

**Пламена Добрева Йовчева**

**Обобщеномрежово моделиране на невронни мрежи**

**АВТОРЕФЕРАТ**

На дисертационен труд за присъждане на образователна и научна степен  
„доктор” по научна специалност „Компютърни системи и технологии”,  
Област на висшето образование 5. Технически науки,  
Професионално направление 5.3. Комуникационна и компютърна техника

**Научни ръководители**

чл.-кор. проф. дмн дтн Красимир Т. Атанасов

проф. д-р Сотир Н. Сотиров

2020

Представеният дисертационен труд беше обсъден на разширен катедрен съвет на катедра „Компютърни системи и технологии“, в Университет „Проф. д-р Асен Златаров“ – гр. Бургас, на заседание, състояло се на 18.12.2020 г. и е насрочен за разкриване на процедура за защита пред научно жури със заповед ..... г. на Ректора на Университет „Проф. д-р Асен Златаров“ – гр. Бургас.

Дисертационният труд съдържа 121 страници, от които 35 фигури и са използвани 216 литературни източника. Резултатите са публикувани в 7 статии.

Защитата на дисертационния труд ще се състои на ..... от ..... часа в зала ....., Университет „Проф. д-р Асен Златаров“ – гр. Бургас.

Материалите по защитата са предоставени за заинтересованите в деловодството на Университет „Проф. д-р Асен Златаров“ – гр. Бургас.

Автор: Пламена Добрева Йовчева

Заглавие: Обобщеномрежово моделиране на невронни мрежи

*Изказвам искрената си благодарност на моите научни ръководители – чл.-кор. проф. д-мн д-тн Красимир Атанасов и проф. д-р Сотир Сотиров за споделените знания и опит, ценните им съвети и препоръки по време на изготвянето на този дисертационен труд.*

*Благодаря също на всички мои колеги от катедра „Компютърни системи и технологии“ и катедра „Компютърни и информационни технологии“.*

## Увод

В последните години изкуственият интелект и невронните мрежи се развиват бурно. Специалистите от различни области като конструиране, философия, физиология, психология са заинтересувани от възможностите на тази технология и търсят нейното приложение в собствените си области.

Този интерес е предизвикан от теоретичните и практическите възможности на невронните мрежи. Откриват се възможности за решаване на задачи в сфери, които до този момент се отнасят към областта на човешкия интелект, способността за научаване и запаметяване на образи, напомнящи човешките мисловни способности. Възможността за използването на невронните мрежи за разрешаването на различни типове задачи ги прави широко приложими в много сфери. За всяка задача съществуват няколко различни типа невронни мрежи, като за по доброто им решаване е важно да се разберат в детайли всяка една от тях. Нарастващия интерес към тази материя предполага нейното развитие с все по бързи темпове.

Обобщените мрежи са средство за конструиране на структурирани модели на комплексни системи, в които протичат паралелни във времето процеси и са изградени от множество взаимодействащи си компоненти. Това прави обобщените мрежи подходящ инструмент за моделирането на невронни мрежи тъй като невронната мрежа сама по себе си представлява паралелно взаимодействаща система. Обобщеномрежовия модел дава възможност за детайлно, стъпка по стъпка разглеждане на процеса, както от гледна точка на получените резултати така и във времето.

## **Цел и задачи на дисертационния труд**

Основната цел на изследванията, представени в дисертационния труд, е да се изследват различни невронни мрежи и алгоритми чрез моделирането им с помощта на Обобщени мрежи и програмната реализация на основните от тях. За да се постигне тази цел, са поставени следните задачи:

1. Разработване на обобщеномрежов модел на Неокогнитрон;
2. Разработване на обобщеномрежов модел на Дълбока невронна мрежа;
3. Разработване на обобщеномрежов модел на методи на обучение на невронни мрежи - Метод на Отпадане;
4. Разработване на обобщеномрежов модел за разпознаване на пръстов отпечатък с интуиционистки размити множества;
5. Предсказване на мощността на Термоелектричния генератор на базата на изкуствена невронна мрежа;
6. Разработване на обобщеномрежови модел на Стохастично спускане по градиента в комбинация с метод на Отпадане;
7. Програмна реализация на алгоритмите Стохастично спускане по градиента и метод на Отпадането;
8. Разработване на обобщеномрежови модел на Стохастично спускане по градиента в комбинация с метод на Отпадане с интуиционистки размита оценка.

В текста с [n\*] са означени статиите на автора, включени в дисертационния труд.

## 1. Въведение в Обобщените мрежи и Невронните мрежи

В тази глава са дадени основни дефиниции, които са необходими за изложението по-нататък, свързани с теорията на Обобщените мрежи и на Невронните мрежи.

### 1.1 Невронни мрежи

#### 1.1.1 Въведение в невронните мрежи

Представено е понятието биологичен неврон, изкуствен неврон, както и изкуствени невронни мрежи.

#### 1.1.2 Видове невронни мрежи

Представени са видовете невронни мрежи, както и основните два типа за обучение – с учител и без учител.

#### 1.1.3. Дълбоки невронни мрежи

Разяснена е същността на Дълбока невронна мрежа и разликата с традиционните невронни мрежи.

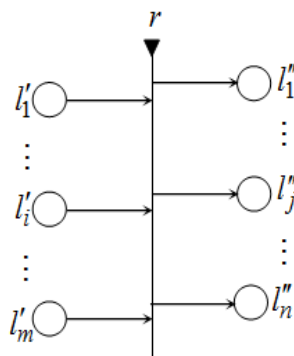
## 1.2 Обобщени мрежи

### 1.2.1 Въведение в теорията на обобщените мрежи

Обобщените мрежи (ОМ) са средства за моделиране на паралелно протичащи във времето процеси. Те включват като частни случаи мрежите на Петри и всички техни модификации.

#### Дефиниция на преход и обобщена мрежа

Неформално е описано понятието Обобщена Мрежа (ОМ) и са описани преходите ѝ. Графично преходът ОМ в се представя като съвкупност от два елемента:  $\bigcirc$  и  $\uparrow$  (Фиг. 1).



Фиг. 1 Представяне на преход в ОМ

Дадена е формална дефиниция на понятията преход и ОМ.

### 1.2.2 Алгоритми за функциониране на преход и обобщена мрежа

Описани са основни алгоритмите за движение на ядрата през дадени преход и ОМ.

### **1.3. Съществуващи обобщеномрежови модели на невронни мрежи**

Описани са основните обобщеномрежови модели на невронни мрежи, създадени досега.

### **1.4 Извод**

В Глава първа от дисертационния труд е направен литературен обзор на същността на невронните мрежи и теорията на обобщените мрежи, като по този начин се полага основата на анализът и разработването на обобщеномрежови модели, които ще бъдат представени в следващите две глави на разработката.

## **2. Моделиране на невронни мрежи с обобщени мрежи**

### **2.1 Обобщеномрежов модел на невронна мрежа Неокогнитрон**

Мрежата, наричана "Неокогнитрон", е самоорганизираща се мрежа и има способността да разпознава модели въз основа на разликата в тяхната форма. Всеки шаблон, изчислен като принадлежност към определена категория, се определя най-точно от Неокогнитрон в същата категория.

Неокогнитрон е в състояние правилно да идентифицира изображението, дори ако има нарушение или движение в положението. Самоорганизацията в Неокогнитрон също се осъществява неконтролируемо – за обучението на самоорганизацията се неокогнитрон са необходими само колекция от повтарящи се модели на разпознаваемия образ и не е необходима информация за категории, които включват шаблони. Неокогнитрон придобива способността сам да класифицира и разпознава тези структури въз основа на различия във формите. Неокогнитрон разпознава правилно изображения без влиянието на смени и дори при сериозно изкривяване на формата на обекта.

Неокогнитрон има йерархична структура. Разпознаваемата информация за изображение, която попада във входния слой на Неокогнитрон, се обработва стъпка по стъпка, на всеки етап от многослойната мрежа. Клетките на по-дълбок етап обикновено са склонни да реагират селективно със сложната функция и в същото време има чувствителни области и те са чувствителни към промени в позицията на изображението. По този начин всяка клетка в по-дълбок слой е отговорна само за определен модел.

$C$ -клетките, които наподобяват сложни клетки във визуалната кора, се вмъкват в мрежата, за да позволят позиционни грешки в характеристиките на стимула. Входните връзки на  $C$ -клетките, които идват от  $S$ -клетките на предходния слой, са фиксирани и неизменни. Всяка  $C$ -клетка получава възбудителни входни въздействия от група  $S$ -клетки, които извличат една и съща характеристика, но от малко по-различни позиции.  $C$ -клетката отговаря, ако поне една от тези  $S$ -клетки дава резултат. Дори ако функцията за стимулиране се измества в положение и друга  $S$ -клетка идва да реагира вместо първата, същата  $C$ -клетка продължава да реагира. По този начин, реакцията на  $C$ -клетката е по-малко чувствителна към изместване в положението на входния модел. Можем също да кажем, че  $C$ -клетките правят операция на замъгляване, тъй като

