10- Дәріс

Кері таралу алгоритмі

Дәрістің мақсаты

Студенттерде көпқабатты перцептрондардың (MLP) сызықтық емес есептерді (мысалы, XOR мәселесі) шешу мен кез келген үзіліссіз функцияларды жуықтау (аппроксимациялау) мүмкіндіктері туралы білім қалыптастыру, сондай-ақ күрделі дифференциалдаудың тізбек ережесіне (Chain Rule) негізделген қатені кері тарату (Backpropagation) алгоритмінің математикалық аппараты мен сезімталдықтарды есептеу тетіктерін меңгерту.

Оқу нәтижелері

Дәрісті меңгеру нәтижесінде студенттер:

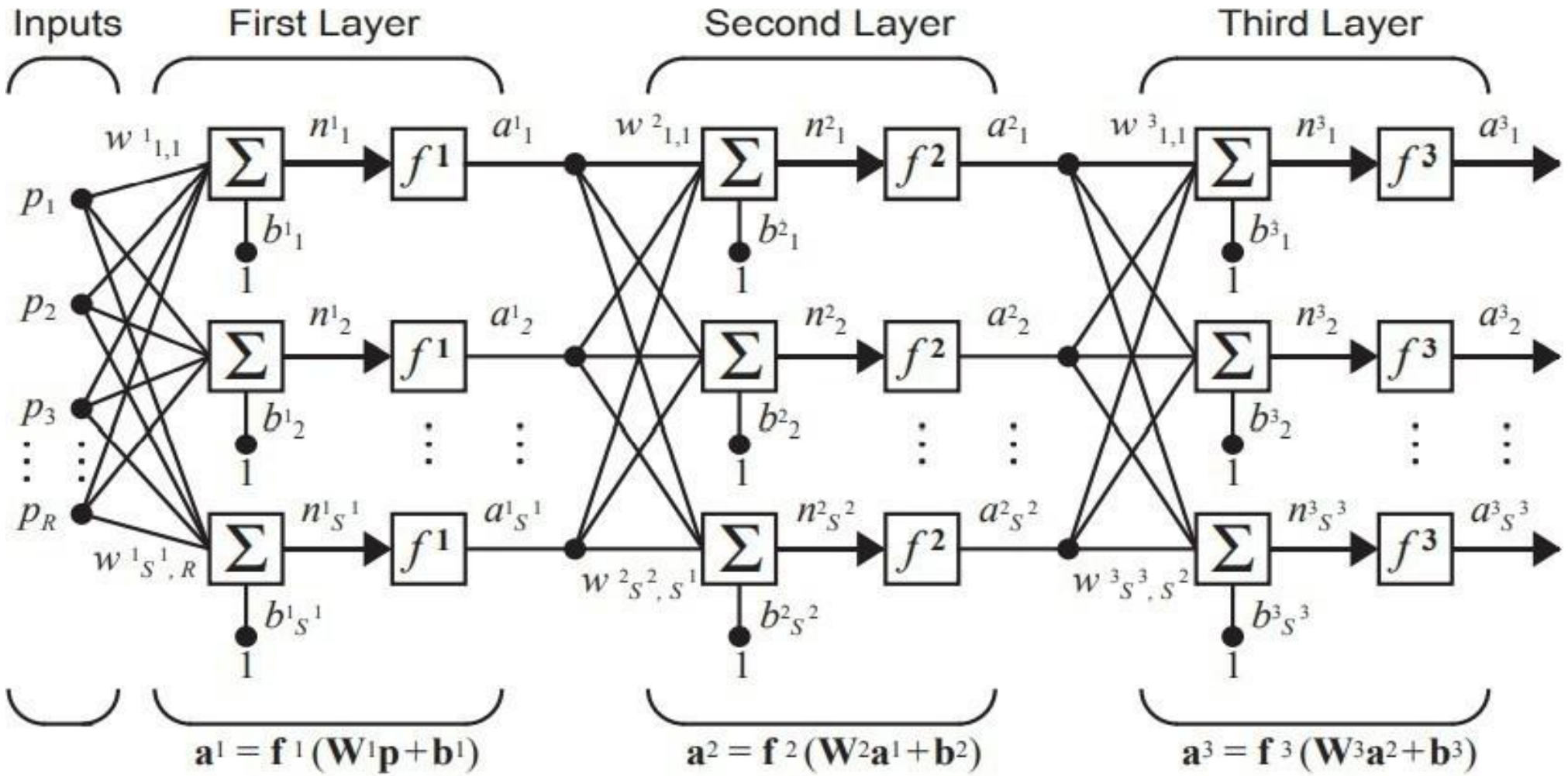
- **Түсінеді:** Бір қабатты желілердің шектеулерін (сызықтық бөлінбейтін XOR есебі), көпқабатты перцептронның (MLP) әмбебап аппроксиматорлық қасиетін және кері таралу алгоритмінің даму тарихын (П. Вербос, Д. Румельхарт).
- **Қолдана алады:** Күрделі функция туындысын табудың тізбек ережесін, Якобиан матрицасын және қабаттар бойынша сезімталдықтарды (\(s^{m}\)) кері бағытта қайта есептеу рекуррентік формулаларын.
- **Іске асыра алады:** Кері таралу алгоритмінің негізгі үш кезеңін (сигналды тура тарату, сезімталдықтарды кері тарату, салмақтар мен ығысуларды градиенттік жаңарту) нақты есептерде қолдана алады.

