

№13 Дәріс. Терең нықталандырып оқыту (Deep Reinforcement Learning)

Дәрістің мақсаты:

Терең нықталандырумен оқытудың (Deep Reinforcement Learning, DRL) негізгі қағидаларын, оның классикалық нықталандырумен оқыту (Reinforcement Learning) мен терең нейрондық желілерді біріктіру ерекшеліктерін түсіндіру. Сонымен қатар, бұл әдістің қазіргі заманғы жасанды интеллект жүйелеріндегі қолдану салаларын талдау.

Кілт сөздер:

Терең нықталандырумен оқыту, агент, орта, күйлер, әрекеттер, марапаттау сигналы, нейрондық желі, Deep Q-Network, саясатты оңтайландыру, тәжірибе реплей буфері, тұрақтандыру әдістері.

Дәріс жоспары:

1. Терең нықталандырумен оқытуға кіріспе
2. Нықталандырумен оқыту мен терең оқытудың байланысы
3. DRL жүйесінің жалпы құрылымы
4. Deep Q-Learning алгоритмі және оның ерекшеліктері
5. Саясатты тікелей оңтайландыру әдістері
6. Қолдану салалары мен мысалдар
7. Артықшылықтары мен қиындықтары
8. Қорытынды

Кіріспе

Қазіргі уақытта жасанды интеллект (AI) саласында нықталандырумен оқыту (RL) мен терең оқытудың (Deep Learning) бірігуі ерекше қарқынмен дамуда. Бұл бағыт терең нықталандырумен оқыту (Deep Reinforcement Learning, DRL) деп аталады. DRL адамның тәжірибеден үйрену қабілетін модельдеу арқылы интеллектуалды агенттерге қоршаған ортадан ақпарат алып, тиімді шешім қабылдауға мүмкіндік береді.

DRL-дің басты мақсаты — агентті күрделі, көпөлшемді орталарда тиімді әрекет етуге үйрету. Бұл әдіс дәстүрлі RL тәсілдерінің шектеулерін жеңіп, бейсызық тәуелділіктер мен үлкен деректер көлемін өңдеуге мүмкіндік береді. Терең нейрондық желілер агенттің саясат функциясын, мән функциясын немесе екісінің комбинациясын жуықтауға қолданылады.

1. Терең нықталандырумен оқытуға кіріспе

Терең нықталандырумен оқыту — агенттің қоршаған ортамен өзара әрекеттесу процесі негізінде үйренуін сипаттайтын әдіс. Агент әрбір қадамы үшін марапат (немесе жаза) алады және осы ақпарат арқылы болашақ әрекеттерін жақсартады. Мұнда классикалық RL-дегі күй мен әрекеттердің арасындағы байланыстарды сипаттау үшін терең нейрондық желілер қолданылады.

2. Нықталандырумен оқыту мен терең оқытудың байланысы

RL агенттің шешім қабылдау процесін сипаттайды, ал Deep Learning күрделі деректерден белгілерді (features) автоматты түрде бөліп алуға мүмкіндік береді. DRL осы екі бағытты біріктіру арқылы бейнелер, дыбыс немесе көпөлшемді сенсорлық ақпарат негізінде тікелей шешім қабылдауға жағдай жасайды.

3. DRL жүйесінің жалпы құрылымы

DRL жүйесі үш негізгі элементтен тұрады:

- Агент (Agent): шешім қабылдайтын жүйе.
- Орта (Environment): агент әрекет ететін кеңістік.
- Марапаттау сигналы (Reward): агенттің әрекетінің тиімділігін көрсететін кері байланыс.

Агент ортамен өзара әрекеттесу кезінде күйді қабылдайды, әрекет таңдайды және марапат алады. Бұл процесс бірнеше рет қайталанып, тәжірибе жинақтау арқылы агенттің саясаты жетілдіріледі.

