

Система распознавания семиографических песнопений

Одной из важных и в высшей степени интересных задач в изучении музыкальной культуры древней Руси является раскрытие мелодического содержания рукописных певческих книг XII—XVII веков, в большом количестве сохранившихся до наших дней.

Мелодии в этих книгах записаны посредством безлинейных нотных систем, вырабатывавшихся на Руси на протяжении нескольких столетий. Древние нотации давно уже вышли из употребления и забыты, но с развитием семиотики и музыкальной палеографии они все чаще становятся объектами исследований.

Слово «семиография» означает известный условно принятый способ письма и выражения определенных музыкальных звуков и их взаимоотношений. Фрагмент семиографического песнопения представлен на рис.1. Структурно песнопение состоит из следующих частей:

- строк знамен (семиографических знаков), обозначающих нотные знаки;
- строк текста, которым сопоставляются знамена;
- в структуре знамен можно особо выделить пометы, обозначающие длительность или высотность нот;

Разработанная система предназначена для автоматизации процесса распознавания данного типа рукописей. На текущем этапе основной задачей явилось распознавание наиболее простой части песнопений – помет, общее число которых не превышает десяти.

В статье основное внимание уделено описанию математических моделей и алгоритмов, легших в основу Системы распознавания семиографических песнопений (СРСП). В заключение приводится их сравнение для решения задач распознавания помет.

В качестве теоретической базы выбран математический подход к распознаванию, который основывается на математических правилах классификации, которые формулируются и выводятся в терминах математической статистики.

В качестве методов распознавания были выбраны нейронные сети (НС) трех типов:

- Многослойный перцептрон
- НС с общей регрессией
- Вероятностная НС

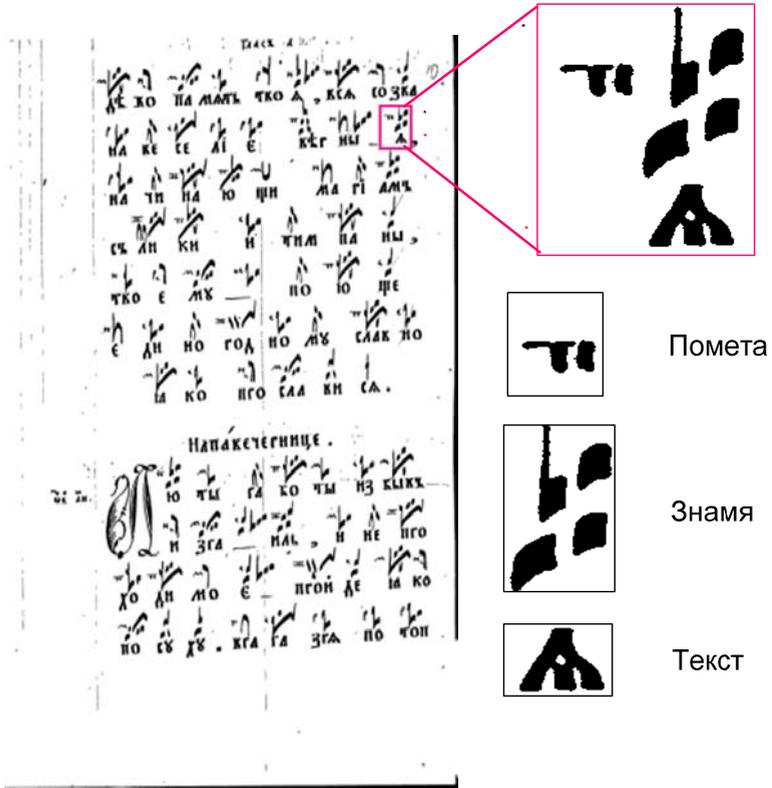


Рис. 1. Структура семиографического песнопения

Многослойный перцептрон

В качестве теоретического основания этой модели НС выступают теорема А.Н.Колмогорова, которая гласит, что любую непрерывную функцию на d -мерном кубе можно представить в виде суперпозиции функций двух переменных, и теорема Новикова (применительно к перцептрон Розенблатта).

Теорема А.Н.Колмогоров, 1957. Для любой непрерывной функции на d -мерном кубе существует вычисляющая ее трехслойная нейронная сеть с $(d+1)(2d+1)$ скрытыми нейронами.

