



ЦЕНТЪР ЗА ОБУЧЕНИЕ – БАН

1000 София
ул. „Сердика“ № 4
<http://edu.bas.bg>

email: tdc-phd@cu.bas.bg
тел.: 02 987 31 67
02 979 52 60

Основна информация:

Име на курса: РАЦИОНАЛНОСТ И ИЗКУСТВЕН ИНТЕЛЕКТ

Лектор: Доц. Борис Грозданов

Телефон: 0887 701 443

Имейл: grozdanoff@ips.bas.bg

Хорариум: 30 учебни часа

Анотация (до 150 думи):

Целта на курса е да въведе докторантите в съвременния интердисциплинарен дебат върху изкуствения интелект (ИИ), който включва изчислителни подходи, символни подходи и най-новите хибридни подходи. Курсът ще изложи историята на проблемното поле на ИИ, ще даде водещите позиции по проблема за природата на рационалността и интелигентността, и ще изложи на достъпен език водещите подходи в практическото разработване на ИИ, като изкуствени невронни мрежи и архитектури на ИИ, базирани на тях и водещите проблеми и предизвикателства пред тях. Ще бъдат изложени и символите подходи към ИИ, както и тяхната най-модерна интеграция в хибридните подходи. Ще бъдат разгледани и редица хардуерни аспекти на архитектурите и обучението на ИИ.

Тематично съдържание на курса (кратко описание по теми или модули):

Тема / Модул 1: Интелигентност: история и водещи теории

Тема / Модул 2: История на ИИ

Тема / Модул 3: Съвременни архитектури на ИИ и техния успех

Форми на обучение и оценяване:

Присъствено

Презентация и курсова работа

.....

Компетентности, придобити в резултат на обучението (3-5 точки):

Докторантите, завършили успешно курса, ще познават историята на възникването на различните ИИ архитектури, техните възможности и успехи и техните ограничения. Ще могат да се ориентират в съвременната ситуация на динамично развитие на нови ИИ модели и да преценяват добре кои архитектури какви задачи могат да изпълняват с какъв успех.

Литература:

R. Sutton and A. Barto (2018) Reinforcement learning: An Introduction (2nd Ed). MIT Press.

D. Silver, et al., (2017) Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm, Deepmind, .



ЦЕНТЪР ЗА ОБУЧЕНИЕ – БАН

1000 София
ул. „Сердика“ № 4
<http://edu.bas.bg>

email: tdc-phd@cu.bas.bg
тел.: 02 987 31 67
02 979 52 60

- J. Schrittwieser, et al. (2020) Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model, Deepmind, .
- Allen, C., G. Varner, and J. Zinser (2000) Prolegomena to any future artificial moral agent. Journal of Experimental & Theoretical Artificial Intelligence 12 (3): 251 – 261.
- Anderson, M., and S. Anderson (2007) Machine ethics: Creating an ethical intelligent agent. AI Magazine 28 (4): 15 – 26 .
- Beavers, A. (2001) Kant and the problem of ethical metaphysics. Philosophy in the Contemporary World 7 (2): 47 – 56 .
- Beavers, A., ed. (2010) Robot ethics and human ethics. Special issue of Ethics and Information Technology 12 (3).
- Floridi, L., and J. Sanders (2004) On the morality of artificial agents. Minds and Machines 14 (3): 349 - 379.
- Gips, J. (1995) Towards the ethical robot . In Android Epistemology, ed. K. Ford, C. Glymour, and P. Hayes, 243 – 252. Cambridge, MA : MIT Press .
- Grozdanoff, Popov, Serafimova, Eds. (2023) Rationality and Ethics in Artificial Intelligence, Cambridge Scholars Publishing.pdf
- Grozdanoff, Boris D. (2020) «Prospects for a Computational Approach to Savulescu's Artificial Moral Advisor», in Ethical Studies, vol. 5/book 3, 2020, pp. 121–140.
- Grozdanoff, Boris D. (2021) «Graph Formalism for normative ethical tasks in Artificial Intelligence» in Ethical Studies, vol. 6/book 2, 2021, pp. 109–120.
- Kurzweil, R. (1990) The Age of Intelligent Machines . Cambridge, MA : MIT Press .
- Mittelstadt et al. (2018) The ethics of algorithms: Mapping the debate, in Big Data and Society, Volume: 3 issue: 2; <https://doi.org/10.1177/2053951716679679>.
- Moor, J. (2006) The nature, importance, and difficulty of machine ethics. IEEE Intelligent Systems 1541 – 1672 : 18 – 21 .
- Powers, T. (2006) «Prospects for a Kantian machine». IEEE Intelligent Systems 1541 – 1672 : 46 – 51
- Savulescu J., Maslen H. (2015) Moral Enhancement and Artificial Intelligence: Moral AI?. In: Romportl J., Zackova E., Kelemen J. (eds) «Beyond Artificial Intelligence. Topics in Intelligent Engineering and Informatics», vol 9. Springer, Cham. https://doi.org/10.1007/978-3-319-09668-1_6
- Todorova, M (2020) The Artificial Intelligence: A Brief history of its development and the ethical aspects of the problem, Sofia, East-West Publishing, in Bulgarian.
- Turing, A. (1950) Computing machinery and intelligence. Mind 59 (236): 433 – 460.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.

Допълнителна информация (по желание) (например: специални изисквания, лабораторно оборудване, предварителни знания):

Курсът се води на английски език, докторантите трябва да владеят английски на академично ниво (разбиране на текст, разбиране на говор и писане на английски).

.....
.....