

№8 Дәріс. Шешімдерді зерттеу және пайдалану диллемасы (Exploration vs Exploitation Dilemma)



Дәріскер: Даркенбаев Д.К. PhD, қауымдастырылған профессор

Дәрістің мақсаты

Бұл дәрістің мақсаты — күшейтілген оқыту (Reinforcement Learning) саласындағы негізгі мәселелердің бірі болып табылатын зерттеу мен пайдалану диллемасын (exploration vs exploitation dilemma) түсіндіру.



Студенттер агенттің ортада әрекет ету кезінде жаңа мүмкіндіктерді зерттеу мен бұрыннан белгілі тиімді шешімдерді пайдалану арасындағы тепе-теңдікті табудың маңызын меңгереді.

Кілттік сөздер

Күшейтілген оқыту

Зерттеу (exploration)

Пайдалану (exploitation)

Агент

Орта

Саясат (policy)

Тәжірибеден үйрену

Тепе-теңдік

Шешім қабылдау

Стратегия

Дәріс жоспары

01	02
Кіріспе Күшейтілген оқытудың негізгі мақсаты, агент пен ортаның өзара әрекеттесуі, диллема ұғымының мәні	Зерттеу және пайдалану ұғымдары Зерттеу (exploration) деген не, пайдалану (exploitation) деген не, екеуінің өзара байланысы
03	04
Зерттеу мен пайдаланудың диллемасы Агенттің алдында туындайтын шешімдік қайшылық, тәуекел мен табыс арасындағы тепе-теңдік, уақыт пен тәжірибе факторы	Зерттеу мен пайдаланудың маңыздылығы мен мысалдары Күнделікті өмірдегі мысалдар, интеллектуалды жүйелердегі көрінісі, әртүрлі стратегиялар (ε-greedy, softmax және т.б. идеялар)
05	06
Зерттеу мен пайдаланудың тепе-теңдігін табу жолдары Агенттің оқыту процесіндегі рөлі, стратегияларды бейімдеу, тәжірибе жинақтаудың әсері	Қолдану салалары мен практикалық маңызы Өндіріс және басқару жүйелері, робототехника, қаржы және экономикалық модельдер, ойындар мен симуляциялар

7. Қорытынды	8. Бақылау сұрақтары	9. Пайдаланылған әдебиеттер
--------------	----------------------	-----------------------------

Кіріспе

Күшейтілген оқытуда агент қоршаған ортамен өзара әрекеттесе отырып, шешім қабылдайды. Оның басты мақсаты — ұзақ мерзімді марапатты барынша арттыру. Бірақ бұл жолда агент әр сәт сайын екі түрлі таңдау алдында қалады:

	
Зерттеу (exploration) Жаңа мүмкіндіктерді байқап көру, бұрын қолданылмаған әрекеттерді сынау	Пайдалану (exploitation) Бұрын жақсы нәтиже берген, белгілі әрекетті қайталау арқылы пайда табу

Міне, осы екі бағыттың арасындағы қайшылық **«зерттеу мен пайдалану диллемасы»** деп аталады. Бұл диллема — тек машиналық оқытуда емес, адам өміріндегі кез келген шешім қабылдау процесінде де маңызды рөл атқарады.

Мысалы, адам жаңа жұмысты іздеуде немесе таныс мейрамханадан тамақ істеуде де дәл осы диллемамен бетпе-бет келеді: жаңа нәрсені байқап көру (зерттеу) немесе бұрынғысын жалғастыру (пайдалану)?

Зерттеу және пайдалану ұғымдары

Зерттеу (exploration)

Агенттің жаңа немесе аз зерттелген әрекеттерді орындауы. Мақсаты — орта туралы көбірек мәлімет алу және ұзақ мерзімді перспективада тиімді шешім табу.

❏ **Мысалы:** робот белгісіз жолмен жүру арқылы жаңа аймақтарды таниды.

Пайдалану (exploitation)

Агенттің бұрынғы тәжірибесіне сүйеніп, ең пайдалы әрекетті қайталап орындауы.

