

## MEGHÍVOTT PLENÁRIS ELŐADÓ A MTNE 22. FÓRUMÁN



**BODÓ ZALÁN** a Babeş–Bolyai Tudományegyetem Matematika és Informatika Karának dékánhelyettese, a Magyar Matematika és Informatika Intézet docense. Egyetemi (BSc, MSc) tanulmányait a BBTE Matematika és Informatika Karán végezte, 2009-ben ugyanitt védte meg *Semi-supervised learning with kernels* című doktori dolgozatát. 2008-ban tanársegédként csatlakozott a kar oktatói gárdájához, 2011-ben adjunktussá, 2017-ben pedig docenssé lépett elő. Tudományos kutatásai a mesterséges intelligenciához kapcsolódnak, pontosabban a gépi tanulási módszerekhez, a még szűkebb területek közül pedig megemlíthetjük a természetesnyelv-feldolgozást, a félig felügyelt tanulási módszereket, a hashing algoritmusokat, illetve a gépi látást. A kutatás mellett fontosnak tartja a színvonalas egyetemi oktatást, illetve a hallgatók mihamarabbi bevonását a kutatási tevékenységekbe.

A Magyar Tudomány Napja Erdélyben 22. fórumán elhangzó plenáris előadása:

**Bodó Zalán (informatikus, egyetemi docens, Babeş–Bolyai Tudományegyetem, Matematika és Informatika Kar, Magyar Matematika és Informatika Intézet, Kolozsvár): *Az értelmezhetőség minőségének mérése interpretálható konvolúciós neurális hálókbán***

A gépi tanuló modellek értelmezhetősége nagyon sok alkalmazás esetében fontos kérdés, például biztonságkritikus rendszerekben, ahol hibás döntés esetén az ok-okozati elemzéshez elengedhetetlen a predikció magyarázata. Például önvezető autók esetében, ahhoz, hogy megtaláljuk egy hibás döntés okát vagy okait, hogy ezt javítani tudjuk a rendszerben, értenünk kell, hogy hogyan jutottunk az adott döntéshez. Míg egyszerűbb modellek esetében (pl. logisztikus regresszió, szupport vektor gépek, döntési fák stb.) az értelmezhetőség könnyen megvalósítható, mély neurális hálók esetében ez közel sem egyértelmű. Habár osztályozási feladatok esetében például az osztályozási pontosság az egyik legfontosabb kiértékelési mérték, interpretálható modellek esetében az értelmezhetőség minőségének mérése is fontos. E rövid előadásban ilyen módszereket mutatok be és hasonlítok össze egymással.