

ТЕОРЕТИЧНИ ОСНОВИ НА АЛГОРИТМИТЕ ЗА ИЗКУСТВЕН ИНТЕЛЕКТ

Николай Николов¹

¹ Икономически университет – Варна/Катедра „Информатика“, Варна, България,
nikolay.nikolov@ue-varna.bg

РЕЗЮМЕ

Настоящият доклад изследва теоретичните принципи, които стоят в основата на алгоритмите за изкуствен интелект (ИИ). Фокусът е поставен върху математическите и алгоритмичните модели, които позволяват на машините да придобиват, представят и използват знания. Разглеждат се концепции от линейна алгебра, теория на вероятностите и статистика, които са критични за разбирането на машинното обучение и дълбоките невронни мрежи. Анализират се различни методи за обучение, включително контролирано, неконтролирано и усилващо обучение. Докладът също така обсъжда ролята на алгоритмичната сложност и изчислителната ефективност при разработката на мащабируеми ИИ системи. В заключение се представят бъдещи направления и изследователски въпроси, свързани с теоретичните аспекти на изкуствения интелект.

КЛЮЧОВИ ДУМИ: изкуствен интелект, алгоритми, машинно обучение, невронни мрежи.

THEORETICAL FOUNDATIONS OF ARTIFICIAL INTELLIGENCE

Nikolay Nikolov¹

¹ Department of Informatics, University of Economics - Varna, Varna, Bulgaria,
nikolay.nikolov@ue-varna.bg

ABSTRACT

This paper explores the theoretical principles underlying artificial intelligence (AI) algorithms. The focus is on the mathematical and algorithmic models that enable machines to acquire, represent, and utilize knowledge. Concepts from linear algebra, probability theory, and statistics, which are critical for understanding machine learning and deep neural networks, are examined. Various learning methods, including supervised, unsupervised, and reinforcement learning, are analyzed. The paper also discusses the role of algorithmic complexity and computational efficiency in the development of scalable AI systems. In conclusion, future directions and research questions related to the theoretical aspects of artificial intelligence are presented.

KEYWORDS: artificial intelligence, algorithms, machine learning, neural networks.

ВЪВЕДЕНИЕ

През последните десетилетия изкуственият интелект (ИИ) премина от теоретична концепция към всеобхватна технология, която влияе върху различни аспекти на съвременния живот. Теоретичните основи на алгоритмите за ИИ са от съществено значение за разбирането на това как интелигентните системи учат, адаптират се и вземат решения. Дълбокото познаване на тези основи позволява на изследователите и практиците да разработват по-стабилни, ефективни и етични ИИ системи.

Областта на ИИ е по своята същност интердисциплинарна, обединявайки концепции от компютърните науки, математиката, статистиката и когнитивните науки. Фундаментални теории като теорията на изчислителното обучение (Valiant 1984), статистическата теория на обучението (Vapnik 1995) и вероятностното разсъждение (Pearl 1988) значително допринесоха за развитието на алгоритмите за ИИ.

С появата на дълбокото обучение ИИ претърпя революция, въвеждайки сложни архитектури на невронни мрежи, способни да учат йерархични представяния (LeCun, Bengio и Hinton 2015). Разбирането на математическите основи на тези модели е от критично значение за напредъка на областта и за справяне с предизвикателства, свързани с интерпретируемостта и обобщаването на моделите.

Настоящият доклад има за цел да изследва теоретичните принципи, които стоят в основата на алгоритмите за изкуствен интелект. Чрез разглеждане на ключови математически модели и алгоритмични стратегии се стреми да обясни как машините придобиват, представят и използват знания. Ще бъдат разгледани концепции от линейна алгебра, теория на вероятностите и оптимизационни методи, които са критични за машинното обучение и невронните мрежи. Освен това ще бъдат анализирани различни парадигми на обучение, включително контролирано, неконтролирано и усилващо обучение, и ще се обсъди ролята на алгоритмичната сложност при разработката на мащабируеми ИИ системи.

1. МАТЕМАТИЧЕСКИ ОСНОВИ НА АЛГОРИТМИТЕ ЗА ИЗКУСТВЕН ИНТЕЛЕКТ

Разбирането на математическите принципи е от съществено значение за разработването и анализирането на алгоритми за изкуствен интелект. Три основни области допринасят за тази основа: линейна алгебра, теория на вероятностите и оптимизационни методи.

1.1. Линейна алгебра

Линейната алгебра предоставя инструменти за представяне и обработка на данни чрез вектори и матрици. Тези структури са ключови при разработката на модели за машинно обучение и невронни мрежи (Strang, 2016). Например, матричните операции се използват за изчисляване на изходите на слоевете в невронните мрежи.

Векторите и матриците са основни елементи за представяне на данни в пространството на характеристиките. Собствените стойности и собствени вектори са критични при намаляване на размерността и анализа на данни чрез методи като анализ на главните компоненти (PCA) (Jolliffe, 2002). Тези концепции позволяват извличане на най-съществената информация от големи и сложни набори от данни.