Теорема Новикова. Пусть задана бесконечная последовательность примеров с элементами, удовлетворяющими неравенству $|u_i| < D$, где U - пространство признаков. Предположим, что существует гиперплоскость с коэффициентами W_0 , которая точно разделяет элементы последовательности и удовлетворяет условию

$$\min_{(y,u) \in \{\hat{Y}, \hat{U}\}} \frac{y(w_0 * u)}{|w_0|} \geq \rho_0, \text{ где } \left\{ \hat{Y}, \hat{U} \right\} = (y_1, u_1), \dots, (y_l, u_l),$$

$\rho_0 > 0$

Затем, используя итеративно процедуру

$$w(t) = \begin{cases} w(t-1), & y_i(w(t-1) * u_i) > 0 \\ w(t-1) + y_i u_i, & \text{else} \end{cases},$$

персептрон строит

гиперплоскость, которая корректно разделяет все примеры бесконечной последовательности. Для создания такой гиперплоскости персептрону требуется $M = \left[\frac{D^2}{\rho_0^2} \right]$ коррекции.

На рис. 2 представлена типовая схема персептронной сети, реализация которой описывается следующим формализмами:

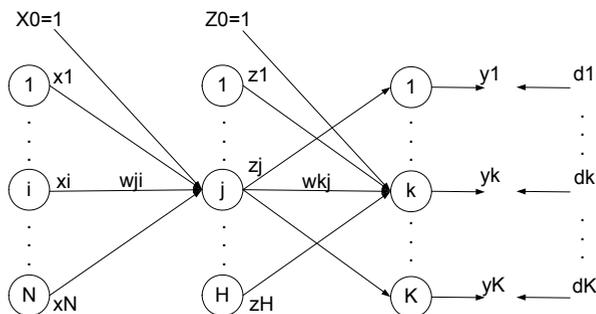


Рис. 2. Структура персептрона

где X_i – i -ый нейрон входного слоя, Z_j – j -ый нейрон скрытого слоя, Y_k – k -ый нейрон выходного слоя, D_k – k -ый желаемый выход.

$$net_j = \sum_{i=0}^N w_{ji} x_i$$

Состояние j -го нейрона скрытого слоя: , вы-

ход j -го нейрона скрытого слоя: $z_j = f_h(net_j)$

$$net_k = \sum_{j=0}^H w_{kj} z_j$$

Состояние k -го нейрона выходного слоя:

выход k -го нейрона выходного слоя: $y_k = f_o(net_k)$

$$E(w) = \frac{1}{2} \sum_{k=1}^K (d_k - y_k)^2$$

Расчет ошибки сети:

Пересчет весов для выходного слоя:

$$\Delta w_{kj} = -\eta \frac{\partial E}{\partial w_{kj}} = \eta (d_k - y_k) f'_o(net_k) z_j$$

Пересчет весов для скрытого слоя:

$$\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial z_j} \frac{\partial z_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}}$$

$$\frac{\partial E}{\partial z_j} = -\sum_{k=1}^K (d_k - y_k) \frac{\partial y_k}{\partial z_j} = -\sum_{k=1}^K (d_k - y_k) f'_o(net_k) w_{kj}$$

$$\Delta w_{ji} = \eta \left[\sum_{k=1}^K (d_k - y_k) f'_o(net_k) w_{kj} \right] f'_h(net_j) x_i$$

Обучение и работа сети

Обучение НС осуществляется с помощью алгоритма обратного распространения ошибки, который схематично представлен на рис. 3

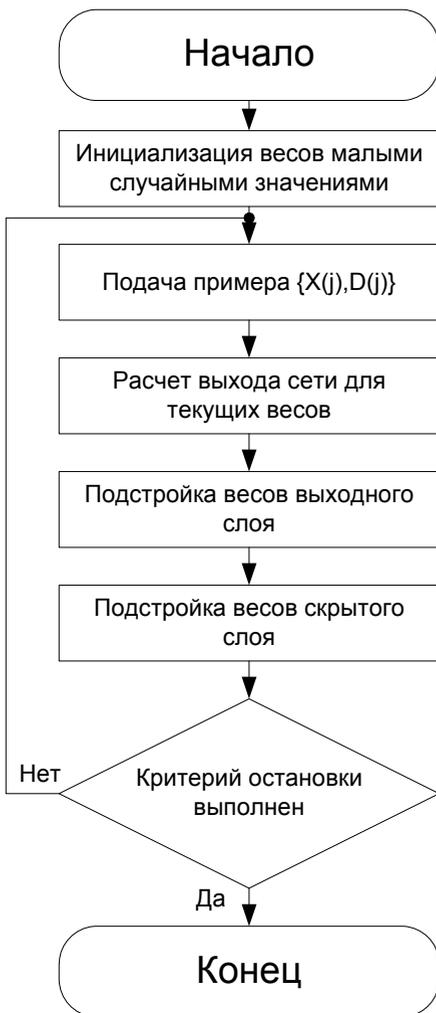


Рис. 3. Алгоритм обратного распространения

Шаг 1. Инициализация весов малыми случайными значениями.