Кіріспе:

- Фрэнк Розенблаттың перцептронды үйрену ережесі және Бернард Уидроу мен Хоффтың LMS алгоритмі бір Керсептрон тәрізді желілерді оқытуға арналған. Алдыңғы дәрістерде талқылағанымыздай, бұл бірқабатты желілердің кемшілігі бар, олар тек сызықты түрде бөлінетін жіктеу мәселелерін шеше алады. Розенблат та, Уидроу да осы шектеулерді білетін және оларды жеңе алатын көпқабатты желілерді ұсынды, бірақ олар осы қуатты желілерді үйрету үшін алгоритмдерін жалпылай алмады.

- Көпқабатты желілік жұмыстарды үйрету алгоритмінің алғашқы сипаттамасы 1974 жылы Пол Вербостың тезисінде қамтылған болса керек [Werbo74]. Бұл жұмыс алгоритмді жалпы желілер контекстінде, нейрондық желілерді ерекше жағдай ретінде көрсетті және нейрондық желі қауымдастығында таралмады. 1980 жылдардың ортасында ғана кері таралу алгоритмі қайта ашылып, кеңінен жарияланды.

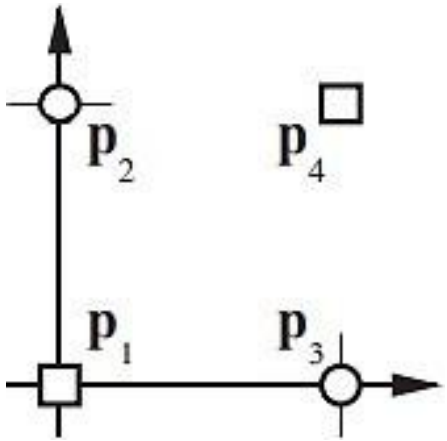
Көпқабатты перцептрондар

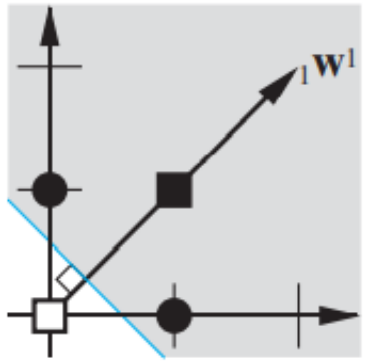


$$\mathbf{a}^3 = \mathbf{f}^3(\mathbf{W}^3\mathbf{f}^2(\mathbf{W}^2\mathbf{f}^1(\mathbf{W}^1\mathbf{p}+\mathbf{b}^1)+\mathbf{b}^2)+\mathbf{b}^3)$$