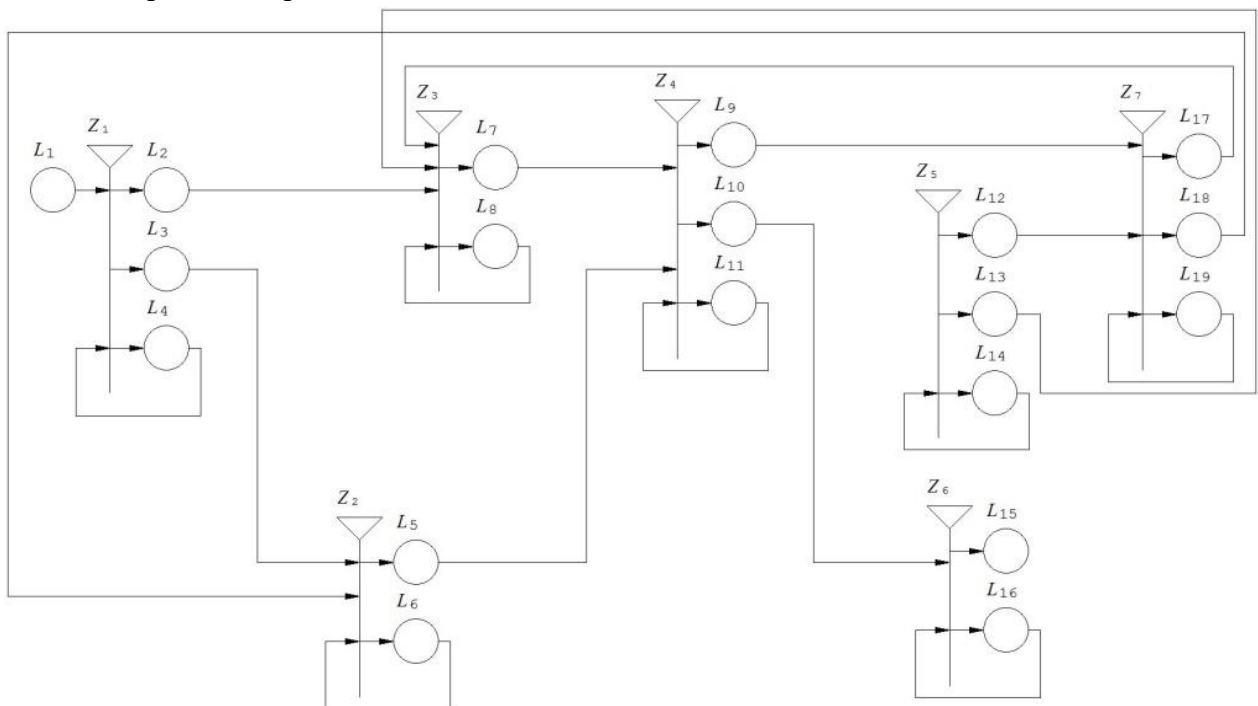
реакцията на слой от  $S$ -клетки е пространствено замъглена в отговора на следващия слой от  $C$ -клетки.

Обобщената мрежа (фиг. 2.1.2) съдържа следните множество от преходи:

- $Z_1$  - Разделяне на модела на въвеждане;
- $Z_2$  - Изчисляване на изходната реакция на  $V$  възли;
- $Z_3$  - Изчисляване на параметъра за  $S$  възли;
- $Z_4$  - Формиране на мащабиран вход на  $S$  възли;
- $Z_5$  - Инициализиране на теглата за слоевете  $S$  и  $C$ ;
- $Z_6$  - Разпознаване от мрежата;
- $Z_7$  - Изчисляване на стойностите от  $C$  възлите.

Първоначално следните ядра влизат в Обобщената мрежа.

- В позиция  $L_1$  едно ядро с първоначална характеристика „Входен модел за неокогнитрон“.
- В позиция  $L_{14}$  едно ядро с първоначална характеристика „Структура на невронната мрежа“.



Фигура 2.1.2. Обобщен модел на невронната мрежа Неокогнитрон

Преходът  $Z_1$  има следната форма:

$$Z_1 = \langle \{L_1, L_4\}, \{L_2, L_3, L_4\}, R_1, \vee (L_1, L_4), \rangle$$



$$R_1 = \begin{array}{c|ccc} & L_2 & L_3 & L_4 \\ \hline L_1 & false & false & true, \\ L_4 & W_{4,2} & W_{4,3} & true \end{array}$$

$W_{4,2} = W_{4,3} =$  "Има  $\alpha$  ядро в позиция  $L_1$ ".

Ядро  $\alpha$  от позиция  $L_1$ , което постъпва в позиция  $L_4$ , не получава нова характеристика.

Ядро  $\alpha$  от позиция  $L_4$  се разделя на две ядра и получава нова характеристика.

$$x_{cu}^{\alpha'} = \text{"Входна мрежа за обучение"}$$

Ядро  $\alpha$  от позиция  $L_4$ , което постъпва в позиция  $L_4$ , не получава нова характеристика.

Ядро  $\alpha'$  от позиция  $L_4$ , което постъпва в позиция  $L_2$  и  $L_3$ , не получава нова характеристика.

Преходът  $Z_2$  има следната форма:

$$Z_2 = \langle \{L_3, L_6, L_{18}\}, \{L_5, L_6\}, R_2, \vee (L_3 \wedge (L_6, L_{18})) \rangle,$$

$$R_2 = \begin{array}{c|cc} & L_5 & L_6 \\ \hline L_3 & false & true \\ L_6 & W_{6,5} & true \\ L_{18} & false & true \end{array}$$

$W_{6,5} =$  "Стойностите на  $V$  възлите са изчислени".

Ядро  $\alpha'$  от позиция  $L_3$ , което постъпва в позиция  $L_6$ , не получава нова характеристика.

Ядро  $\eta$  от позиция  $L_{18}$ , което постъпва в позиция  $L_6$ , не получава нова характеристика.

Ядрото  $\delta$  от позиция  $L_6$ , което постъпва в позиция  $L_6$ , получава нова характеристика

$$x_{cu}^{\delta} = \sqrt{\sum t_i c_i^2}$$

Ядро  $\delta$  от позиция  $L_6$ , което влиза в позиция  $L_5$ , не получава нова характеристика.

Преходът  $Z_3$  има следната форма:

$$Z_3 = \langle \{L_2, L_8, L_{13}, L_{17}\}, \{L_7, L_8\}, R_3, \vee (L_2 \wedge (L_8, L_{13}, L_{17})) \rangle,$$

$$R_3 = \begin{array}{c|cc} & L_7 & L_8 \\ \hline L_2 & false & true \\ L_8 & W_{8,7} & true \\ L_{13} & false & true \\ L_{17} & false & true \end{array} ,$$

$W_{8,7}$  = "Стойностите за  $e$  параметъра са изчислени".

Ядро  $\alpha'$  от позиция  $L_2$ , което влиза в позиция  $L_8$ , не получава нова характеристика.

Ядро  $\eta$  от позиция  $L_{17}$ , което влиза в позиция  $L$ , не получава нова характеристика.

Ядро  $\beta''$  от позиция  $L_{13}$  се обединява с ядрата  $\alpha'$  и  $\eta$  от позиция  $L_8$  и получава нова характеристика

$$x_{cu}^{\delta} = \sum_i c_i w_i$$

Ядро  $\varepsilon$  от позиция  $L_8$ , което влиза в позиция  $L_7$ , не получава нова характеристика.

Преходът  $Z_4$  има следната форма:

$$Z_4 = \langle \{L_5, L_7, L_{11}\}, \{L_9, L_{10}, L_{11}\}, R_4, \vee (L_{11} \wedge (L_5, L_7)) \rangle,$$

$$R_4 = \begin{array}{c|ccc} & L_9 & L_{10} & L_{11} \\ \hline L_5 & false & false & true \\ L_7 & false & false & true \\ L_{11} & W_{11,9} & W_{11,10} & true \end{array} ,$$

$W_{11,9}$  = "Стойностите за  $S$  възлите са изчислени";

$W_{11,10}$  = "Стойностите за  $S_3$  възлите са изчислени".

Ядрата  $\delta$  и  $\varepsilon$  от позиция  $L_5$  и  $L_7$ , които постъпват в позиция  $L_{11}$  се обединяват в ядро  $\phi$ , което получава нова характеристика

$$x_{cu}^{\phi} = \frac{1 + pr_1 x_{cu}^{\varepsilon}}{1 + pr_1 x_{cu}^{\delta} w_0}$$

Ядро  $\phi$  от позиция  $L_{11}$ , което постъпва в позиция  $L_9$ , не получава нова характеристика.

Ядрата, които постъпват в позиция  $L_{10}$  от позиция  $L_{11}$  получават нови характеристики

$$x_{cu}^{\phi'} = \text{"Стойности за } S_3 \text{ единиците"}.$$

Ядрата  $\varepsilon$  и  $\delta$ , които постъпват в позиция  $L_{11}$  от позиция  $L_{11}$ , не получават нови характеристики.

Преходът  $Z_5$  има следната форма:

$$Z_5 = \langle \{L_{14}\}, \{L_{12}, L_{13}, L_{14}\}, R_5, \vee (L_{14}) \rangle,$$

$$R_5 = \frac{\quad}{L_{14}} \left| \begin{array}{cc} L_{12} & L_{13} & L_{14} \\ W_{14,12} & W_{14,13} & true \end{array} \right.,$$

$W_{14,12}$  = "Теглата на  $C$  възлите са инициализирани";

$W_{14,13}$  = "Теглата на  $S$  възлите са инициализирани".

Ядро  $\beta$  от позиция  $L_{14}$ , което постъпва в позиция  $L_{12}$ , получава нова характеристика

$$x_{cu}^{\beta'} = \frac{1}{1 + |k| + |h|}$$

Ядро  $\beta$  от позиция  $L_{14}$ , което постъпва в позиция  $L_{13}$ , получава нова характеристика

$$x_{cu}^{\beta''} = \langle w_0, w_i \rangle$$

Ядро  $\beta$  от позиция  $L_{14}$ , което постъпва в позиция  $L_1$ , не получава нова характеристика.

Преходът  $Z_6$  има следната форма:

$$Z_6 = \langle \{L_{10}, L_{16}\}, \{L_{15}, L_{16}\}, R_6, \vee (L_{10}, L_{16}) \rangle,$$

$$R_6 = \frac{\quad}{L_{10}} \left| \begin{array}{cc} L_{15} & L_{16} \\ false & true \\ L_{16} & W_{16,15} & true \end{array} \right.,$$

$W_{16,15}$  = "Мрежата е разпозната".

Ядро  $\phi'$  от позиция  $L_{10}$ , което постъпва в позиция  $L_{16}$ , не получава нова характеристика.

Ядро  $\phi'$  от позиция  $L_{16}$ , което постъпва в позиция  $L_{15}$ , получава нова характеристика

$$x_{cu}^{\phi''} = \text{"Мрежата е разпозната"}.$$

Преходът  $Z_7$  има следната форма:

$$Z_7 = \langle \{L_9, L_{12}, L_{19}\}, \{L_{17}, L_{18}, L_{19}\}, R_7, \vee (L_{12} \wedge (L_9, L_{19})) \rangle,$$

|          |              |              |             |
|----------|--------------|--------------|-------------|
|          | $L_{17}$     | $L_{18}$     | $L_{19}$    |
| $L_9$    | <i>false</i> | <i>false</i> | <i>true</i> |
| $L_{12}$ | <i>false</i> | <i>false</i> | <i>true</i> |
| $L_{19}$ | $W_{19,17}$  | $W_{19,18}$  | <i>true</i> |

$W_{19,17}$  = "Стойностите на  $S$  възлите са изчислени";

$W_{19,18}$  = "Стойностите на  $V$  възлите са изчислени".