Шаг 2. Задать значения входов $\{x_1, \dots, x_N\}$ и соответствующих желаемых выходов $\{d_1, \dots, d_K\}$

Шаг 3. Посчитать ответ сети для текущих значений весов

$$y_k = f_o\left(\sum_{j=0}^H w_{kj} f_h\left(\sum_{i=0}^N w_{ji} x_i\right)\right)$$

Шаг 4. Построить веса выходного слоя

$$\Delta w_{kj}(t) = \eta \delta f'_o(net_k) z_j + \mu \Delta w_{kj}(t-1), \text{ где } \delta_k = d_k - y_k$$

Шаг 5. Подстроить веса скрытого слоя

$$\Delta w_{ji}(t) = \eta \delta f'_h(net_{kj}) x_i + \mu \Delta w_{ji}(t-1), \text{ где}$$

$$\delta_j = \sum_{k=1}^K (d_k - y_k) f'_o(net_k) w_{kj}$$

Шаг 6. Если критерий остановки не выполнен, то Шаг 2, иначе СТОП.

Нейронная сеть с общей регрессией

НС с общей регрессией была впервые представлена Надараяем и Ватсоном. Регрессионная функция, выполненная на независимой переменной X , вычисляет наиболее вероятное значение зависимой переменной Y , основанной на конечном множестве наблюдений X и соответствующих значений Y .

Пусть $f(x, y)$ совместная непрерывная функция плотности вероятности случайного вектора X и скалярного случайного значения переменной y . Пусть X - частное значение случайной переменной X . Тогда регрессия y данного X будет

$$E[y / X] = \frac{\int_{-\infty}^{+\infty} y * f(X, y) dy}{\int_{-\infty}^{+\infty} f(X, y) dy}$$

Обычно совместная плотность $f(x, y)$ неизвестна. Тогда оценочная функция будет иметь вид

$$f_n(x) = \frac{1}{n\lambda} \sum_{i=1}^n \varphi\left(\frac{x - x_i}{\sigma}\right), \text{ где } x_i - \text{ независимые одина-}$$

ково распределенные случайные переменные с абсолютно непрерывной функцией распределения. Весовая функция φ должна быть ограниченной и удовлетворять следующим условиям:

$$\int_{-\infty}^{+\infty} |\varphi(y)| dy < \infty ; \lim_{y \rightarrow \infty} |y\varphi(y)| = 0 ; \int_{-\infty}^{+\infty} \varphi(y) dy = 1$$

Функция $\sigma = \sigma(n)$ должна быть выбрана таким образом, чтобы

$$\lim_{n \rightarrow \infty} \sigma(n) = 0 ; \lim_{n \rightarrow \infty} n\sigma^2(n) = \infty$$

Одной полезной формой весовой функции φ является функция плотности ядра (Гауссиан). Парзен показал, что эти оценочные функции состоятельны.

$$\hat{f}(X, Y) = \frac{1}{(2\pi)^{\frac{(p+1)}{2}} \sigma^{p+1}} * \frac{1}{n} \sum_{i=1}^n \exp\left[-\frac{(X - X^i)^T (X - X^i)}{2\sigma^2}\right] * \exp\left[-\frac{(Y - Y^i)^2}{2\sigma^2}\right], \text{ где } p \text{ является размерностью вектора } X, \text{ а } n$$

количеством наблюдений.

Определим скалярную функцию D_i^2

$$D_i^2 = (X - X^i)^T (X - X^i)$$

и, подставив ее, получим

$$\hat{Y}(X) = \frac{\sum_{i=1}^N Y^i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^N \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}$$

Данная регрессия носит название оценочной функции Надарая-Ватсона. Оценка $\hat{Y}(X)$ может быть представлена как взвешенное среднее всех наблюдаемых значений Y^i .