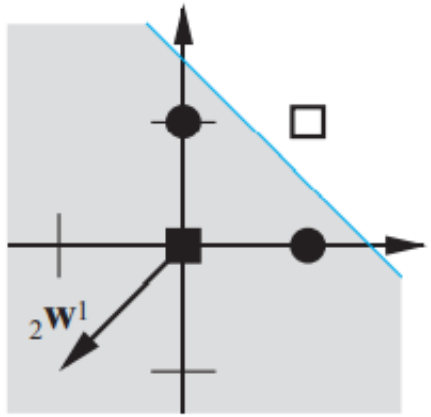
- Үлгі классификациясы үшін көп қабатты перцептронның мүмкіндіктерін көрсету үшін классикалық эксклюзивті немесе (XOR) мәселесін қарастырыңыз. XOR элементы үшін кіріс/мақсат жұптары болып табылады

$$\left\{\mathbf{p}_1 = \begin{bmatrix}0\\0\end{bmatrix}, t_1 = 0\right\}\left\{\mathbf{p}_2 = \begin{bmatrix}0\\1\end{bmatrix}, t_2 = 1\right\}\left\{\mathbf{p}_3 = \begin{bmatrix}1\\0\end{bmatrix}, t_3 = 1\right\}\left\{\mathbf{p}_4 = \begin{bmatrix}1\\1\end{bmatrix}, t_4 = 0\right\}.$$