Ядро  $\phi$  от позиция  $L_{12}$  се обединява с ядро  $\beta'$  от позиция  $L_{19}$  и получава нова характеристика

$$x_{cu}^{\eta} = \sum_i s_i u_i$$

Ядро  $\eta$  от позиция  $L_{19}$  се разделя на две ядра, които влизат в позиция  $L_{17}$  и  $L_{18}$ , където не получават нова характеристика.

За следващите обобщеномрежови модели преходите са описани по аналогичен начин.

## 2.2 Обобщеномрежов модел на Дълбока невронна мрежа

Дълбокото обучение позволява създаването на многослойни изчислителни модели, които дават възможност на невронната мрежа да бъде обучена да представя данни с голяма абстракция. Този метод значително подобрява разпознаването на речта, разпознаването на обекти и може да се използва и в медицината, откриването на нови лекарства и т.н. Дълбокото обучение открива сложна структура в голям набор от данни, използвайки алгоритъм за обратно разпространение, за да покаже как една машина трябва да променя вътрешните си параметри, които се използват за изчисляване на производителността във всеки слой.

Методите за дълбоко обучение са представителни методи на многостепенно обучение, съставени от прости, но нелинейни модули, всеки преобразуващ резултата от едно ниво в представяне на друго. Например, изображението идва под формата на масив от пикселни стойности, а функциите в първия слой обикновено представляват наличието или отсъствието на ръбове на определени места в изображението. Във втория слой обикновено има мотиви, като се забелязват определени подредби на ръбовете, независимо от малките вариации в позициите на ръба. В третия слой мотивите могат да бъдат комбинирани в по-големи комбинации, които съответстват на части от известни обекти, а следващите слоеве ще намерят обекти, които са комбинации от тези фрагменти. Ключовият аспект тук е, че тези слоеве не са проектирани от хора, а извличат данните, като използват процес на обучение. Не е нужно да описваме на компютър какво е легло, можем просто да го включим в интернет и да му дадем милиони снимки на легла, така че машината получава обща представа какво е легло. След това го тестваме с още повече снимки и поправяме грешките, което подобрява разпознаването.

Терминът "дълбоко" обикновено се отнася до броя на скритите слоеве в невронната мрежа. Традиционните невронни мрежи съдържат само 2-3 скрити слоя, докато дълбоките мрежи могат да имат до 100. Моделите се обучават чрез използване на голям набор от данни и архитектури на невронни мрежи, които учат функции директно от данните, без да е необходимо ръчно извличане на функции. Обобщената мрежа (фиг. 2.2.1) съдържа следното множество от преходи

$$A = \{Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7\},$$

Преходите са описани, както следва:

$Z_1$  - обобщаващ случаен вектор за стойности на матрицата на тегло  $W$  и отклонение  $b$ ;

$Z_2$  - изчисляване на сумата от входните въздействия;

$$n_M = \sum_{i=1}^{R_M} (P_i W_{Mi}^2) + b_M$$

за  $M$ -тия слой от неврони, където:

- $R_M$  е броят на невроните в  $M$ -тия слой;
- $P$  е входния вектор на мрежата;
- $W$  е матрица с тегловните коефициенти за всеки сигнал и слой;
- $b$  е отклонение на слоевете;

$Z_3$  – изчисляване на изходната реакция  $a_M = F_M(n_M)$  за  $M$ -тия слой;

$Z_4$  – определяне на разликата между получената стойност ( $a^3$ ) и фиксираната учебна цел и изчисляване на най-малката квадратична грешка между тях;

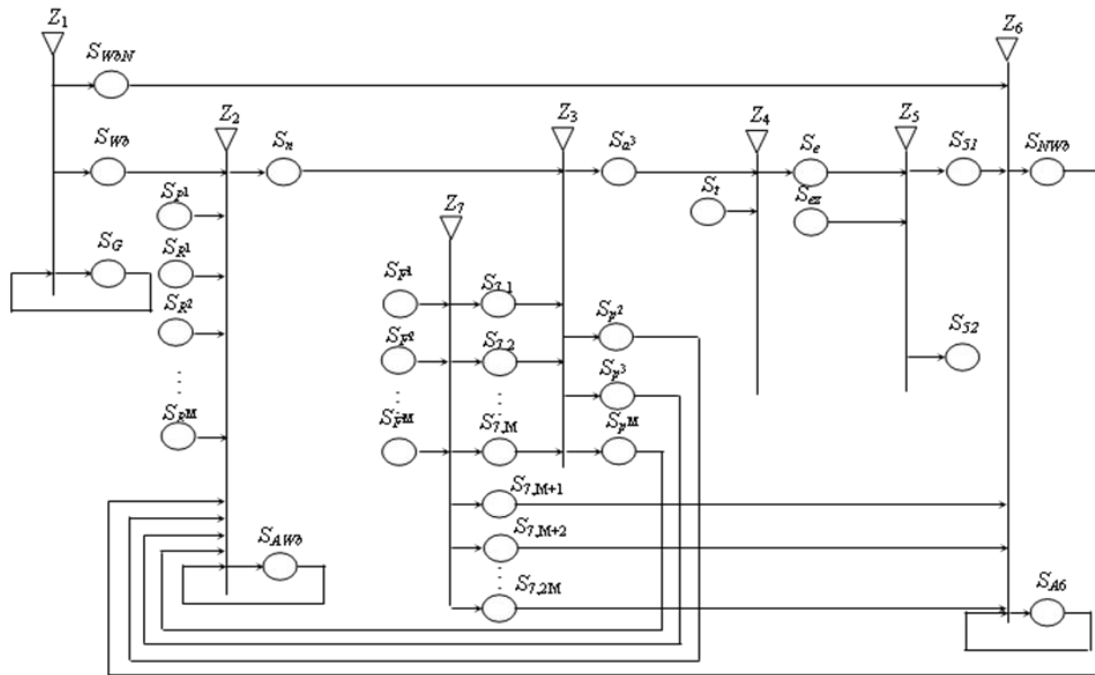
$Z_5$  – определяне дали невронната мрежа е обучена или не;

$Z_6$  – изчисляване на нови тегловни коефициенти и отклонение на невронната мрежа;

$Z_7$  – предаване на трансфериращите функции за изчисляване на новите тегловни коефициенти и отклонения на невронната мрежа на изходите  $a^3$ .

В мрежата има следните основни ядра:

- В позиция  $S_G$  – едно  $\alpha_G$ -ядро с характеристика "Случаен генератор на числа" за обобщаване на коефициентите на теглата и отклонението на невронната мрежа,
- Във всяка от позициите  $S_{F1}$ ,  $S_{F2}$  и  $S_{FM}$  има по едно  $\alpha_{Ri}$  ядро,  $1 \leq i \leq M$ , с характеристика "Предавателна функция от  $i$ -тия слой на невронната мрежа",
- В позиция  $S_i$  – едно  $\alpha_i$  ядро с характеристика "Цел на обучението на невронната мрежа, която трябва да бъде получена на изхода",
- В позиция  $S_{ez}$  – едно  $\alpha_{ez}$  ядро с характеристика "Предварително фиксирана грешка за обучението на невронната мрежа".



Фигура 2.2.1. Обобщеномрежов модел на мрежа за дълбоко обучение

### 2.3 Обобщеномрежов модел на методи на обучение на невронни мрежи - Метод на Отпадане

Метод на Отпадане е алгоритъм за обучение на невронни мрежи чрез случайно изключване на неврони по време на тренировка. Метод на Отпадане в дълбокото обучение работи по следния начин - един или повече неврони в невронната мрежа периодически се изключват, така че да не взаимодействат с мрежата. Чрез този метод обучените тегловни коефициенти на невроните стават малко по-малко чувствителни към теглата на другите и по-малко зависими от другите неврони, с които са свързани.

По принцип изпускането помага на мрежата да обобщи по-добре и да увеличи точността, тъй като въздействието на неврони намалява поради отпадането.

Метод на Отпадане тренира ансамбъл, състоящ се от всички подмрежи, които могат да бъдат структурирани чрез премахване на несъществуващи неврони от основната базова мрежа. Изборът кои неврони да отпаднат е напълно случаен. Вероятността даден неврон да бъде изключен е стойност между 0 и 1.

Обобщената мрежа съдържа следните преходи:

$$A = \{Z_1, Z_2, Z_3, Z_4, Z_5, Z_6\},$$

Преходите са описани както следва:

$Z_1$  – обобщаващ случаен вектор за стойности на тегловната матрица  $W$ ;

$Z_2$  – изчисляване на  $avgk$ ;

$Z_3$  – изчисляване на изходна реакция  $a_k = F_k(n_k)$  от  $k$ -тия слой;

$Z_4$  – определяне на разликата между получената стойност ( $L_0$ ), фиксираната учебна цел и най-малката квадратична грешка;

$Z_5$  – определяне дали невронната мрежа е достатъчно обучена;

$Z_6$  – изчисляване на новите тегловни коефициенти.