Первая реализация сети впервые была представлена Спехтом и описывается ниже:

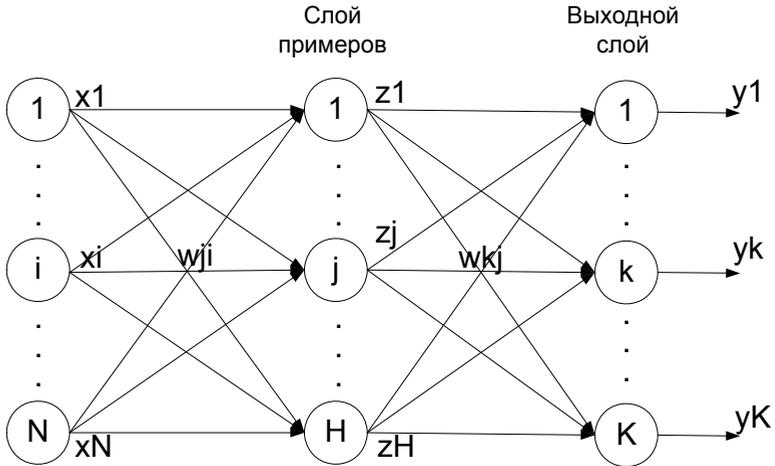


Рис. 4. Структура нейронной сети с общей регрессией

где X_i – i -ый нейрон входного слоя, Z_j – j -ый нейрон слоя примеров, Y_k – k -ый нейрон выходного слоя.

Выход j -го нейрона слоя примеров:

$$z_j = \exp\left(-\frac{\sum_{i=1}^N (w_{ji} - x_i)^2}{2\sigma^2}\right)$$

Выход k -го нейрона выходного слоя:

$$y_k = \frac{\sum_{j=1}^H w_{kj} z_j}{\sum_{j=1}^H z_j}$$

Вероятностная нейронная сеть

Вероятностная нейронная сеть, часто называемая байесовской, основана на аналогичных теоретических принципах, что и сеть с общей регрессией, т.к. в ней используются оценочные функции распределения вероятности Парзена. Правдоподобие неизвестного вектора, принадлежащего данному классу, может быть выражено как

$$f_i(x) = \frac{1}{(2\pi)^{p/2} \sigma^p} * \frac{1}{n_i} \sum_{i=1}^{n_i} \exp \left[-\frac{(x - x_{ij})^T (x - x_{ij})}{2\sigma^2} \right],$$

где i – номер класса, j – номер паттерна, n_i - количество обучающих векторов в i – классе, X - тестовый вектор.

Ниже представлена реализация сети

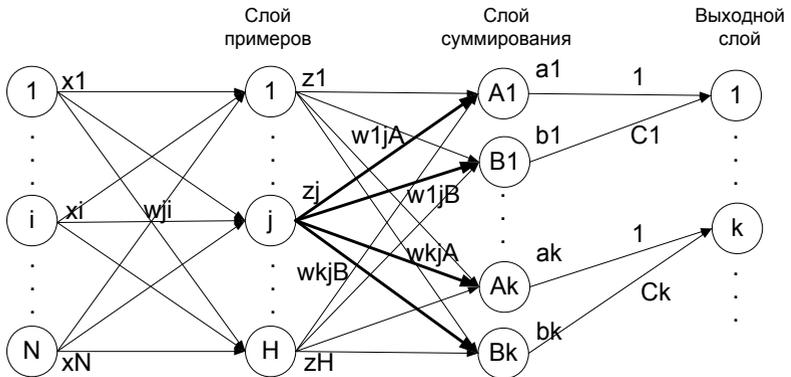


Рис. 5. Структура вероятностной нейронной сети

где X_i – i -ый нейрон входного слоя, Z_j – j -ый нейрон слоя примеров, A_k – k -ый нейрон слоя суммирования, принадлежащий данному классу, B_k – k -ый нейрон слоя суммирования, не принадлежащий данному классу, Y_k – k -ый нейрон выходного слоя.