❏ **Мысалы:** бір жол бұрын жақсы нәтиже берген болса, агент сол бағытты қайта таңдайды.



Бұл екі тәсілдің үйлесімі агенттің оқыту сапасын анықтайды.

- Егер агент тек пайдалану жолымен жүрсе — ол жаңа мүмкіндіктерді байқамай, дамудан қалады.
- Керісінше, егер ол тек зерттеумен айналысса — тәжірибесі тұрақсыз болып, нақты жетістікке жетпейді.

Зерттеу мен пайдаланудың диллемасы

Зерттеу мен пайдалану — бір-бірін толықтыратын, бірақ көбіне қарама-қарсы бағыттар. Агенттің алдында әрбір сәтте шешім тұрады: жаңа әрекетті байқап көру ме, әлде бұрынғы тиімді әрекетті қайталау ма?



Уақыт факторы

Бұл диллема уақыт өте келе күрделене түседі



Мәлімет жинақтау

Агент орта туралы көбірек мәлімет алған сайын пайдалану тиімді бола бастайды



Тым ерте пайдалану

Болашақтағы мүмкіндіктерді шектеуі мүмкін

Практикалық мысал

Мысалы, интернет-дүкен ұсынатын тауарды таңдау жүйесін қарастырайық.



Жүйе қолданушының ұнатқан тауарларын пайдалана отырып, белгілі ұсыныстар береді (**пайдалану**), бірақ кейде жаңа өнімдерді де ұсынуы керек (**зерттеу**), өйткені олар әлеуетті пайдалы болуы мүмкін.

Диллеманы шешудің тәсілдері

Зерттеу мен пайдалану арасындағы тепе-теңдікті табу — күшейтілген оқыту алгоритмдерінің негізгі міндеті. Ол үшін әртүрлі стратегиялар қолданылады:

1

Тепе-тең стратегиялар

Агент белгілі бір ықтималдықпен жаңа әрекеттерді зерттеп отырады (мысалы, ϵ -greedy тәсілінде кездейсоқ зерттеу жасалады)

