



где  $W^{(m)}$  – матрица весов слоя MAXNET,  $p$  – количество нейронов сети,  $\varepsilon$  – случайная величина с достаточно малой амплитудой.

По результатам тестирования разработанной системы можно сделать вывод, что сеть Хемминга превосходит сеть Хопфилда по точности распознавания образов.

Sign	/	%
Total	81 / 140	57,86
а	11 / 20	55
г	18 / 20	90
е	13 / 20	65
к	9 / 20	45
л	9 / 20	45
н	13 / 20	65
ч	8 / 20	40

Рис.1 Сеть Хопфилда

Sign	/	%
Total	108 / 140	77,14
а	15 / 20	75
г	18 / 20	90
е	19 / 20	95
к	15 / 20	75
л	15 / 20	75
н	16 / 20	80
ч	10 / 20	50

Рис. 2 Сеть Хемминга

Даже при сопоставимых результатах точности распознавания образов, сеть Хемминга превосходит сеть Хопфилда по рациональности использования ресурсов ЭВМ.

### Литература

1. Осовский, С. Нейронные сети для обработки информации [Текст]/ Пер. с польского И.Д. Рудинского – М.: Финансы и статистика, 2002. – 344 с.: ил.
2. Хайкин, С. Нейронные сети: полный курс, 2е издание [Текст]/ С. Хайкин – М.: Издательский дом "Вильямс", 2006 – 1104 с.

Е.А. Назарова, И.М. Куликовских

### НЕЧЕТКИЙ АЛГОРИТМ ОБУЧЕНИЯ НА ОСНОВЕ ЭФФЕКТА ЗАБЫВАНИЯ, ВЫЗВАННОГО ИЗВЛЕЧЕНИЕМ ИНФОРМАЦИИ

(Самарский национальный исследовательский университет  
им. академика С.П. Королева, Самара, Россия)

В настоящий момент помимо непосредственно алгоритмов машинного обучения интерес представляют также любые их модификации, позволяющие улучшить их качество работы, добиться более высокой точности вычислений.

Во время обучения модели многослойного персептрона нередко возникает проблема переобучения, когда модель предсказывает выходные значения для известных ей примеров из тренировочной выборки с высокой точностью, но на новых примерах (из тестовой выборки) точность сильно снижается. То есть модель подстраивается под каждый конкретный пример из тренировочной выборки.



ки и ее обобщающая способность теряется. Чтобы преодолеть проблему переобучения, необходимо препятствовать модели подстраиваться под каждый конкретный пример. Для этого можно увеличить размер тренировочной выборки, уменьшить количество признаков, воспользоваться техниками регуляризации (L1, L2, dropout) [1] или приостанавливать обучение модели на этапе, когда средняя ошибка на валидационной выборке начинает увеличиваться (early stopping) [2].

В данной работе будет предложен алгоритм обучения с использованием модификаций сигмоидальной функции с введенным нечетким параметром  $c$  и группой нечетких параметров  $c_g$  и  $c_f$  в качестве функции активации [3,4].

Идея добавления данных нечетких параметров возникла на основе RIF-эффекта (retrieval-induced forgetting) [5], который также называют эффектом забывания, вызванного извлечением информации. Суть его в следующем: были описаны эксперименты, во время которых испытуемым были представлены категории предметов и несколько предметов из каждой категории. Затем часть представленных предметов одной из категорий испытуемые пытались запомнить, то есть набор данных был разбит таким образом:

- тренируемые предметы из тренируемой категории,
- не тренируемые предметы из тренируемой категории,
- не тренируемые предметы из не тренируемой категории.

В итоге хуже всего испытуемым удавалось вспомнить именно не тренируемые предметы из тренируемой категории (55%), результаты для остальных групп были следующими: тренируемые предметы из тренируемой категории – 81%, не тренируемые предметы из не тренируемой категории – 68%.