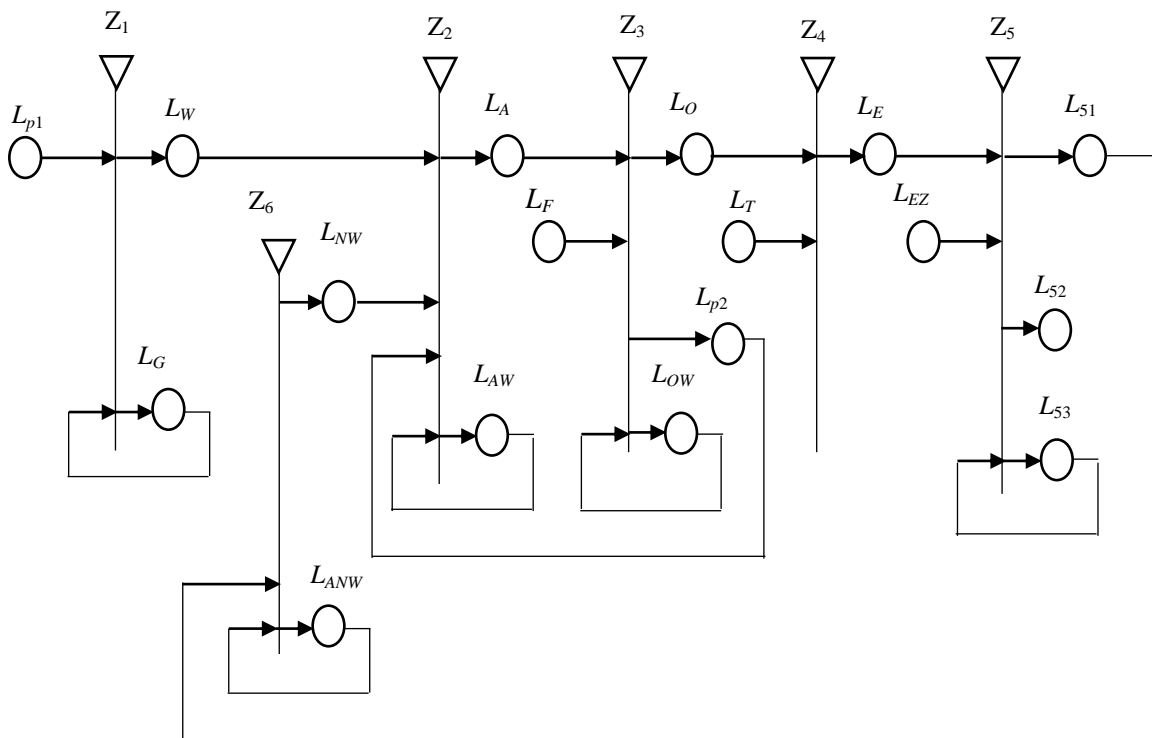
Обобщената мрежа съдържа следните ядра :

В позиция  $L_G$  – едно  $\alpha_G$ -ядро с характеристика "Случаен генератор на числа" за тегловни коефициенти,

Във всяка позиция  $L_F$  има едно  $\alpha_i$ -ядро,  $1 \leq i \leq k$ , с характеристика "Предавателна функция от  $i$ -тия слой на невронната мрежа",

В позиция  $L_T$  – едно  $\alpha_t$  - ядро с характеристика "Учебна цел за изходна реакция на невронната мрежа",

В позиция  $L_{EZ}$  – едно  $\alpha_{ez}$  - ядро с характеристика "Предварително фиксирана грешка при обучението на невронната мрежа".



Фигура 2.3.1. Обобщеномрежов модел на метод на Отпадане

## 2.4 Обобщеномрежов модел за разпознаване на пръстов отпечатък с интуиционистки размити множества

Един от най-популярните биометрични методи се основава на сравняване на пръстови отпечатъци, като се използват две основни техники за разпознаване. Първата взема и сравнява детайлите на гребените на линиите, където се разминават и прекъсват. Втората техника измерва и сравнява направленията на линиите на пръстовия отпечатък, напр. промени, гребени и арки, тяхната ширина и дълбочина. Възможно е да възникнат някои трудности при автоматичното разпознаване, когато пръстите са замърсени, мокри или покрити с масло. Освен това, отпечатъкът зависи от това специфичното използване на ръката в някои професии, като например строителство, при раняване, често се наблюдават абразия или други подобни промени на кожата на ръката. Също, когато кожата остарява, са възможни промени в линиите на пръстовия отпечатък. Това

обаче е поправимо, ако периодично базата данни се актуализира, което е важно за сравняване на идентификация. За да е сигурно разпознаването, необходимата база данни, с която сравнението ще бъде направено, трябва да бъде с много високо качество.

Предложен е метод за оценка на еквивалентността на пръстовите отпечатьци. За оценка на еквивалентността на съответната оценка се използва теорията за Интуиционистки размити множества. Интуиционистки размитите множества, дефинирани от Красимир Атанасов, представляват разширение на концепция за размити множества, показваща функция  $\mu_A(x)$ , определяща принадлежността на елемент  $x$  към множество  $A$ , оценена в интервала  $[0; 1]$ .

Получената интуиционистки размита оценка може да влияе или не на разпознаването в системата. Също така има известна степен на неопределеност, когато информацията не е достатъчна. В този случай се използват прагови стойности за минималната и максималната степен на принадлежност и непринадлежност. За описанието на целия процес е използван обобщеномрежов модел

Обобщена мрежа включва набора от преходи:

$$A = \{Z_1, Z_2, Z_3, Z_4\},$$

където преходите описват следните процеси:

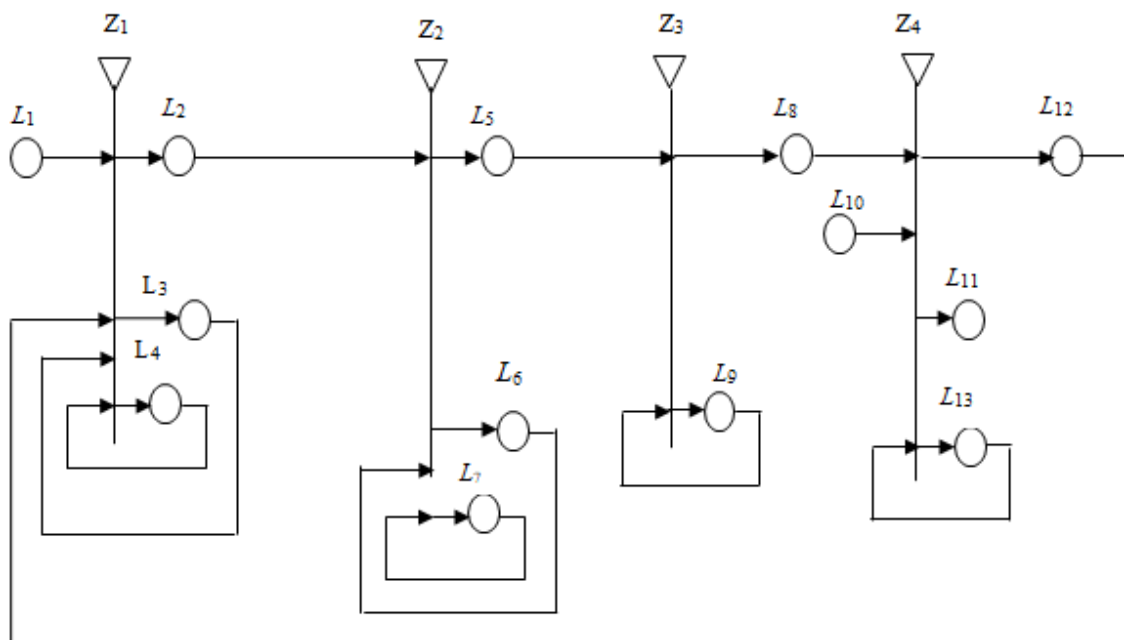
$Z_1$  – Проверява се дали пръстът е в правилна позиция;

$Z_2$  – Проверява се дали пръстът има правилно въртене;

$Z_3$  - Корекция на интензивността;

$Z_4$  - Интуиционистки размита оценка на пръстови отпечатьци.

Ядро влиза в мрежата от позиция  $L_1$  с първоначална характеристика: "нов сканиран отпечатък".



Фигура 2.4.1. Обобщеномрежов модел за разпознаване на пръстов отпечатък с интуиционистки размити множества



## 2.5 Изводи

Във втората глава са разработени са четири обобщеномрежови модела.

В първият от тях е реализиран обобщеномрежов модел описващ структурата на невронната мрежа Неокогнитрон. Разглежда се действието на всеки слой и ориентиращата подсистема. Алгоритъмът, използван за тази структура, се формира като многослойна мрежа. Самоорганизацията с неконтролирано обучение с помощта на повтарящи се модели е необходим елемент от мрежата и елиминира необходимостта от информация за класа, към който принадлежи изображението.

Вторият модел разглежда алгоритъма на обучение на Дълбока невронна мрежа описан чрез Обобщена мрежа. Моделът дава възможност да се разгледат действието на отделните слоеве.

Третият модел представя алгоритъма метод на Отпадане. Действието е описано детайлно чрез инструмента на Обобщените мрежи. Трябва да се отбележи, че големите многослойни мрежи са бавни и прилагането на този метод е една добра техника за справянето с този проблем.

В четвъртият модел се разглежда разпознаване на пръстов отпечатък. Моделът дава възможност да се разгледат различните етапи на идентификация на пръстови отпечатъци. Отпечатъкът е една от най-уникалните части на човешкото тяло и затова се използва в устройства за идентификация. За целта използваме оценка с интуиционистки размити множества. Получените интуиционистични размити оценки отразяват разпознаването на системата.

## 3. Обобщеномрежови модели на приложения на невронни мрежи

### 3.1 Предсказване на мощността на Термоелектричния генератор на базата на изкуствена невронна мрежа

През последните години активно се използват алтернативни източници на енергия, в това число и термоелектрическите генератори (ТЕГ), известни още с наименованието - модули на Зеебек.

Модулите на Зеебек се различават от модулите на Пелтие по състава на материалите, от които са изработени. За ТЕГ се използват високотемпературни полупроводници, легирани със съответните примеси. Основните материали с висока термоелектрическа ефективност, респективно висок качествен фактор  $Z_0$ , използвани за производство на ТЕГ са съединенията  $\text{BiTe}$ ;  $\text{PbSnTe}$ ;  $\text{SnTeSe}$  и др.

