát megbecsülni rá vonatkozó mérési adatokból gépi tanulási módszerekkel. Korábban prediktáltak már életkort például blogbejegyzésekből 4,14 év [1], arcképekből 5,78 év [2], vérképadatokból 5,55 év [3], MRI fevételekből 4,16 év [4] átlagos abszolút hibával. Az alkalmazások szempontjából rendkívül hasznos lenne olyan mérhető jellemzőket és gépi tanulási módszert találni, amellyel ezeknél jelentősen kisebb hibával lehet prediktálni. E kérdéskör kutatásában vesz részt a jelölt. Két irány közül választhat. Az egyik irány, hogy egy adott publikus adathalmazon végzett legjobb életkorpredikciót (pl. az[1]-ben kévőt) próbálja megjavítani olyan módszerekkel amit korábban az adott adathalmazra még nem használtak. A másik irány, hogy keres olyan publikus adathalmazokat, amelyhez van csatolva életkor, de még nem végeztek rajta életkorpredikciót. Ígéretes lenne egy olyan adathalmazt, amelyben az egyénekről többféle típusú adat áll rendelkezésre. Lehet akár nem emberre vonatkozó adatokkal (pl. egyéb élőlény, gép, operációs rendszer, stb...) dolgozni, ha vannak csatolva élettartam adatok és érdekes vizsgálni az entitás életkorát.

## A jelentkezővel szemben támasztott elvárások:

- Önálló munka (beleértve a szükséges adatbányászati és gépi tanulási ismeretek önálló elsajátítását)
- Programozási tapasztalat az adatelemzéshez (preferált nyelv: Python)
- Angol nyelvismeret cikkek alapos tanulmányozásához

## Kiindulási irodalom:

[1] Nguyen, D., Smith, N.A. and Rosé, C.P., 2011, June. Author age prediction from text using linear regression. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 115-123). Association for Computational Linguistics.

[2] Nguyen, D.T., Cho, S.R., Shin, K.Y., Bang, J.W. and Park, K.R., 2014. Comparative study of human age estimation with or without preclassification of gender and facial expression. *The Scientific World Journal*, 2014.

[3] Putin, E., Mamoshina, P., Aliper, A., Korzinkin, M., Moskalev, A., Kolosov, A., Ostrovskiy, A., Cantor, C., Vijg, J. and Zhavoronkov, A., 2016. Deep biomarkers of human aging: application of deep neural networks to biomarker development. *Aging (Albany NY)*, 8(5), p.1021.

[4] Cole, J.H., Poudel, R.P., Tsagkrasoulis, D., Caan, M.W., Steves, C., Spector, T.D. and Montana, G., 2017. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, 163, pp.115-124.