В классической модели многослойного персептрона в качестве функции активации используется сигмоидальная функция:

$$F(x) = \frac{1}{(1+e^{-x})} \quad (1)$$

С ее помощью удастся получить степень принадлежности конкретного примера из выборки к каждому из классов. Но вместо одного конкретного значения степени принадлежности к классу хотелось бы иметь возможность работать с диапазоном таких значений и выбирать в итоге наиболее оптимальное в данной ситуации. Введение нечетких параметров позволяет этого достичь. При этом введенный параметр  $c$  (или пара параметров:  $c_g$  и  $c_f$ ) добавляет в исходную функцию некоторое отклонение, что и моделирует «забывание», убирая излишнее внимание модели к каждому конкретному примеру из выборки. Множество всех измененных функций, параметр  $c$  в которых принимает значение из диапазона  $[0;1]$ , образует область определения нечеткого множества [4].

Нечеткие параметры были добавлены разными способами:

$$F_1(x) = c + \frac{1-c}{(1+e^{-x})} \quad (2)$$

$$F_2(x) = \frac{1-c}{(1+e^{-x})} \quad (3)$$

$$F_3(x) = c + \frac{1-2c}{(1+e^{-x})}, \quad (4)$$



включая добавление пары нечетких параметров  $c_g$  и  $c_f$ :

$$F_4(x) = c_g + \frac{1-(c_g+c_f)}{1+e^{-x}} \quad (5)$$

Среднее значение ошибки [6] для одного примера по выборке оценивалось в виде [6],

$$Error = \frac{1}{2}MSE = \frac{1}{n} \sum_{i=1}^m \frac{1}{2} (y_i - \hat{y}_i)^2, \quad (6)$$

где  $n$  – количество всех примеров в выборке (тренировочной или тестовой),  $m$  – количество классов, принадлежность к которым определяется для каждого примера,  $y_i$  – идеальное значение выходного параметра (принадлежность классу  $i$ ),  $\hat{y}_i$  – оцениваемое значение  $y_i$ , предсказанное сетью.

Параметры обучения модели были заданы следующим образом: количество итераций: 500, коэффициент обучения: 0.001, размер тренировочного набора данных по отношению ко всему набору данных: 80%. Обучение проводилось на наборе данных «Statlog (Heart) Data Set» [7].

Кривая обучения (график значений средней ошибки в зависимости от числа итераций) на тренировочной и тестовой выборках для сигмоидальной функции активации и ее модификации (5) показана на рис. 1. Здесь можно заметить, что хоть ошибка вычисления значений для примеров из тестовой выборки снижается незначительно, добавление нечетких параметров  $c_g$  и  $c_f$  позволяет повысить обобщающую способность.

