3-Дәріс

Перцептронды оқыту

Дәрістің мақсаты

Студенттерде нейрондық желілерді оқыту парадигмалары (бақыланатын, бақылаусыз және күшейтілген оқыту) туралы жүйелі түсінік қалыптастыру, сондай-ақ бір нейронды және көп нейронды перцептронның салмақтары мен ығысуын (bias) реттейтін оқыту ережесін құру және қолдану дағдыларын дамыту.

Оқу нәтижелері

Дәрісті меңгеру нәтижесінде студенттер:

- **Ажырата алады:** Нейрондық желілерді оқытудың негізгі үш санатын (бақыланатын, бақылаусыз, күшейту арқылы оқыту) және олардың қолдану аясын.
- **Түсінеді:** Бір нейронды және көп нейронды перцептрон архитектурасын, шектік активация функциясының (hardlim) жұмысын және оқыту қатесінің ($e = t - a$) мәнін.
- **Қолдана алады:** Перцептронды оқытудың бірыңғай ережесін ($W^{\text{жаңа}} = W^{\text{ескі}} + e p^T$), $b^{\text{жаңа}} = b^{\text{ескі}} + e$) нақты жаттығу мысалдарында салмақтар мен ығысуды есептеу үшін қолдана алады.

Оқыту ережесінің анықтамасы және оқыту санаттары

- **Оқыту ережесі (алгоритмі):** Желіге белгілі бір тапсырманы орындауды үйрету мақсатында оның салмақ коэффициенттерін (W) және ығысу (bias, b) мәндерін мақсатты түрде өзгерту процедурасы.

- **Бақыланатын оқыту (Supervised Learning):** Желіге кіріс деректерімен қатар дұрыс мақсатты шығыс мәндері (t) беріледі ($\{p_1, t_1\}, \{p_2, t_2\}, \dots$). Алгоритм нақты шығыс (a) пен мақсатты мән (t) арасындағы айырмашылықты азайту үшін салмақтарды түзетеді. *Перцептронды оқыту осы санатқа жатады.*

- **Күшейту арқылы оқыту (Reinforcement Learning):** Желіге әр қадамда дұрыс жауап берілмейді, оның орнына кірістер тізбегі бойынша желінің жалпы өнімділігін бағалайтын балл (сыйақы немесе айыппұл) беріледі. Көбінесе басқару жүйелерінде қолданылады.

- **Бақылаусыз оқыту (Unsupervised Learning):** Мақсатты жауаптар мүлдем болмайды. Салмақтар мен ығысулар тек кіріс сигналдарының құрылымына жауап ретінде өзгереді. Мұндай алгоритмдер деректерді кластерлеуге және векторлық кванттауға арналған.

Перцептрон архитектурасы

- **Құрылымы:** Перцептрон кіріс векторын (p өлшемі R), салмақ матрицасын (W өлшемі $S \times R$), ығысу векторын (b өлшемі $S \times 1$) және шектік белсендіру функциясын (**hardlim** немесе **hardlims**) қамтиды.
- **Шығыс есептеу формуласы:** $a = \text{hardlim}(Wp + b)$. Егер $Wp + b \ge 0$ болса, $a = 1$; егер $Wp + b < 0$ болса, $a = 0$.
- **Бір нейронды және көп нейронды перцептрон:** Бір нейрон ($S=1$) кіріс кеңістігін бір түзумен (гипержазықтықпен) екі класқа бөледі. Көп нейронды перцептронда ($S > 1$) әрбір нейрон өз бөлуші шекарасын құрып, кеңістікті 2^{S} аймаққа (класқа) бөлуге мүмкіндік береді.