1.2. Теория на вероятностите и статистика

Теорията на вероятностите позволява моделиране на неопределеността и случайността в данните. Статистическите методи са от съществено значение за оценка на моделите и правене на изводи от данните (Casella и Berger, 2002). Вероятностните разпределения описват как стойностите на случайна променлива се разпределят, което е важно при моделирането на стохастични процеси.

Байесовата теория и вероятностните модели позволяват актуализиране на вероятностите въз основа на нови данни (Murphy, 2012). Това е особено полезно при разработването на адаптивни системи, които се учат в динамични среди.

1.3. Оптимизационни методи

Оптимизационните алгоритми са в основата на процеса на обучение на ИИ моделите. Градиентният спуск се използва за минимизиране на функцията на загубите чрез актуализиране на параметрите в посока на най-стръмния наклон (Ruder, 2016). Стохастичният градиентен спуск (SGD) е вариация, която използва случайни подмножества от данни за ускоряване на обучението, което е особено полезно при големи набори от данни.

2. АЛГОРИТМИ ЗА МАШИННО ОБУЧЕНИЕ

Машинното обучение е основен компонент на изкуствения интелект, който позволява на системите да се учат от данни и да правят прогнози или решения без изрично програмиране (Mitchell, 1997). Съществуват три основни парадигми на обучение: контролирано, неконтролирано и усилващо обучение.

2.1. Контролирано обучение

При контролираното обучение моделите се обучават върху етикетирани данни, където входовете са свързани с изходни стойности. Регресията предвижда непрекъснати стойности, като например линейна регресия, докато класификацията разпределя данни в дискретни категории, като използва методи като логистична регресия и опорни векторни машини (SVM) (Bishop, 2006).

2.2. Неконтролирано обучение

Неконтролираното обучение работи с неетикетирани данни и цели откриване на скрити структури или модели. Клъстеризацията групира подобни данни, като например метода k-средни (Hastie, Tibshirani и Friedman, 2009). Намаляването на размерността, като при PCA, намалява сложността на данните, запазвайки съществената информация.

2.3. Усилващо обучение

Усилващото обучение включва обучение на агент, който взаимодейства с околната среда и се учи от последствията на своите действия чрез награди или наказания (Sutton и Barto, 2018). Този подход е особено ефективен при задачи, където последователностите от решения влияят върху крайната цел.

3. НЕВРОННИ МРЕЖИ, ДЪЛБОКО ОБУЧЕНИЕ, АЛГОРИТМИЧНА СЛОЖНОСТ

Невронните мрежи са модели, вдъхновени от биологичните неврони, които са способни да учат сложни функции и представяния (Goodfellow, Bengio и Courville 2016).

3.1. Основи на невронните мрежи

Изкуствените неврони са основните единици, които приемат входове и произвеждат изходи чрез функция на активация. Многослойните персептрони (MLP) са невронни мрежи с един или повече скрити слоя, позволяващи моделиране на нелинейни зависимости (Hornik, Stinchcombe и White, 1989). Теоремата за универсална апроксимация гласи, че невронните мрежи с достатъчен брой неврони могат да апроксимират произволна непрекъсната функция (Hornik, 1991).

3.2. Архитектури на дълбоки невронни мрежи

Сверточните невронни мрежи (CNN) са специализирани за обработка на изображения, като улавят пространствени зависимости (LeCun и др., 2015). Рекурентните невронни мрежи (RNN) са подходящи за последователни данни като текст и аудио, тъй като вземат предвид времевите зависимости.

3.3. Теоретични аспекти и предизвикателства

Теорема за универсална апроксимация: Невронните мрежи с достатъчен брой неврони могат да апроксимират произволна непрекъсната функция (Hornik 1991).

Предизвикателства при обучението: Проблеми като изчезващия градиент изискват техники като нормализация на партидите (Ioffe и Szegedy 2015).

3.4 Алгоритмична сложност и ефективност

Анализът на алгоритмичната сложност е критичен за оценка на ефективността на ИИ алгоритмите, особено при работа с големи обеми данни. Голямото O обозначение се използва за

описване на производителността на алгоритмите (Cormen и др., 2009). Мащабируемостта се постига чрез използване на паралелни изчисления и оптимизации, което позволява обработка на големи данни в реално време.

4. ЕТИЧНИ И СОЦИАЛНИ АСПЕКТИ. НАПРАВЛЕНИЕ ЗА РАЗВИТИЕ

Развитието на изкуствения интелект повдига важни етични въпроси, свързани с прозрачността, обяснимостта и пристрастията в алгоритмите (Floridi и др. 2018). Необходимостта от създаване на алгоритми, които са справедливи и прозрачни, е от критично значение за изграждането на отговорен ИИ. Въпросите, свързани с поверителността на данните и автоматизацията на работните места, имат значително влияние върху обществото и изискват внимателно разглеждане.