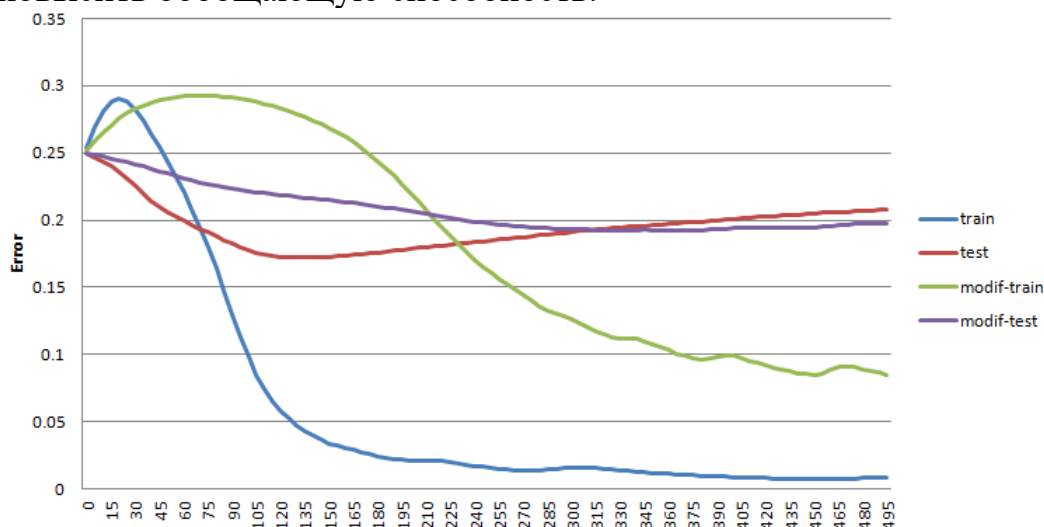


Рисунок 1 – График ошибки обучения от числа итераций

### Благодарности

Работа выполнена при поддержке гранта Президента РФ № МК-6218.2018.9.

### Литература

1. Deep Learning, Regularization [Электронный ресурс]. – URL: <http://www.deeplearningbook.org/contents/regularization.html> (дата обращения: 20.03.2017)



2. Wikipedia, Early stopping [Электронный ресурс]. – URL: [https://en.wikipedia.org/wiki/Early\\_stopping](https://en.wikipedia.org/wiki/Early_stopping) (дата обращения: 20.03.2017)
3. Kulikovskikh I.M. Cognitive validation map for early occupancy detection in environmental sensing // Engineering Applications in Artificial Intelligence – 2017. – 65. – pp. 330-335.
4. Prokhorov S.A., Kulikovskikh I.M. Fuzzy learning performance assessment based on decision making under internal uncertainty/7th Computer Science and Electronic Engineering Conference (CEEC 2015). – Colchester, UK, 2015. – pp. 65-70.
5. Wikipedia, RIF-effect [Электронный ресурс]. – URL: [https://en.wikipedia.org/wiki/Retrieval-induced\\_forgetting](https://en.wikipedia.org/wiki/Retrieval-induced_forgetting) (дата обращения: 20.03.2017)
6. Jerry M. Mendel, Robert I. Bob John. Type-2 Fuzzy Sets Made Simple [Текст] // IEEE Transactions on Fuzzy Systems. – 2002. – vol. 10, no. 2. – С. 117-127.
7. Wikipedia, MSE [Электронный ресурс]. – URL: [https://en.wikipedia.org/wiki/Mean\\_squared\\_error](https://en.wikipedia.org/wiki/Mean_squared_error) (дата обращения: 20.03.2017)
8. UCI Machine learning repository, Statlog (Heart) Data Set [Электронный ресурс]. – URL: [http://archive.ics.uci.edu/ml/datasets/statlog+\(heart\)](http://archive.ics.uci.edu/ml/datasets/statlog+(heart)) (дата обращения: 20.03.2017)

А.О. Новиков, Н.Г. Чернобровин

## АВТОМАТИЗИРОВАННАЯ УСТАНОВКА ДИАГНОСТИКИ БИПОЛЯРНЫХ ТРАНЗИСТОРОВ

(Самарский университет)

Известен способ отбраковки дефектных биполярных структур, основанная на эффекте отрицательного дифференциального сопротивления в динамическом режиме питания высокочастотным напряжением [1]. Типовая S-образная вольт-амперная характеристика в координатах  $U_b$  ( $U_k$ ), где  $U_b = I_b R_b$  амплитуда напряжения на токоизмерительном резисторе в цепи базы,  $U_k$  – амплитуда напряжения питания на частоте 1 МГц, представлена на Рис.1.

Разработана автоматизированная установка диагностики биполярных транзисторов, в основу которой положено устройство для измерения параметров полупроводниковых приборов с S-образной вольт-амперной характеристикой [2].

Установка является периферийным устройством ПЭВМ типа IBM PC и реализует автоматизированное измерение напряжений и токов исследуемого транзистора, соответствующих характерным точкам динамической S-образной вольт-амперной характеристики. Функциональная схема установки представлена на рис. 2.