Целта е да представи невронна мрежа за предсказване на генерираното напрежение от термоелектрическа батерия (ТЕБ), чиито основен възел е модул на Зеебек, при различен температурен градиент –  $(0\div 60)^\circ\text{C}$ . За нуждите на моделирането са използвани реални данни от предварително проведен експеримент. Изчислен е коефициента на полезно действие  $\eta$ .

$$\eta = \frac{T_h - T_l \sqrt{1 + \frac{Z(T_h + T_l)}{2}} - 1}{T_h \sqrt{1 + \frac{Z(T_h + T_l)}{2}} + \frac{T_l}{T_h}}$$

където:  $T_h$  - висока температура;  $T_l$  - ниска температура;  $Z$  - коефициент "стойност на качеството", дефинирана като:

$$Z = \frac{\alpha^2 \sigma}{k}$$

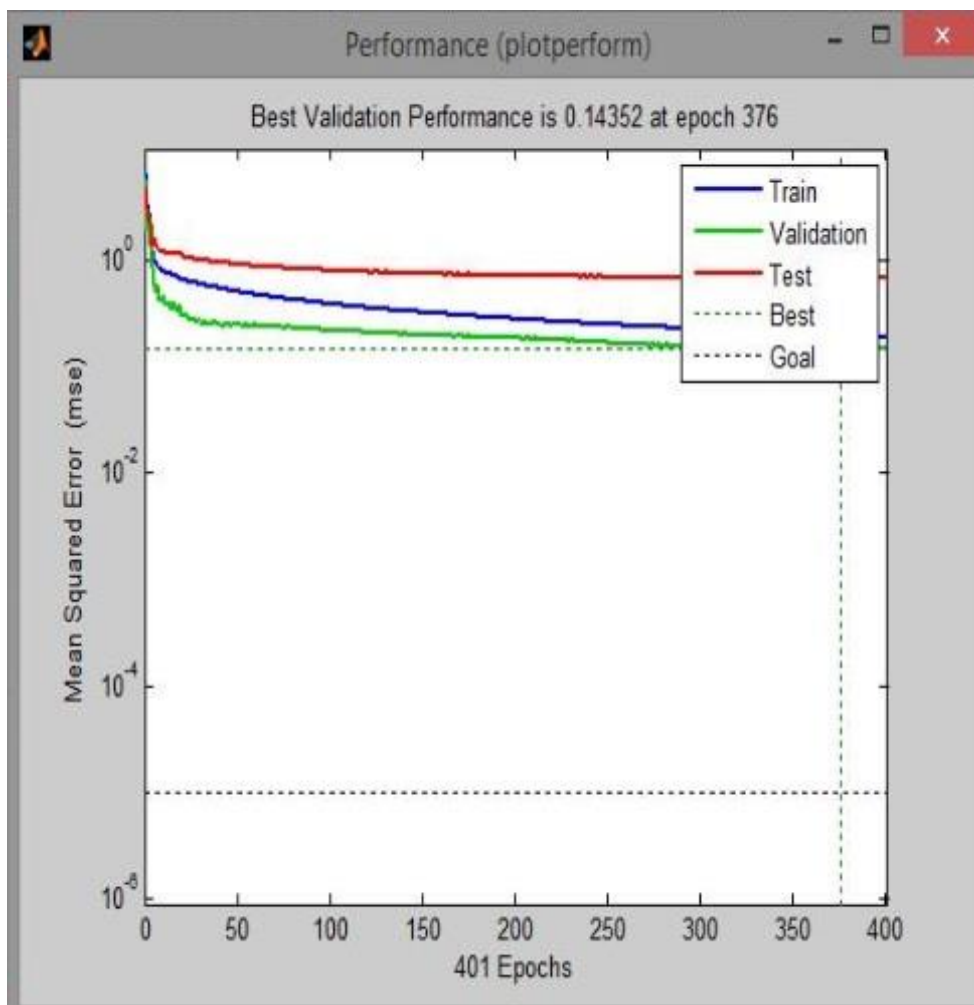
където:  $\alpha$  - коефициент на Зеебек;  $\sigma$  - електрическа проводимост;  $k$  - топлопроводимост.

Целта тук бе да се използва контролирана изкуствена невронна мрежа, за да се разпознае генерираното електродвижещо напрежение на термоелектрически модул на базата на температурна разлика  $\Delta T$  от модула като вход вектор. Целта е електродвижещото напрежение  $E$ . Има един вид модул и за този вид има 30 измервания. Таблица 3.1.1 показва векторите, използвани за трениране на невронната мрежа.

|                            |      |      |      |      |      |      |      |      |
|----------------------------|------|------|------|------|------|------|------|------|
| $\Delta T, ^\circ\text{C}$ | 0    | 2    | 4    | 6    | 8    | 10   | 12   | 14   |
| $E, \text{V}$              | 0    | 0,16 | 0,31 | 0,44 | 0,61 | 0,84 | 0,9  | 0,96 |
| $\Delta T, ^\circ\text{C}$ | 16   | 18   | 20   | 22   | 24   | 26   | 28   | 30   |
| $E, \text{V}$              | 0,99 | 1,11 | 1,2  | 1,31 | 1,38 | 1,47 | 0,55 | 1,63 |
| $\Delta T, ^\circ\text{C}$ | 32   | 34   | 36   | 38   | 40   | 42   | 44   | 46   |
| $E, \text{V}$              | 1,72 | 1,83 | 1,92 | 2,01 | 2,1  | 2,22 | 2,31 | 2,39 |
| $\Delta T, ^\circ\text{C}$ | 48   | 50   | 52   | 54   | 56   | 58   | 60   |      |
| $E, \text{V}$              | 2,44 | 2,51 | 2,61 | 2,72 | 2,81 | 2,92 | 3,03 |      |

Таблица 3.1.1

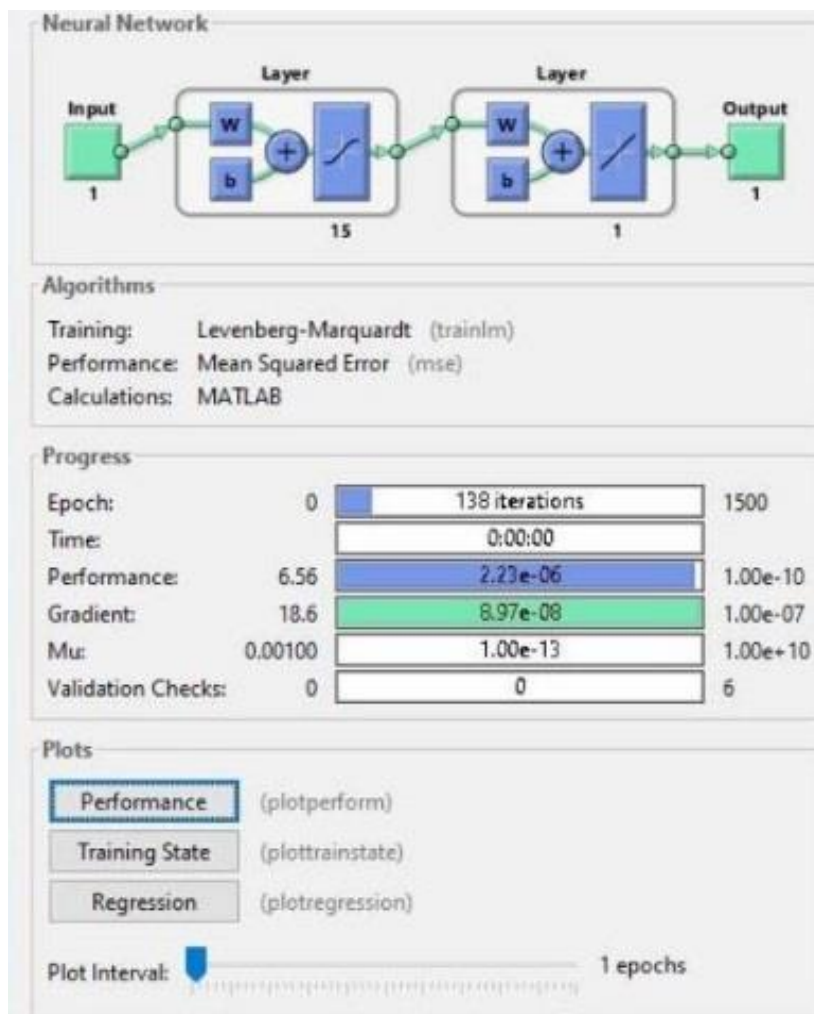
За симулиране на невронната мрежа се използва средата MATLAB. Процесът на обучение е показан на Фигура 3.1.4. Структурата на невронната мрежа е 1: 15: 1 (1 вход, 15 скрити неврона и 1 изход) и е показана на фигура 3.1.5.  $\Delta T$  се използва за входните данни. За изходните данни се използва  $E$ . Данните, събрани от ТЕГ, се използват за създаване на двойки, които съдържат разлика в температурата и EMF напрежение.



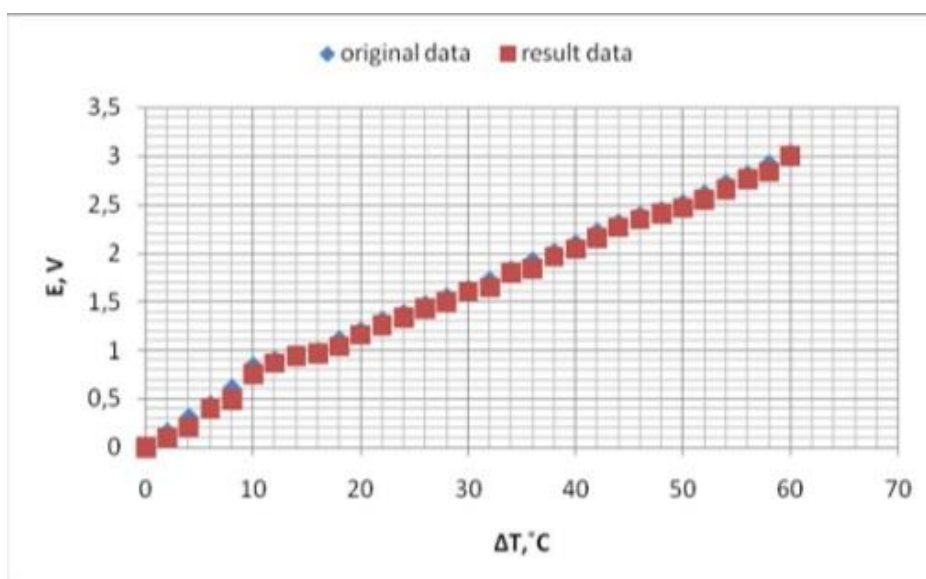
Фигура 3.1.4 Процесът на обучение

### 3.1.1. Резултати при обучението на невронната мрежа

Обучената невронна мрежа е показана на Фигура 3.1.5. Тук сме задали 1500 итерации на мрежата за обучение. Фигура 3.1.5 показва, че нашата цел за изпълнение е изпълнена само с 138 повторения от 1500 повторения. Показват се и други параметри на мрежата. За прекратяване на обучението се използват големината на градиента и броят на проверките за валидиране. Градиентът има тенденция да става много малък, когато тренировката достигне минимални показатели. Обучението спира, когато величината на градиента е под  $1e-7$ . Броят на последователните повторения, които ефективността на валидирането не намалява, се представя от броя на проверките за валидиране. Когато това число достигне стойността по подразбиране (6), обучението спира.