Выход j -го нейрона слоя примеров:

$$z_j = \exp\left(-\frac{\sum_{i=1}^N (w_{ji} - x_i)^2}{2\sigma^2}\right),$$

$$a_k = \sum_{j=1}^H w_{kj}^A z_j,$$

Для всех нейронов слоя суммирования:

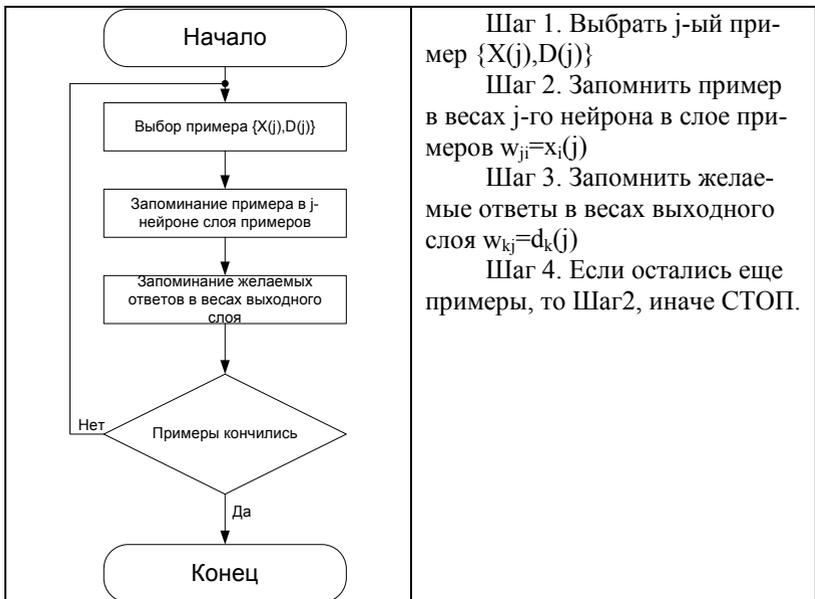
$$b_k = \sum_{j=1}^H w_{kj}^B z_j$$

Для всех нейронов выходного слоя:

$$y_k = \text{sign}(a_k + C_k b_k),$$

$$C_k = -\frac{l_{Bk} h_{Bk} n_{Ak}}{l_{Ak} h_{Ak} n_{Bk}}$$

Обучение вероятностных (байесовых) сетей осуществляется по следующему алгоритму:



Результаты обучения и сравнение нейронных сетей

На базе анализа литературы были выявлены достоинства и недостатки сетей, которые представлены в таблице ниже.

	Рыхлые данные	Малое количество данных	Локальные минимумы	Большое число входов
Персептрон	-	+	+	+
НС с общей регрессией	+	+	-	-
Вероятностная НС	+	+	-	-

В процессе разработки системы был проведен ряд сравнительных запусков системы при различных параметрах нейросетей, признаков и т.д.

Для персептрона скорость и момент обучения были выбраны на основе следующих данных. Полученных в результате экспериментов:

- Зависимость количества неверных ответов от скорости обучения (момент=0.9)
- Зависимость количества неверных ответов от момента обучения (скорость =0.01)

При этом следует отметить, что при одновременном изменении этих параметров в допустимых рамках, можно улучшить результат.

Например, при скорости обучения 0.1 и моменте 0.5 количество неверных ответов сети будет равно 0.1.