2

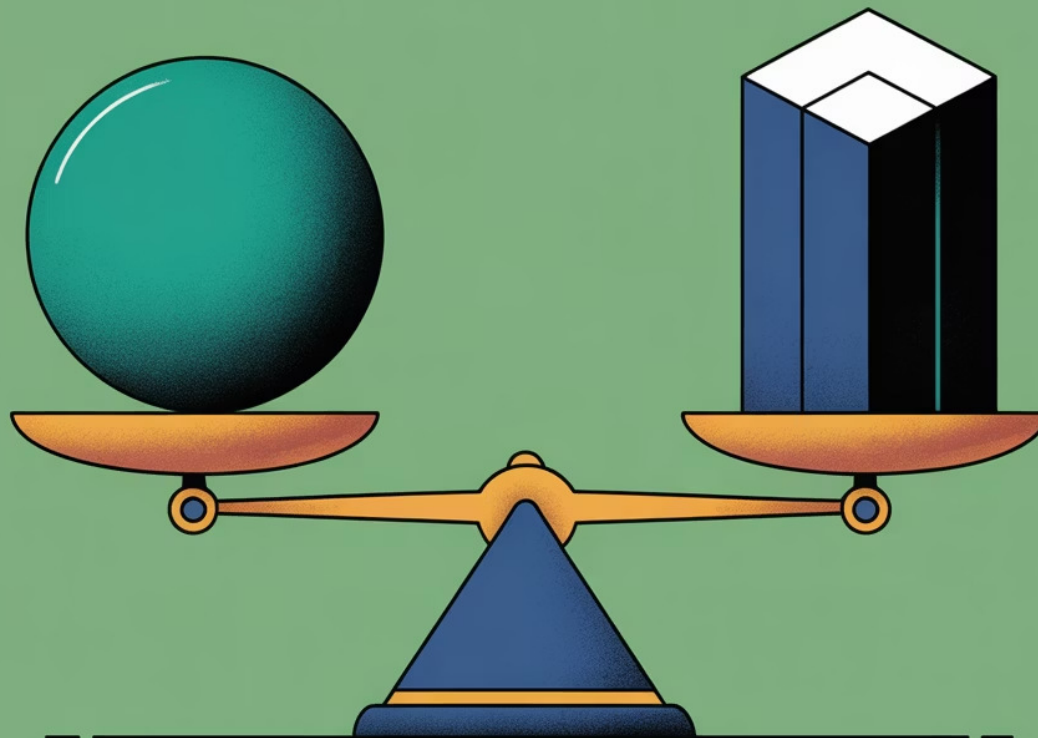
Уақытқа байланысты бейімделу

Бастапқыда зерттеу деңгейі жоғары болады, ал уақыт өте келе пайдалану жиілейді

3

Болашақ марапатты болжау

Агент қысқа мерзімді емес, ұзақ мерзімді табысқа ұмтылады



Бұл тәсілдердің барлығы агенттің үйрену сапасын жақсартып, тиімді шешім қабылдауға мүмкіндік береді.

Қолдану салалары

Зерттеу мен пайдалану диллемасы көптеген нақты қолданбаларда кездеседі:



Робототехникада

Робот белгісіз аймақты зерттей отырып, ең қысқа жолды табады.



Қаржы саласында

Инвестор жаңа активтерді зерттеу мен белгілі қауіпсіз салымдарды пайдалану арасындағы баланс табады.



Деректерді талдау жүйелерінде

Алгоритм жаңа гипотезаларды тексеру немесе дәлелденген үлгіні қолдану арасында таңдау жасайды.



Ойындарда

Ойыншы жаңа стратегияны сынай ма, әлде бұрынғы жеңіске жеткізген тәсілді қолдана ма деген шешім қабылдайды.



Зерттеу мен пайдалану диллемасы — күшейтілген оқыту жүйелеріндегі ең өзекті мәселелердің бірі

Ол агенттің үйрену сапасын, икемділігін және ұзақ мерзімді тиімділігін анықтайды.

Тиімді тепе-теңдік

Агент ортадан ең пайдалы тәжірибе жинай алады

Жаңа мүмкіндіктер

Дер кезінде пайдаланып, өз саясатын жетілдіреді

Қазіргі заманғы жасанды интеллект жүйелерінің жетістіктері — дәл осы диллеманы тиімді шешудің арқасында мүмкін болды.

Бақылау сұрақтары

- 1 Зерттеу мен пайдалану деген ұғымдар нені білдіреді?
- 2 Бұл диллема күшейтілген оқытуда неліктен маңызды?
- 3 Агент неге тек пайдалану стратегиясымен үйрене алмайды?
- 4 Зерттеу мен пайдалану тепе-теңдігін қалай табуға болады?
- 5 Күнделікті өмірден зерттеу мен пайдалану диллемасына мысал келтіріңіз.
- 6 Бұл диллеманы шешудің қандай әдістері бар?
- 7 Неліктен зерттеу тым көп болса, оқыту процесі тиімсіз болуы мүмкін?
- 8 Бұл диллеманың нақты қолданылатын салалары қандай?

Пайдаланылған әдебиеттер

1. Sutton, R. S., & Barto, A. G. Reinforcement Learning: An Introduction. MIT Press, 2018.

2. Kaelbling, L. P., Littman, M. L., & Moore, A. W. Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research, 1996.

3. Russell, S., & Norvig, P. Artificial Intelligence: A Modern Approach. Pearson, 2021.

4. Busoniu, L., Babuska, R., De Schutter, B., & Ernst, D. Reinforcement Learning and Dynamic Programming Using Function Approximators. CRC Press, 2010.

5. Silver, D. Lecture Notes on Reinforcement Learning. University College London, 2020.

6. Li, L., Szepesvári, C., & Schuurmans, D. A Unified Framework for Exploration in Reinforcement Learning. ICML, 2012.