Перцептронды оқыту ережесін құру

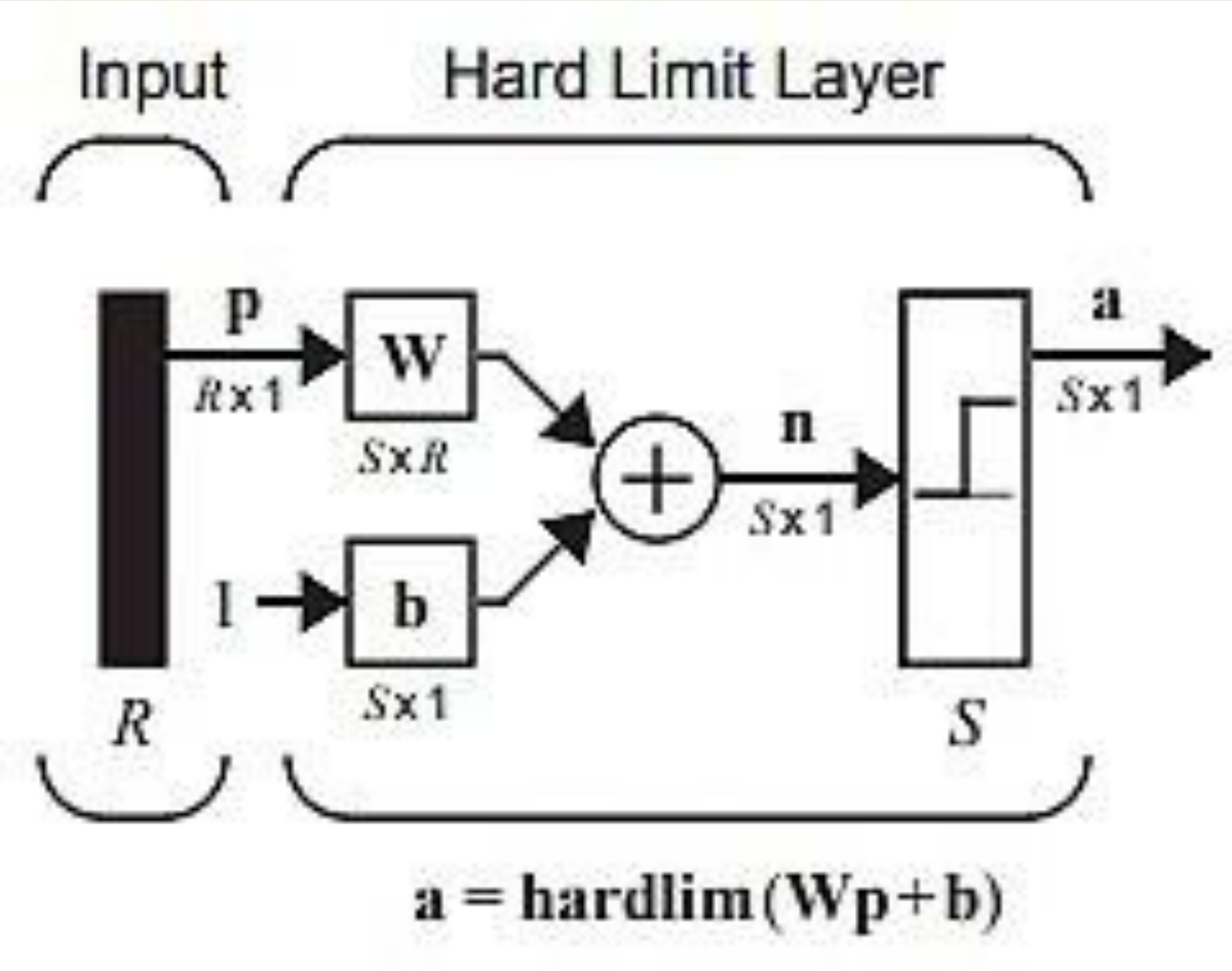
- **Оқыту қатесі (e):** Мақсатты мән (t) мен желінің нақты шығысы (a) арасындағы айырмашылық: $e = t - a$.
- **Мүмкін болатын үш жағдай ($t \in \{0, 1\}$, $a \in \{0, 1\}$):**
  - **$t = a$ ($e = 0$):** Желі дұрыс жіктеді — салмақ пен ығысу өзгермейді.
  - **$t = 1, a = 0$ ($e = 1$):** Кіріс векторы p бөлуші шекараның теріс жағында қалып қойды. Шекараны дұрыс бағытқа бұру үшін салмақ векторына кіріс векторы қосылады: $w^{\text{жаңа}} = w^{\text{ескі}} + p$, $b^{\text{жаңа}} = b^{\text{ескі}} + 1$.
  - **$t = 0, a = 1$ ($e = -1$):** Кіріс векторы p қате түрде оң аймаққа түсті. Салмақ векторынан кіріс векторы алынады: $w^{\text{жаңа}} = w^{\text{ескі}} - p$, $b^{\text{жаңа}} = b^{\text{ескі}} - 1$.

- **Оқыту ережесі** арқылы біз желінің салмағын және bias өзгерту процедурасын түсінеміз. (Бұл процедураны оқыту алгоритмі деп те атауға болады.)
- **Оқыту ережесінің мақсаты** - желіні қандай да бір тапсырманы орындауға үйрету. Нейрондық желіні оқыту ережелерінің көптеген түрлері бар.
- Олар үш үлкен санатқа бөлінеді:
    - *бақыланатын оқыту,*
    - *бақылаусыз оқыту және*
    - *күшейту (немесе деңгейлеп оқыту).*
- **Бақыланатын оқытуда (supervised learning)** оқу ережесі желіні арнайы мысалдарымен (жаттығулар жиынтығымен) қамтамасыз етеді.
- Оқыту ережесі желінің шығысын мақсатты көрсеткіштерге жақындату үшін желінің салмағын және бейімділігін (биас) реттеу үшін қолданылады. Перцептронды оқыту ережесі осы бақыланатын оқыту санатына жатады.