Фигура 3.1.5 Обучение на невронната мрежа



Фигура 3.1.6 Показва оригиналните данни от ТЕГ модула и резултатите от тестваните данни

Основната цел на разработката бе да се обучи невронна мрежа за моделиране на генерираното електромоторно напрежение на термоелектрически модул въз основа на температурната разлика  $\Delta T$ . Създадена е и е обучава невронна мрежа с реални експериментални данни. Фигура 3.1.6 показва, че така обучената невронна мрежа може да се приложи за прогнозиране на ЕМП на модул Seebeck.

### 3.2 Стохастично спускане по градиента в комбинация с метод на Отпадане

Стохастичното спускане по градиента се отнася за алгоритми за оптимизация и често се използва за конфигуриране на параметри на модела за машинно обучение. При стандартно Спускане по градиент се използва градиент за промяна на параметрите на модела. Градиентът обикновено се изчислява като сумата от градиентите, получени от всеки учебен елемент. Векторът на параметрите се променя в посока на антиградиента с дадена стъпка. Следователно, стандартното Спускане по градиента изисква едно преминаване през тренировъчните данни, преди да може да промени параметрите. При Стохастичното спускане по градиента стойността на градиента се изчислява приблизително от наклона на функция на загубата, изчислена само за един учебен елемент. Тогава параметрите се променят пропорционално на приблизителния наклон. По този начин параметрите на модела се променят след всеки учебен обект. За големи набори от данни Стохастичното градиентно спускане може да осигури значително предимство на скоростта спрямо стандартното спускане по градиента.

Предложена е комбинация на двата алгоритъма – Стохастично спускане по градиента и метод на Отпадане, като стремежа е да се взаимства положителната част на всеки от методите. В следствие, на което бе предложен нов алгоритъм.

Обобщеномрежовия модел включва следните преходи:

$$A = \{Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7\},$$

$Z_1$  – обобщаване на случаен вектор за тегловната матрица  $W$ ,

$Z_2$  – изчисляване на  $avg_K$ ;

$Z_3$  – изчисляване на градиента;

$Z_4$  – изчисляване на изходана реакция  $a_k = F_K(n_k)$  от  $k$ -тия слой;

$Z_5$  – определяне на разликата между получената стойност ( $S_O$ ), фиксираната учебна цел при обучението и най – малката квадратична грешка;

$Z_6$  – определяне дали невронната мрежа е обучена или не;

$Z_7$  – изчисляване на новите тегловни коефициенти;

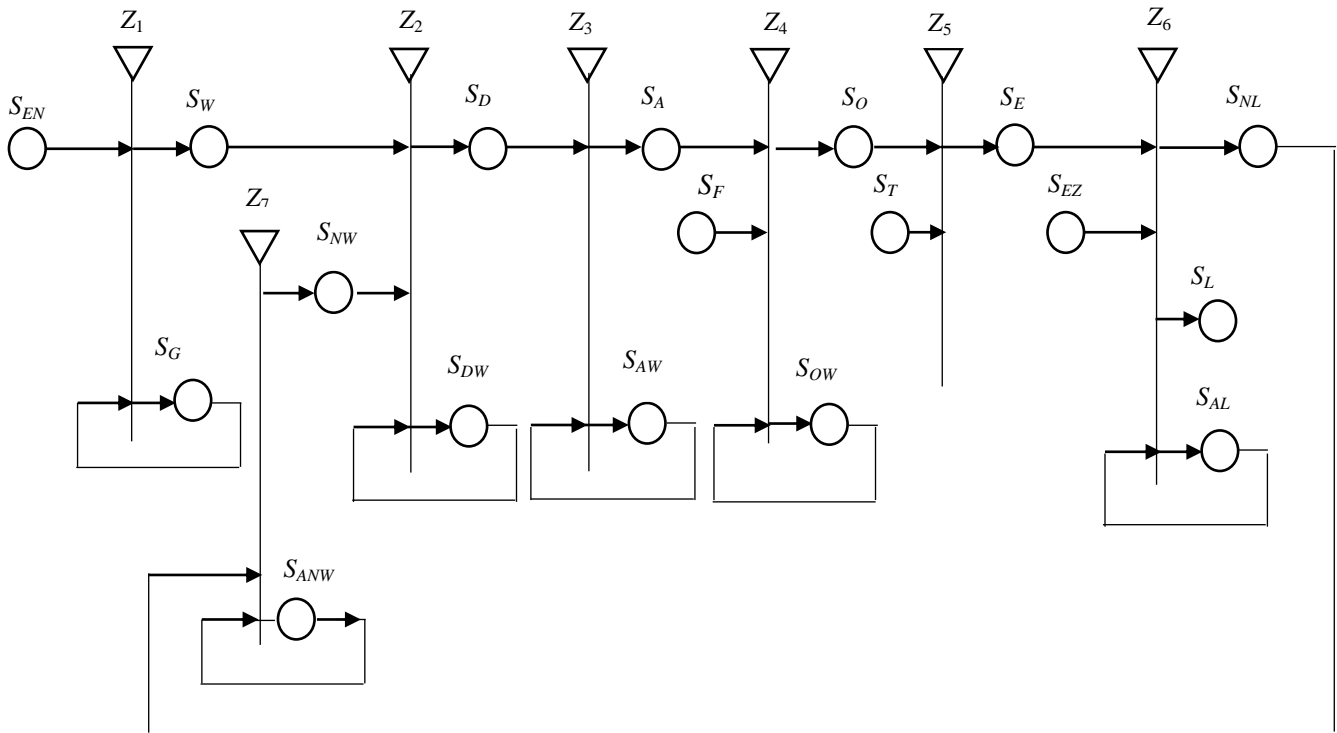
Мрежата съдържа следните ядра:

В позиция  $S_G$  - едно  $\alpha_G$  – ядро с характеристика "Генератор на случайни числа" за тегловните коефициенти,

Във всяка позиция  $S_F$  - едно  $\alpha_i$  - ядро,  $1 \leq i \leq k$ , с характеристика "Предавателна функция от  $i$ -тия слой на невронната мрежа",

В позиция  $S_T$  - едно  $\alpha_t$  - ядро с характеристика "Учебна цел за изходната реакция на невронна мрежа",

В позиция  $S_{EZ}$  - едно  $\alpha_{ez}$  - ядро с характеристика "Предварително фиксирана грешка в обучението по невронна мрежа".

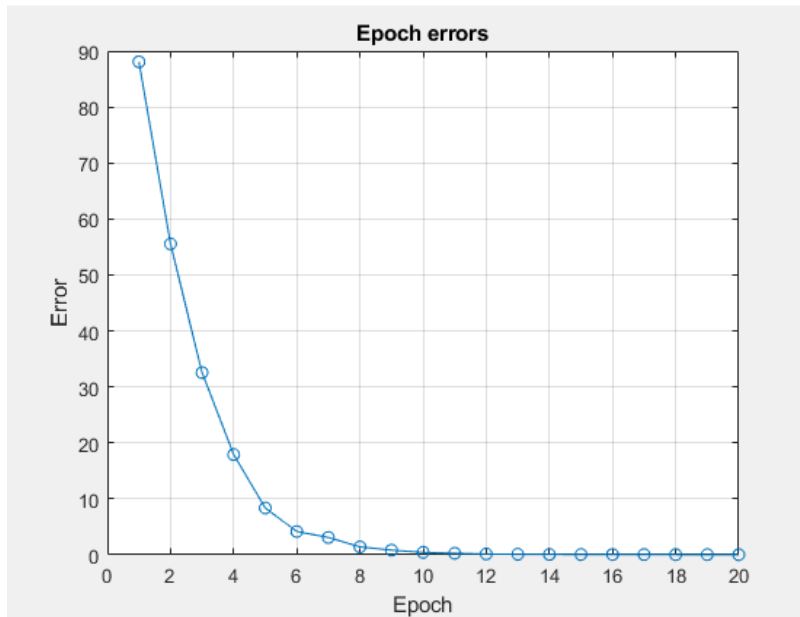


Фигура 3.2.1 Обобщеномрежови модел на Стохастично спускане по градиента в комбинация с метод на Отпадането

### 3.2.1. Резултати при обучението на невронната мрежа

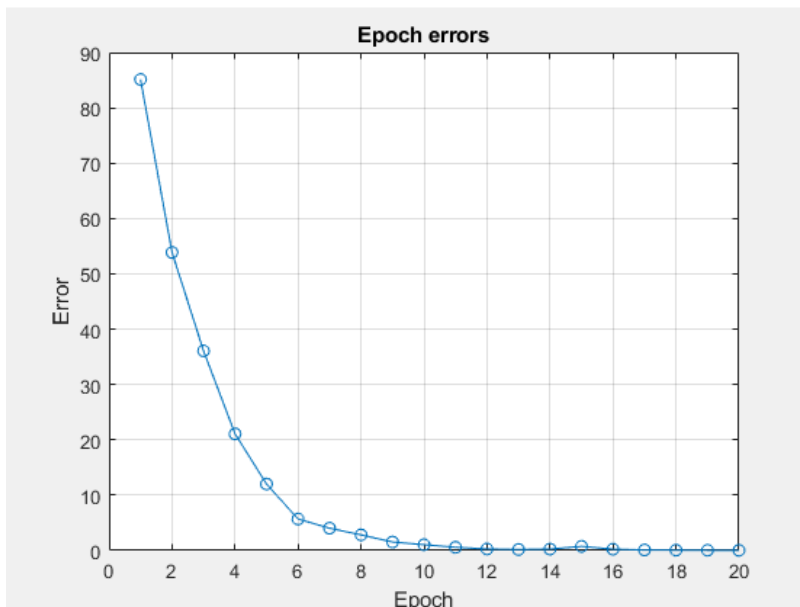
Направен е експеримент с различни стойности за вероятността даден неврон да бъде изключен при тренировъчния процес.