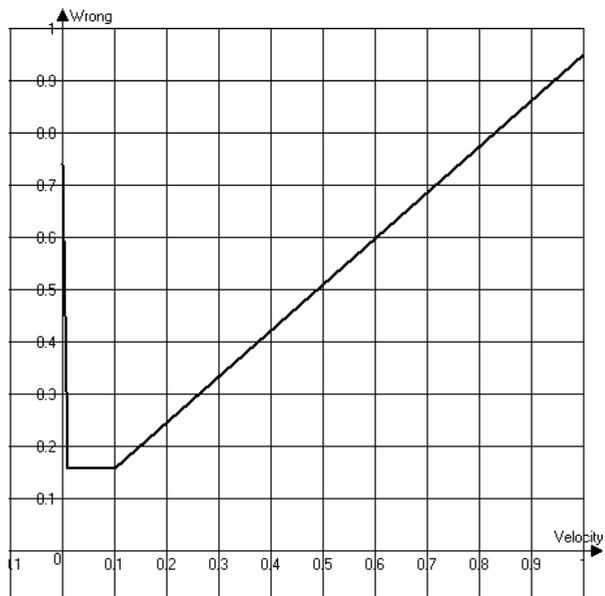


Рис. 6. Зависимость процента распознавания от скорости при постоянном моменте

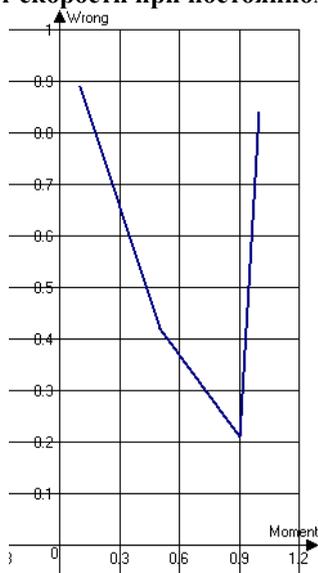


Рис. 6. Зависимость процента распознавания от момента при постоянной скорости

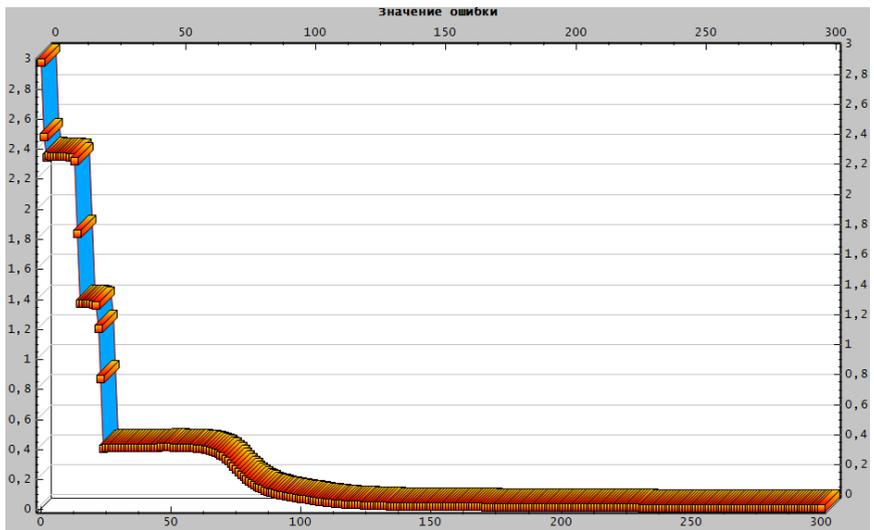


Рис. 7. Персептрон. Уменьшение ошибки в процессе обучения

Ниже приведены результаты работы сетей при количестве признаков=25

Вероятностная НС(1 эпоха)

Слой	Входной слой	Скрытый слой	Выходной слой
Нейронов	25	110	7
Символ	Среднее значение выхода	Максимальное значение выхода	Миним-ое значение выхода
С	77.77800735	99.99999987	33.33402255
М	99.99999994	100	99.99999981
Н	99.99999999	100	99.99999995
П	79.16640769	99.99999961	50.00310248
Р	99.99922433	100	99.99689771
В	99.99999884	99.99999884	99.99999884

Статистика распознавания:

- Правильно – 17
- Неправильно – 2

Перцептрон (300 эпох, скорость=0.01, момент=0.9)

Слой	Входной слой	Скрытый слой	Выходной слой
Нейронов	26	14	7
Функция	Экспоненциальная		
С	0.631956791	0.640731412	0.623182169
М	0.946009168	0.955919016	0.93682458
Н	0.768222496	0.940875576	0.14833814
П	0.777250872	0.900827998	0.469283307
Р	0.843146703	0.947454024	0.622591368
В	0.80032522	0.80032522	0.80032522

Статистика распознавания:

- Правильно – 16
- Неправильно – 3

НС с общей регрессией(1 эпоха)

Слой	Входной слой	Скрытый слой	Выходной слой
Нейронов	25	110	7
Символ	Среднее значение выхода	Максимальное значение выхода	Миним-ое значение выхода
С	0.777780073	0.999999999	0.333340226
М	0.999999999	1	0.999999998
Н	1	1	1
П	0.791664077	0.999999996	0.500031025
Р	0.999992243	1	0.999968977
В	0.999999988	0.999999988	0.999999988

Статистика распознавания:

- Правильно – 17
- Неправильно – 2

В результате было установлено, что при решении данной задачи лучше работают вероятностная нейронная сеть и сеть с общей регрессией. Следует отметить, что особенностью данной задачи являлось ограниченное и весьма небольшое количество примеров в обучающей выборке.

Также было замечено, что количество признаков не должно быть ни слишком большим, ни слишком малым, т.к. при этом классификация существенно затрудняется.