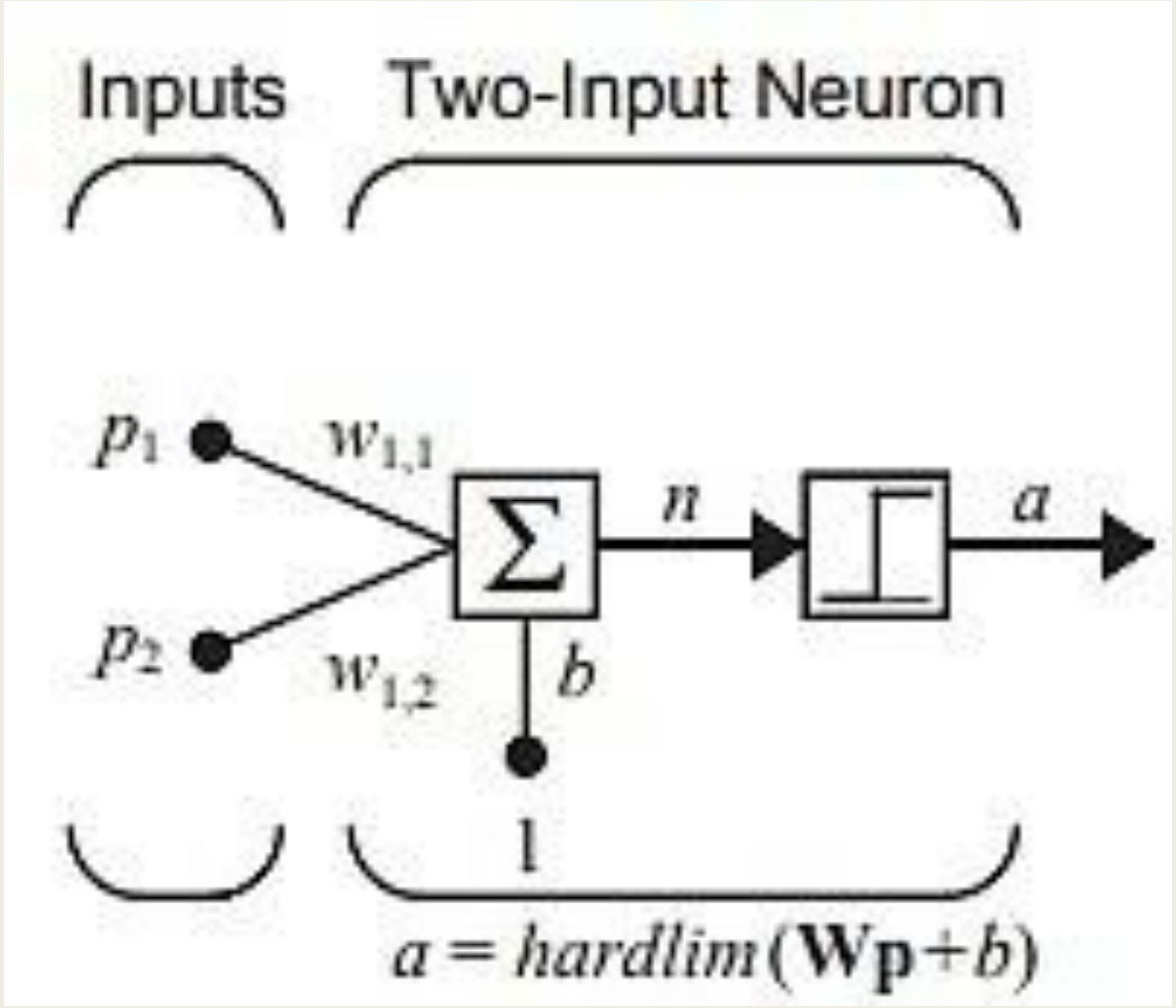
- Күшейту оқытуы (Reinforcement learning) бақыланатын оқытуға ұқсас, тек алгоритмге әр желіге дұрыс шығудың орнына, тек баға қойылады. Дәреже (немесе балл) - бұл кірістің кейбір реттілігі бойынша желінің өнімділігін өлшейтін көрсеткіш. Оқытудың бұл түрі қазіргі уақытта бақыланатын оқытуға қарағанда әлдеқайда сирек кездеседі. Бұл жүйелік қосымшаларды басқару үшін ең қолайлы болып көрінеді

- Бақылаусыз оқытуда (unsupervised learning) салмақ пен ауытқулар тек желілік кірістерге жауап ретінде өзгертіледі. Қол жетімді мақсатты нәтижелер жоқ. Бір қарағанда бұл мүмкін емес сияқты көрінуі мүмкін. Егер сіз не істеу керектігін білмесеңіз, желіні қалай үйретуге болады? Бұл алгоритмдердің көпшілігі қандай да бір кластерлік операцияны орындайды. Олар кіріс үлгілерін сыныптардың шектеулі санатына жатқызуды үйренеді. Бұл әсіресе векторлық квантталу сияқты қосымшаларда пайдалы.