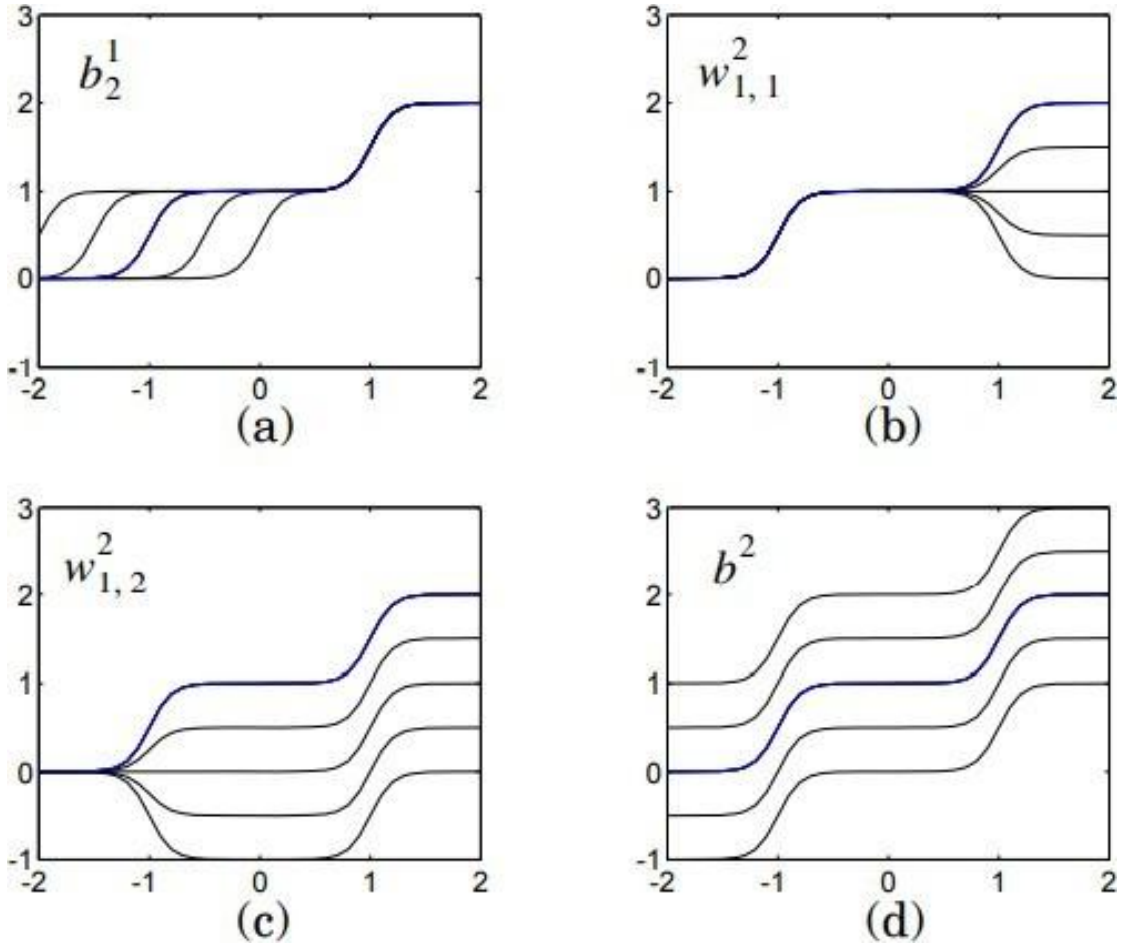


Layer 1/Neuron 1



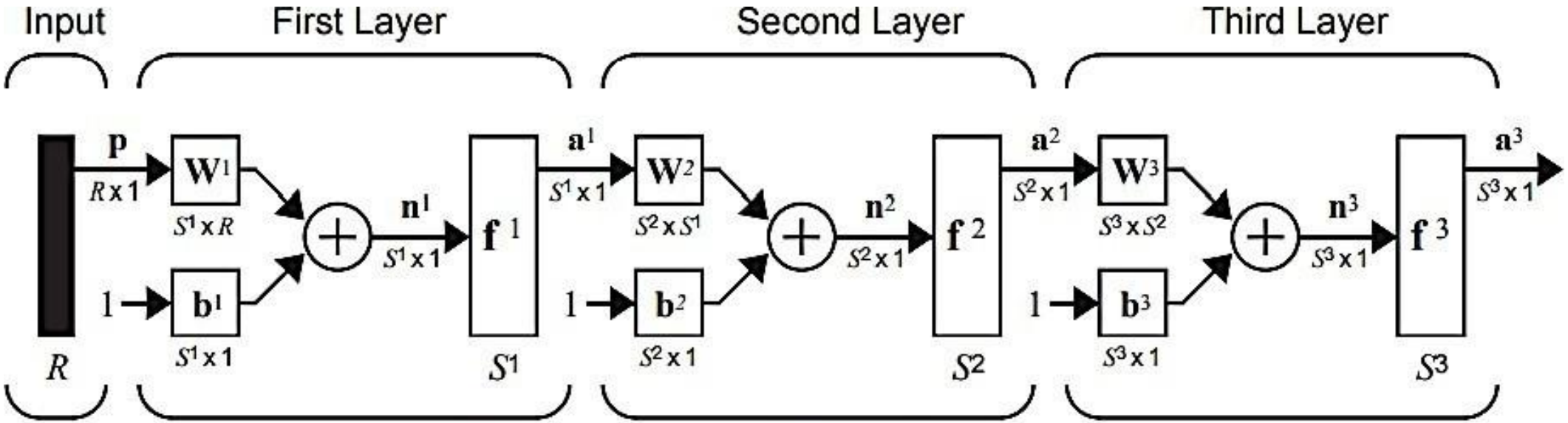
Layer 1/Neuron 2

$$-1 \le w^2_{1,1} \le 1\,,\ -1 \le w^2_{1,2} \le 1\,,\ 0 \le b^1_2 \le 20\,,\ -1 \le b^2 \le 1\,.$$



Енді біз үлгіні тану және функцияларды жақындату үшін көп қабатты перцептрондық желілердің күші туралы біраз түсінікке ие болғаннан кейін, келесі қадам осындай желілерді оқыту алгоритмін әзірлеу болып табылады.

Алгоритмі



$$\mathbf{a}^1 = \mathbf{f}^1(\mathbf{W}^1\mathbf{p}+\mathbf{b}^1) \qquad \mathbf{a}^2 = \mathbf{f}^2(\mathbf{W}^2\mathbf{a}^1+\mathbf{b}^2) \qquad \mathbf{a}^3 = \mathbf{f}^3(\mathbf{W}^3\mathbf{a}^2+\mathbf{b}^3)$$

$$\mathbf{a}^3 = \mathbf{f}^3(\mathbf{W}^3\mathbf{f}^2(\mathbf{W}^2\mathbf{f}^1(\mathbf{W}^1\mathbf{p}+\mathbf{b}^1)+\mathbf{b}^2)+\mathbf{b}^3)$$

$$\mathbf{a}^{m+1} = \mathbf{f}^{m+1}(\mathbf{W}^{m+1}\mathbf{a}^m + \mathbf{b}^{m+1}) \text{ for } m = 0, 1, \ldots, M-1,$$

- **Алгоритмнің жалпы циклі:**
  - *Тура жүріс (Forward pass):* Кіріс сигналы бірінші қабаттан соңғы (M)-ші қабатқа дейін таралып, шығыс (a^{M}) есептеледі.
  - *Кері жүріс (Backward pass):* Шығыс қабаттан бастап ((s^{M})) бірінші қабатқа дейін ((s^{1})) сезімталдықтар кері бағытта есептеледі.
  - *Параметрлерді жаңарту:* Есептелген сезімталдықтар негізінде барлық қабаттардың салмақтары мен ығысулары түзетіледі.

