

No 'melnās kastes' – uz 'stikla kasti': izskaidrojamais mākslīgais intelekts

Specseminārs

Pasniedzējs: Maksims Ivanovs, M.Sc. kognitīvajās zinātnēs, doktorants (LU DF un EDI)

(maksims.ivanovs@edi.lv)

Kopsavilkums. Mākslīgā intelekta metodes, it īpaši mašīnmācīšanās metodes, arvien vairāk ietekmē zinātņi, uzņēmējdarbības vidi un sabiedrību kopumā. Tās tiek izmantotas, lai atpazītu patoloģijas pazīmes mikroskopijas attēlos un sociālu lietotājus viņu kopīgajās bildēs, lai padarītu aizraujošākas datorspēles un izlemtu, vai hipotekārā kredīta pieteikums tiks noraidīts vai apstiprināts. Taču jo plašāku pielietojumu gūst mašīnmācīšanās, jo svarīgāks kļūst jautājums, cik lielā mērā mēs saprotam tās metodes un līdz ar to – cik lielā mērā varam uz tām paļauties. Šis jautājums ir īpaši aktuāls dziļo mākslīgo neironu tīklu gadījumā, jo tās darbojas kā 'melnās kastes', kuru funkcionēšana ir izprasta tikai daļēji un nav caurspīdīga. Līdz ar to pašlaik ir problemātiski izmantot ļoti efektīvus mākslīgo neironu tīklu modeļus daudzu kritiski svarīgu uzdevumu risināšanā, piemēram, pašbraucošo auto vadīšanai vai lēmumu par pacienta ārstēšanu pieņemšanai, jo gan no praktiskā, gan no juridiskā skatpunkta būtu riskanti uzticēt tos kontroles sistēmai, kuru mēs izprotam tikai daļēji. Lai risinātu mākslīgā intelekta necaurspīdīguma problēmu un pārvērstu to no 'melnās kastes' par 'stikla kasti', aktīvi notiek pētījumi izskaidrojamā mākslīgā intelekta (angl. *explainable artificial intelligence* jeb *XAI*) jomā. Šī semināra mērķis ir iepazīstināt studentus ar jaunākajām un nozīmīgākajām publikācijām šajā nozarē un dažādām pieejām mākslīgā intelekta modeļu izskaidrošanai, kā arī gūt ieskatu šo pieeju īstenošanā praksē.

Atslēgas vārdi latviski: mākslīgais intelekts, neironu tīkli, mašīnmācīšanās, izskaidrojamais mākslīgais intelekts.

Atslēgas vārdi angļiski: artificial intelligence, neural networks, machine learning, explainable artificial intelligence, XAI.

Semināra organizācija: lekcijas, diskusijas, studentu prezentācijas, programmēšanas projekts vai referāts.

Laiks un vieta: ceturtdienās no 16:30 līdz 18:00 12. auditorijā. Pirmā nodarbība notiks 17. septembrī. Daļa nodarbību var notikt tiešsaistē.

Pieteikšanās: rakstot uz maksims.ivanovs@edi.lv