Перцептрон архитектурасы



Бір нейронды перцептрон



$$a = hardlim(n) = hardlim(\mathbf{W}\mathbf{p} + b)$$

$$= hardlim({}_1\mathbf{w}^T\mathbf{p} + b) = hardlim(w_{1,1}p_1 + w_{1,2}p_2 + b)$$

$$n = {}_1\mathbf{w}^T\mathbf{p} + b = w_{1,1}p_1 + w_{1,2}p_2 + b = 0 .$$

$$w_{1,1} = 1 , w_{1,2} = 1 , b = -1 .$$

$$n = {}_1\mathbf{w}^T\mathbf{p} + b = w_{1,1}p_1 + w_{1,2}p_2 + b = p_1 + p_2 - 1 = 0 .$$

Қорытынды

Бұл дәрісте нейрондық желілерді оқытудың негізгі парадигмалары қарастырылып, бақыланатын оқыту санатына жататын перцептронды оқыту алгоритмі жан-жақты талданды. Перцептронның қатені (\(e = t - a\)) есептеу арқылы салмақ матрицасы мен ығысу векторын итерациялық түрде түзету принципі желіге сызықтық бөлінетін деректерді автоматты түрде жіктеуге үйренуге мүмкіндік береді.

Оқытудың бірыңғай ережесін меңгеру тек бір қабатты перцептрондарды ғана емес, сонымен қатар күрделірек көпқабатты нейрондық желілер мен градиенттік оқыту алгоритмдерінің (мысалы, Widrow-Hoff / LMS және қатені кері тарату әдістерінің) математикалық логикасын түсінуге берік негіз қалайды.

Өзін-өзі бағалау сұрақтары

1. **Оқыту парадигмалары:** Бақыланатын оқыту, бақылаусыз оқыту және күшейту арқылы оқытудың негізгі айырмашылықтарын түсіндіріңіз. Перцептронды оқыту қай санатқа жатады және неліктен?

2. **Перцептрон архитектурасы:** Бір нейронды перцептрон мен көп нейронды перцептронның құрылымы қандай? Олар белгілер кеңістігін неше аймаққа бөле алады?

3. **Оқыту қатесі:** Шектік функциясы hardlim болатын перцептронда оқыту қатесі $e = t - a$ қандай мәндерді қабылдай алады және әрбір мән нені білдіреді?

4. **Бірыңғай оқыту ережесі:** Перцептронның салмақ матрицасы (W) мен ығысу векторын (b) жаңартудың бірыңғай формулаларын жазыңыз. Неліктен формулада кіріс векторы транспонирленеді (p^{T})?

5. **Геометриялық интерпретация:** Егер желі мақсатты $t = 1$ мәнінің орнына $a = 0$ мәнін шығарса, салмақ векторы W мен шешім қабылдау шекарасы кеңістікте қалай өзгереді?

6. **Жинақталу шарты:** Қандай жағдайда перцептронды оқыту алгоритмі шексіз циклге түспей, дұрыс салмақтарды табуға кепілдік береді?

Әдебиеттер тізімі

1. **Хейган М. Т., Демут Х. Б., Бил М. Х., Де Хесус О.** Нейронные сети. Дизайн и применение (Neural Network Design). — М.: ДМК Пресс, 2021. — 800 с. *(4-тарау: «Правило обучения перцептрона» — оқыту санаттары, перцептрон архитектурасы және бірыңғай оқыту ережесін қорытып шығару).*

2. **Хайкин С.** Нейронные сети: полный курс. — 2-е изд. — М.: Издательский дом «Вильямс», 2018. — 1104 с. *(Розенблатт перцептроны және перцептронның жинақталу теоремасы).*

3. **Осовский С.** Нейронные сети для обработки информации. — М.: Финансы и статистика, 2017. — 448 с. *(Бірқабатты және көпнейронды перцептрондарды оқыту алгоритмдері).*

4. **Гудфеллоу Я., Бенджио И., Курвилль А.** Глубокое обучение. — М.: ДМК Пресс, 2018. — 652 с. *(Машиналық оқыту түрлері: supervised, unsupervised және reinforcement learning негіздері).*

5. **Бейсембекова Р. Н., Әбдірахманова Ж. Б.** Жасанды интеллект және нейрондық желілер негіздері: Оқу құралы. — Алматы: Қазақ университеті, 2021. — 184 б. *(Қазақ тіліндегі перцептрон модельдері және оларды оқыту әдістері).*