4. Deep Q-Learning алгоритмі және оның ерекшеліктері

Deep Q-Learning (DQN) — DRL-дің алғашқы әрі ең танымал түрлерінің бірі. Мұнда нейрондық желі Q-функцияны жуықтау үшін қолданылады. Бұл желі агентке күй мен әрекет арасындағы тәуелділікті үйренуге мүмкіндік береді.

DQN-нің негізгі ерекшеліктері:

- Тәжірибе реплей буфері (Experience Replay): агент өткен тәжірибелерін есте сақтап, оларды қайта қолдану арқылы оқыту процесін тұрақтандырады.
- Мақсатты желі (Target Network): негізгі желінің көшірмесін пайдалану арқылы оқытудағы ауытқуларды азайтады.

5. Саясатты тікелей оңтайландыру әдістері

Кейбір DRL әдістері мән функциясын емес, агенттің саясатын (policy) тікелей үйретеді. Бұл тәсілдер "policy gradient" немесе "actor-critic" әдістеріне негізделеді. Мұнда нейрондық желі агенттің әрбір күй үшін ең тиімді әрекетті таңдауды тікелей үйренеді.

6. Қолдану салалары

Терең нықталандырумен оқыту келесі салаларда кеңінен қолданылады:

- Робототехника: автономды қозғалысты басқару;
- Қаржы жүйелері: нарықтағы шешім қабылдау;
- Компьютерлік ойындар: AlphaGo, Dota2, StarCraft секілді жобалар;
- Инженерлік процестер: энергияны оңтайландыру, желіні басқару;
- Табиғи тілді өңдеу: диалог жүйелерін бейімдеу.

7. Артықшылықтары мен қиындықтары

Артықшылықтары:

- Күрделі, көпөлшемді және бейсызық жүйелермен жұмыс істеуге мүмкіндік береді.
- Үлкен деректерден тиімді саясат үйренеді.
- Өзіндік бейімделу қабілеті жоғары.

Қиындықтары:

- Үлкен есептеу ресурстарын қажет етеді.

- Тұрақсыздық пен оқыту процесінің баяулығы.
- Марапат функциясын дұрыс анықтау қиындығы.

Қорытынды

Терең нықталандырумен оқыту — жасанды интеллекттің ең жылдам дамып келе жатқан бағыттарының бірі. Бұл тәсіл агентке қоршаған ортаны тікелей қабылдап, үздіксіз үйрену арқылы күрделі тапсырмаларды орындауға мүмкіндік береді. DQN, Policy Gradient және Actor-Critic сияқты әдістер DRL-дің практикалық қолдануларын кеңейтті. Болашақта DRL технологиялары автономды жүйелерді басқаруда, өндірістік процестерді оңтайландыруда және интеллектуалды шешім қабылдау жүйелерінде маңызды рөл атқарады.

Бақылау сұрақтары:

1. Терең нықталандырумен оқыту мен классикалық нықталандырумен оқытудың айырмашылығы неде?
2. DRL жүйесінде агент пен ортаның өзара байланысы қалай жүзеге асады?
3. DQN әдісінің негізгі артықшылықтары қандай?
4. Policy-based әдістердің ерекшелігі неде?
5. DRL қандай нақты салаларда қолданылады?

Пайдаланылған әдебиеттер:

1. Sutton, R. S., & Barto, A. G. *Reinforcement Learning: An Introduction*. MIT Press, 2018.
2. Li, Y. *Deep Reinforcement Learning: An Overview*. arXiv:1701.07274, 2017.
3. Silver, D. et al. *Mastering the game of Go with deep neural networks and tree search*. *Nature*, 529(7587), 484–489, 2016.
4. François-Lavet, V. et al. *An Introduction to Deep Reinforcement Learning. Foundations and Trends in Machine Learning*, 2018.
5. Goodfellow, I., Bengio, Y., & Courville, A. *Deep Learning*. MIT Press, 2016.