Освен това, разработването на модели, които са обясними и разбираеми за хората, е ключово предизвикателство пред интерпретируемостта на моделите. Изкуственият общ интелект (AGI), който цели създаването на системи, способни да изпълняват всяка интелектуална задача (Russell и Norvig 2020), представлява дългосрочна цел в областта на ИИ. Изследването на потенциала на квантовите компютри за ускоряване на ИИ алгоритми също е обещаващо направление, което може да доведе до значителни подобрения в производителността и ефективността.

ЗАКЛЮЧЕНИЕ

Теоретичните основи на алгоритмите за изкуствен интелект представляват фундаменталната рамка, върху която се изгражда цялата област на ИИ. В този доклад бяха разгледани ключовите математически концепции и алгоритмични стратегии, включително линейна алгебра, теория на вероятностите и оптимизационни методи, които са критични за разбирането и разработката на съвременните ИИ системи. Анализът на различните парадигми на машинно обучение—контролирано, неконтролирано и усилващо обучение—подчертава значимостта на подходящия избор на методология в зависимост от специфичните задачи и наличните данни.

Разглеждането на невронните мрежи и дълбокото обучение показва, че сложните архитектури и алгоритми могат да улавят нелинейни зависимости и да предоставят високи нива на точност в разнообразни приложения. Алгоритмичната сложност и ефективността на ИИ алгоритмите са от съществено значение, особено при работа с големи обеми от данни и необходимостта от реалновремени решения. Изследването на етичните и социални аспекти подчертава важността на отговорния ИИ и необходимостта от справедливи, прозрачни и обясними модели.

В бъдеще подобряването на интерпретируемостта на моделите и разработката на изкуствен общ интелект (AGI) ще бъдат ключови направления в областта на ИИ. Интегрирането на квантовите изчисления обещава значителни подобрения в производителността и способностите на ИИ алгоритмите. Задълбоченото разбиране на теоретичните основи ще продължи да бъде от решаващо значение за напредъка и устойчивото развитие на изкуствения интелект, като гарантира, че технологиите се развиват по начин, който е от полза за обществото.

БЛАГОДАРНОСТИ

Това изследване е направено в рамките на проект НПИ-65/2023 „Изкуственият интелект в помощ на хората с увреждания при осигуряване на дигитална достъпност в процеса на обучение във висшето образование“.

REFERENCES / ИСПОЛЗВАНА ЛИТЕРАТУРА

1. Bishop, C.M. (2006) Pattern Recognition and Machine Learning. Springer.
2. Casella, G. and Berger, R.L. (2002) Statistical Inference. Duxbury Press.
3. Cormen, T.H., Leiserson, C.E., Rivest, R.L. and Stein, C. (2009) Introduction to Algorithms. 3rd ed. MIT Press.
4. Floridi, L., Cowls, J., Beltrametti, M. et al. (2018) 'AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations', *Minds and Machines*, 28(4), pp. 689–707.
5. Goodfellow, I., Bengio, Y. and Courville, A. (2016) Deep Learning. MIT Press.
6. Hastie, T., Tibshirani, R. and Friedman, J. (2009) The Elements of Statistical Learning. 2nd ed. Springer.
7. Hornik, K. (1991) 'Approximation capabilities of multilayer feedforward networks', *Neural Networks*, 4(2), pp. 251–257.
8. Hornik, K., Stinchcombe, M. and White, H. (1989) 'Multilayer feedforward networks are universal approximators', *Neural Networks*, 2(5), pp. 359–366.
9. Ioffe, S. and Szegedy, C. (2015) 'Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift', *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, pp. 448–456.
10. Jolliffe, I.T. (2002) Principal Component Analysis. 2nd ed. Springer.
11. LeCun, Y., Bengio, Y. and Hinton, G. (2015) 'Deep Learning', *Nature*, 521(7553), pp. 436–444.
12. Mitchell, T.M. (1997) Machine Learning. McGraw-Hill.
13. Murphy, K.P. (2012) Machine Learning: A Probabilistic Perspective. MIT Press.
14. Pearl, J. (1988) Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann.
15. Ruder, S. (2016) 'An overview of gradient descent optimization algorithms'
16. Russell, S. and Norvig, P. (2020) Artificial Intelligence: A Modern Approach. 4th ed. Pearson.
17. Strang, G. (2016) Introduction to Linear Algebra. 5th ed. Wellesley-Cambridge Press.
18. Sutton, R.S. and Barto, A.G. (2018) Reinforcement Learning: An Introduction. 2nd ed. MIT Press.
19. Valiant, L.G. (1984) 'A Theory of the Learnable', *Communications of the ACM*, 27(11), pp. 1134–1142.
20. Vapnik, V.N. (1995) The Nature of Statistical Learning Theory. Springer.