- Бірқабатты сызықтық желі (ADALINE) үшін бұл ішінара туындылар теңдеу арқылы ыңғайлы түрде есептеледі. Көп қабатты желі үшін қате жасырын қабаттардағы салмақтардың айқын функциясы емес, сондықтан бұл туындылар оңай есептелмейді. Қате жасырын қабаттардағы салмақтардың жанама функциясы болғандықтан, біз туындыларды есептеу үшін есептеудің тізбекті ережесін қолданамыз.

Қорытынды

Бұл дәрісте көпқабатты перцептрондардың (MLP) функционалдық мүмкіндіктері мен оларды оқытуға арналған іргелі **кері таралу (Backpropagation)** алгоритмі қарастырылды. Көпқабатты құрылымдардың сызықтық бөлінбейтін XOR есебін шешуге және кез келген үзіліссіз функцияларды дәл жуықтауға қабілетті екені көрсетілді.

Күрделі функцияны дифференциалдаудың тізбек ережесіне негізделген кері таралу алгоритмі қатенің сезімталдығын шығыс қабаттан бастап жасырын қабаттарға қарай тиімді таратуға мүмкіндік береді. Бұл алгоритм LMS (Widrow-Hoff) ережесінің көпқабатты желілерге арналған тікелей жалғасы болып табылады және қазіргі заманғы терең оқыту (Deep Learning) жүйелерінің негізгі есептеу тетігі қызметін атқарады.

Өзін-өзі бағалау сұрақтары

1. **Көпқабатты перцептронның артықшылығы:** Неліктен бір қабатты перцептрон классикалық XOR (эксклюзивті НЕМЕСЕ) мәселесін шеше алмайды және екі қабатты перцептрон бұл шектеуді қалай еңсереді?

2. **Функцияларды жуықтау:** Кез келген үзіліссіз функцияны жоғары дәлдікпен жуықтау (аппроксимациялау) үшін екі қабатты перцептронның жасырын және шығыс қабаттарында қандай белсендіру функциялары қолданылады?

3. **Тізбек ережесінің рөлі:** Неліктен көпқабатты желілерде жасырын қабат салмақтары бойынша қате функциясының дербес туындысын тікелей есептеу мүмкін емес және бұл жерде дифференциалдаудың тізбек ережесі қалай көмектеседі?

4. **Алгоритм кезеңдері:** Кері таралу алгоритмінің бір оқыту итерациясында орындалатын негізгі үш кезеңді ретімен сипаттап беріңіз.

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. *(11-тарау: «Обратное распространение» — XOR мысалы, функцияларды аппроксимациялау, тізбек ережесі және сезімталдықтарды кері тарату).*

2. **Хайкин С.** Нейронные сети: полный курс. — 2-е изд. — М.: Издательский дом «Вильямс», 2018. — 1104 с. *(Көпқабатты перцептрондар және қатені кері тарату алгоритмі).*

3. **Гудфеллоу Я., Бенджио И., Курвилль А.**

4. Глубокое обучение. — М.: ДМК Пресс, 2018. — 652 с. *(6-тарау: «Глубокие сети прямого распространения» — XOR мәселесі, градиенттік оқыту және кері тарату).*

5. **Николенко С., Кадурин А., Архангельская Е.** Глубокое обучение. Погружение в мир нейронных сетей. — СПб.: Питер, 2020. — 480 с. *(Кері таралу алгоритмінің математикалық қорытылуы және есептеу графтары).*

6. **Бейсембекова Р. Н., Әбдірахманова Ж. Б.** Жасанды интеллект және нейрондық желілер негіздері: Оқу құралы. — Алматы: Қазақ университеті, 2021. — 184 б. *(Қазақ тіліндегі көпқабатты перцептрондар мен кері тарату әдістері).*