На Фигура 3.2.2 е показана диаграма, която изобразява стойностите на квадратичната грешка при вероятност за отпадане 0.1.



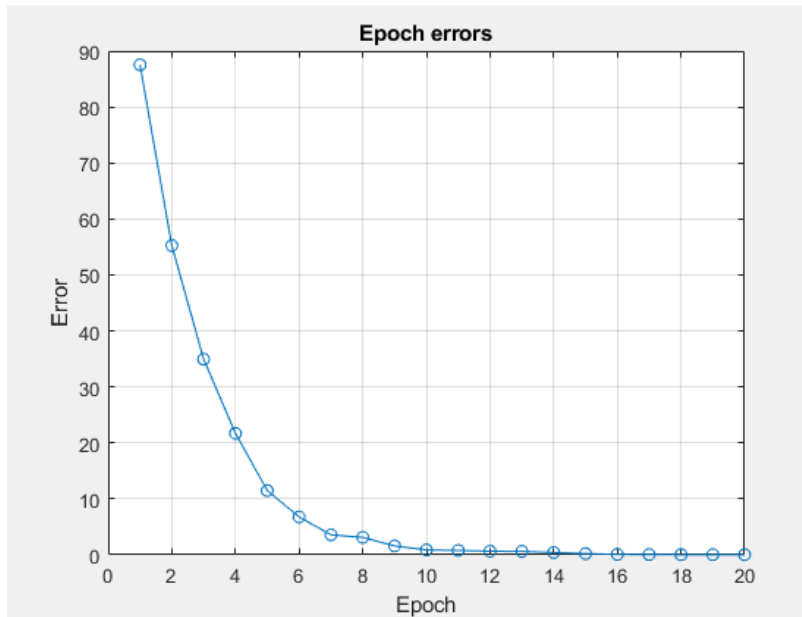
Фигура 3.2.2

На Фигура 3.2.3 е показана диаграма, която изобразява стойностите на квадратичната грешка при вероятност за отпадане 0.5.



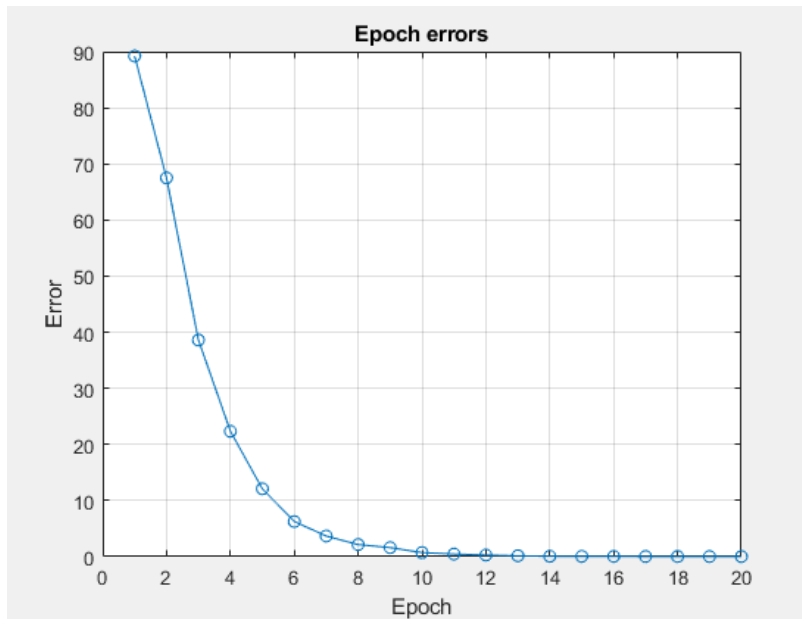
Фигура 3.2.3

На Фигура 3.2.4 е показана диаграма, която изобразява стойностите на квадратичната грешка при вероятност за отпадане 0.9.



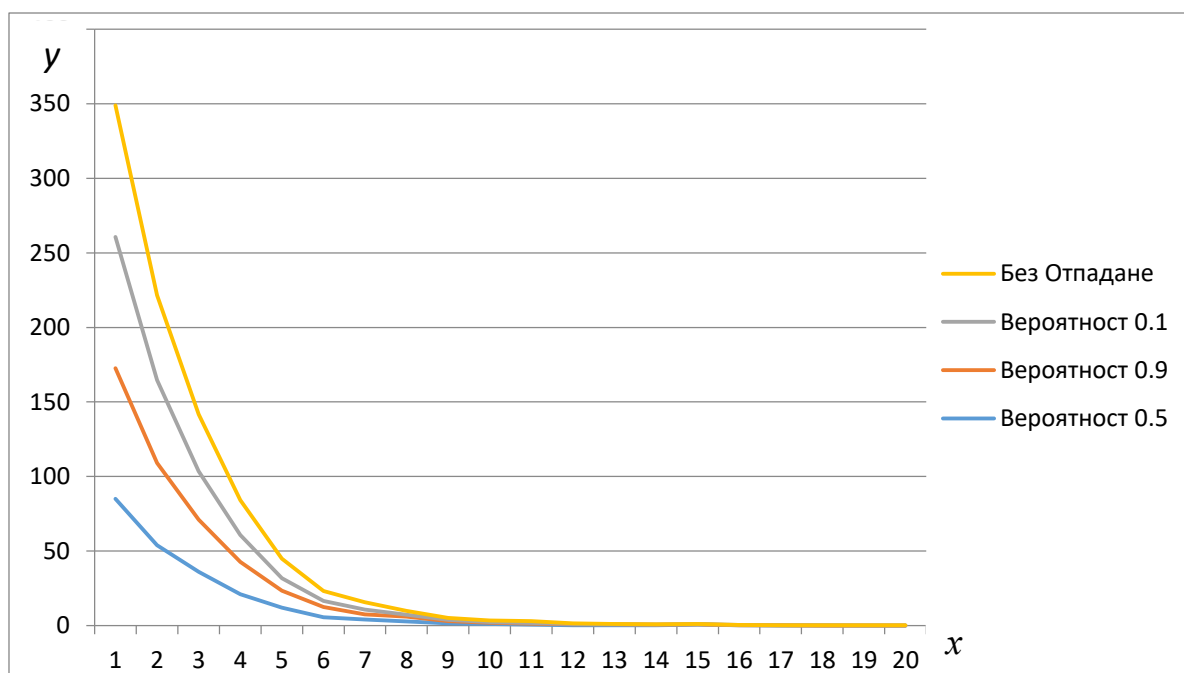
Фигура 3.2.4

На Фигура 3.2.5 е показана диаграма, която изобразява стойностите на квадратичната грешка при прилагане само на Стохастично спускане по градиента.



Фигура 3.2.5





Фигура 3.2.6

На следната диаграма са отбелязани стойностните на квадратичната грешка след всяка епоха. При използване на алгоритъма Стохастично спускане по градиента без комбинация с метод на Отпадане, наблюдаваме по-високи стойности на квадратичната грешка, след всяка епоха, в сравнение с останалите разгледани случаи. Квадратичната грешка е нашия критерий за спиране на обучението.

Тук влияят основно два фактора. Броят неврони, които да обучим, за да може да се възприеме информацията и броят неврони, които "генерират" грешка при използването си.

При първоначалния брой на невроните и съответно тегловните коефициенти към тях информацията се разполага във всички тегловни коефициенти. При началото на обучението грешката е много голяма, защото броя на тегловните коефициенти е голям. При по-малък брой тегловни коефициенти началната грешка е по-малка. При редуциране на 10% и на 90% от тях грешката е подобна поради това, че при премахването на тегловните коефициенти (90%) ние премахваме и грешката, която се натрупва в тях. Оптималния вариант е при вероятност на отпадане 0.5, защото наличната информация успява да се „помести“ в половината от определения брой неврони и паралелно с това да премахне грешката на половината неврони.

Резултатите, при обучение с метод на Отпадане с вероятност 0.5, са по-добри в сравнение с резултатите от останалите експерименти поради, което ви препоръчваме да използвате именно тази вероятност.

### 3.3. Обобщеномрежов модел на Стохастично спускане по градиента в Комбинация с метод на Отпадане с Интуиционистки размити оценки

Стохастичното градиентно спускане е много популярен и често срещан алгоритъм, използван в различни алгоритми за машинно обучение. Градиентното спускане е метод за намиране на локален екстремум (минимум или максимум) на функция чрез движение по градиента. Отпадането работи, като изключва невроните в мрежа по време на обучение, за да принуди останалите неврони да поемат товара на липсващите неврони. Това обикновено се прави на случаен принцип, като се изключва определен процент неврони на слой.

Разгледан е нов Обобщеномрежи модел, симулация на процеса на обучение на невронна мрежа комбинираща Метод на Отпадането и Стохастично спускане по градиента. Разработката показва, че процесът може успешно да се осъществи чрез инструмента на обобщените мрежи. Моделът дава възможност да се разгледат различните етапи при обучението на невронната мрежа. Използвана е оценка с интуиционистки размити множества. Интуиционистки размитите оценки отразяват резултатите на системата. Те могат да повлияят или не на системата. Също така е разгледана степен на неопределеност, в случай, че информацията не е достатъчна. Обобщеномрежовия модел включва следните преходи:

$$A = \{Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7 \},$$

- $Z_1$  – генериране на случаен вектор на матрицата за тегловните коефициенти  $W$ ;
- $Z_2$  – изчисляване на  $avg_k$ ;
- $Z_3$  – изчисляване на градиента;
- $Z_4$  – пресмятане на изходна реакция  $a_k = F_k(n_k)$  за  $k$ -тия слой;
- $Z_5$  – определяне на разликата между получената стойност ( $S_O$ ) и фиксираната учебна цел и най-малката квадратична грешка;
- $Z_6$  – определяне дали невронната мрежа е обучена или не;
- $Z_7$  – изчисляване на новите тегловни коефициенти.

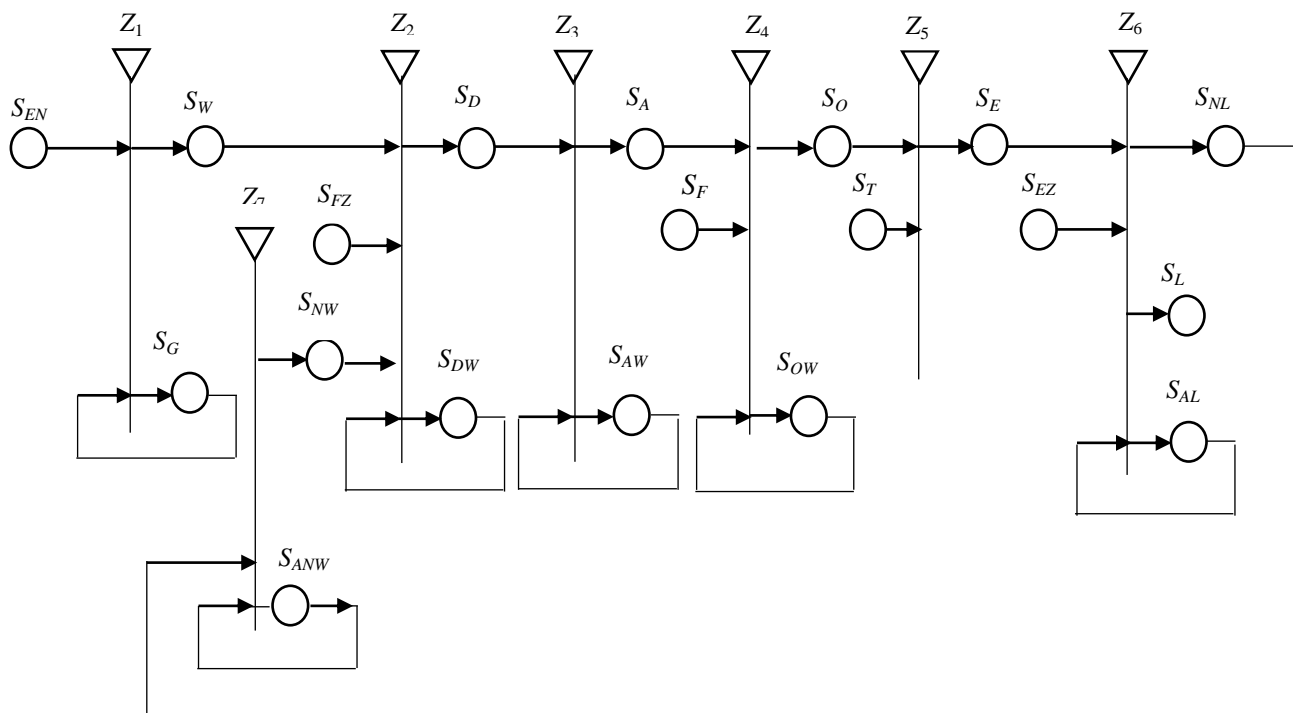
Мрежата съдържа следнит ядра:

В позиция  $S_G$  - едно  $\alpha_G$  – ядро с характеристика "Генератор на случайни числа" за тегловните коефициенти,

Във всяка позиция  $S_F$  едно  $\alpha_i$  - ядро,  $1 \leq i \leq k$ , с характеристика "Предавателна функция от  $i$ -тия слой на невронната мрежа",

В позиция  $S_T$  - едно  $\alpha_i$  – ядро с характеристика "Учебна цел за изход на невронна мрежа",

В позиция  $S_{EZ}$  - едно  $\alpha_{ez}$  – ядро с характеристика "Предварително фиксирана грешка в обучението по невронна мрежа ".



Фигура 3.3.1. Обобщеномрежов модел на Стохастично спускане по градиента в комбинация с метод на Отпадане с интуиционистки размити оценки

### 3.4. Изводи

В третата главата са разработени обобщеномрежови модели описващи приложението на невронните мрежи за различни проблеми.

В първия модел се разглежда приложението на невронна мрежа и термоелектричния генератор. За целта е обучена невронна мрежа за моделиране на генерираното електромоторно напрежение на термоелектрически модул въз основа на температурната разлика  $\Delta T$ . Създава се и се обучава невронна мрежа с реални експериментални данни.

При втория модел се разглежда приложението на невронна мрежа обучена чрез Стохастично спускане по градиента в комбинация с метод на Отпадането. От получените резултати се вижда се, че двата алгоритъма успешно постигат целите си, като по този начин чрез тяхното комбинирание могат да бъдат решени различни проблеми. Процесът е описан чрез помощта на обобщена мрежа.

Последният модел описва обучението на невронна мрежа обучена чрез Стохастично спускане по градиента в комбинация с метод на Отпадането и оценена с интуиционистки размита оценка. Това описание дава възможност да се разгледат детайлно различните етапи при обучението на невронната мрежа. Използвана е оценка с интуиционистки размити множества. Също така е разгледана степен на несигурност, в случай, че информацията е недостатъчна.

### Заклучение

Научно-изследователската работа е проведена в сферата на невронните мрежи. Проблемната област за изследването е моделирането на невронни мрежи. Водещият

научен въпрос на дисертацията е „Как се създава и оптимизира невронна мрежа?“. В търсене на отговора на този въпрос са изследвани и моделирани различни невронни мрежи с помощта на Обобщените мрежи.

Настоящата дисертация представлява научен анализ на текущото състояние в областта на моделирането на невронни мрежи. Създаден е обобщеномрежов модел описващ структурата на невронната мрежа Неокогнитрон. Основно се разглежда действието на всеки слой и ориентиращата подсистема. Алгоритъма, използван за тази структура, се формира като многослойна мрежа. Самоорганизацията с неконтролирано обучение с помощта на повтарящи се модели е необходим елемент от мрежата. Разгледан е обобщеномрежов модел на невронна мрежа за дълбоко обучение, която дава възможност правилно да се идентифицира изображение, разпознаване на реч и обработка на език.. Моделът дава възможност да се разгледат действието на слоевете и ориентиращата подсистема. Друг разгледан алгоритъм е метод на Отпадането. Алгоритъмът е описан детайлно чрез инструмента на обобщените мрежи. Трябва да се отбележи, че големите многослойни мрежи са бавни и прилагането на този метод е една добра техника за справянето с този проблем. Реализиран е модел за разпознаване на пръстов отпечатък Моделът дава възможност да се разгледат различните етапи на идентификация на пръстови отпечатьци. Отпечатъкът е една от най-уникалните части на човешкото тяло и затова се използва в устройства за идентификация. За целта е използвана оценка с интуиционистични размити множества. Получените интуиционистични размити оценки отразяват разпознаването на системата.

Принос на работата е създаването на обобщеномрежов модел описващ структурата на невронна мрежа обучена със Стохастично спускане по градиента в комбинация с метод на Отпадането. Получените резултати при разработката показват, че двата алгоритъма успешно постигат целите си, като по този начин чрез тяхното комбиниране могат да бъдат решени различни проблеми. Също така е използвана оценка с интуиционистки размити множества. Разгледана е степен на несигурност, в случай, че информацията не е достатъчна. Разгледано е приложението на невронна мрежа и термоелектричния генератор, като за целта е обучена невронна мрежа за моделиране на генерираното електромоторно напрежение на термоелектрически модул въз основа на температурната разлика  $\Delta T$ . Създадена е и е обучена невронна мрежа с реални експериментални данни.

### **Приноси към дисертационния труд**

Основните приноси в дисертацията са с научно-приложен характер и се свеждат до предлагането на нови модели. Приносите с научно-приложен характер могат да се формулират като разработване на:

- Обобщеномрежов модел на неокогнитрон;
- Обобщеномрежов модел на дълбока невронна мрежа;
- Обобщеномрежов модел на метод на обучение на невронни мрежи - Метод на Отпадане;
- Обобщеномрежови модел за разпознаване на пръстов отпечатък с интуиционистки размити множества.

Приносите с приложен характер представляват разработване на програми за:

- симулиране поведението на мощността на термоелектричния генератор на базата на изкуствена невронна мрежа;
- симулиране поведението на Обобщеномрежови модел на стохастично спускане по градиента в комбинация с метод на отпадането.

Има разработен програмен код в средата Matlab за невронните мрежи:

- Предсказване на мощността на термоелектричния генератор на базата на изкуствена невронна мрежа;
- Модел на Стохастично спускане по градиента в комбинация с метод на Отпадане.

### **Насоки за бъдещо развитие**

Разработените обобщеномрежови модели дават възможност да се формулират различни идеи за техни бъдещи приложения. Например разработването на обобщеномрежови модел на процеса на разпознаване на изображение с помощта на предложени алгоритъм, също както и оптимизация на алгоритма.

### **ПУБЛИКАЦИИ СВЪРЗАНИ С ДИСЕРТАЦИЯТА**

- 1\*. Belovski, I., P. Yovcheva, S. Surchev. Thermoelectric Generator Power Prediction Based on Artificial Neural Network, Proceedings of the 20th International Symposium on Electrical Apparatus and Technologies (SIELA 2018), 54 - 57.
- 2\*. Bureva, V., P. Yovcheva, S. Sotirov. Generalized Net Model of Fingerprint Recognition with Intuitionistic Fuzzy Evaluations, Proceedings of the 10th Conference of the European Society for Fuzzy Logic and Technology, and the Sixteenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, September 13–15, 2017, Warsaw, Poland, Volume 1, 286 - 294.
- 3\*. Petkov, T.P. Jovcheva, Zh. Tomov, S. Sotirov. A generalized net model of the Neocognitron neural network, Proceedings of the International Conference on Flexible Query Answering Systems, Lecture Notes in Computer Science, Volume 10333, 2017, 249-259.
- 4\*. Yovcheva, P., S. Sotirov. A Generalized Net Model of the Stochastic gradient descent and Dropout algorithm with Intuitionistic fuzzy evaluations. Notes on Intuitionistic Fuzzy Sets, 26, No. 4, 2020 (in press)
- 5\*. Yovcheva, P., S. Sotirov, K. Atanassov. A generalized net model of the Deep learning algorithm. Proceedings of ANNA '18; Advances in Neural Networks and Applications 2018, 59 - 63.
- 6\*. Йовчева П., С. Сотиров. Обобщеномрежови модел на Стохастично спускане по градиента в комбинация с метод на Отпадането, Годишник на секция „Информатика” – Съюз на учените в България, том 10, 2019/2020 (под печат).
- 7\*. Yovcheva, P., S. Sotirov. A generalized net model of the dropout algorithm. Springer, Cham (in press).