



файла преобразуется в бинарный и отправляется запросом к сервису. В ответе будет получен распознанный текст.

Затем возможно несколько способов классификации по жанрам. Одним из которых является — приведение соответствия отдельных слов из текста выборкам, которые содержат характерный набор слов для каждого музыкального жанра.

Недостатком данного способа является то, что он не позволяет корректно обработать данные, в случае нечеткого голоса исполнителя либо ярко выраженного музыкального сопровождения, на фоне которого невозможно распознать голос.

Разрабатываемый продукт сможет помочь людям с достаточной точностью определить жанр музыкальной композиции либо подобрать для себя песню, опираясь на набор слов, который соответствует их любимому жанру.

### Литература

- 1 Теория статистики. – Учебник/ под редакцией проф. Г.Л. Громько.- 2-е изд., перераб. и доп. - М.: ИНФРА–М, 2016 г.- 476 с., гл. 1, стр. 28-37; стр. 149-193.
- 2 Соколович Ю.А., Богданова Г.С. Физика: справочник. – 2-е издание перераб. – Х.: Веста: издательство «Ранок», 2015. – 464 с.
- 3 Распознавание речи. URL: <http://cloud.yandex.ru/docs/speechkit/stt/> (дата обращения: 14.12.2017).

Д.А. Лебедев, В.Г. Литвинов

## ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОГО РАЗМЕРА НЕЙРОННОЙ СЕТИ

(Самарский университет)

Аннотация: в данной статье рассматривается способ определения оптимального размера нейронной сети через сопоставление весов синапса от размера нейросети. Эксперимент показывает, что локальные минимумы хорошо соответствуют свойствам традиционных кривых обучения — зависимостей ошибок обучения от размера нейросети. Такой подход может быть использован для определения оптимального размера сети при отсутствии тестовой выборки.

Ключевые слова: размер сети, кривая обучения.

При использовании искусственных нейронных сетей важной задачей является нахождение оптимального размера (структуры) сети — такого числа скрытых слоев нейронов и нейронов в слоях, которые дадут максимум обобщающих способностей, т.е. минимум ошибки обобщения, особенно в случае отсутствия тестовой выборки или невозможностью разделить выборку данных на обучающую и тестовую части из-за недостаточности общего объема данных.

Поэтому широко используется понятие "кривых обучения — зависимостей ошибок обучения и обобщения от размеров нейросети и обучающей вы-



борки. Оптимуму соответствуют локальные минимумы. Экстраполяции таких графиков [2] позволяют также оценивать необходимые и достаточные для достижения максимума обобщающих способностей объемы обучающих выборок в случае первоначальной недостаточности объемов выборочных данных.

Иным классом кривых обучения являются зависимости внутренних свойств нейросети от её размера, сопоставляемые затем с ошибки обобщения. Варианты — анализ задачи, теоретическая связь [5] ошибки обучения и максимума суммы модулей весов синапсов, приходящих на нейрон сети, градиенты целевой функции обученной сети и позволяющий оценить разницу между ошибками обучения и обобщения. Такие критерии позволяют обходиться без независимой тестовой выборки.

В работе рассматривается вариант кривой обучения — зависимость среднего модуля веса синапса от размера нейросети. Точнее значение длины вектора весов синапсов сети, деленное на общее число синапсов, с целью увеличения влияния наибольших по модулю весов. Этот критерий не является всеобъемлющим, т.к. имеется неоднородность наборов синапсов сети от слоя к слою (для сетей малого размера в часто наблюдалось различие средних модулей и дисперсий весов синапсов выходного и скрытого слоя сети).

Было взято 6 независимых тестовых выборки. Все 6 задач представляют собой задачи классификации с учителем на то или иное число классов. Использовались сети с одним скрытым слоем, число нейронов в котором менялось от 1 до 25. В каждой задаче для каждого размера нейросети обучалось по 25 сетей (с разными начальными случайными значениями синапсов), свойства которых затем усреднялись при построении кривых обучения.

На графике представлена следующая кривая обучения: средние значения ошибок обучения и обобщения. Вдоль осей ординат отложено число нейронов в скрытых слоях сетей. Вокруг каждой точки отложена дисперсия соответствующей выборки из 25 экспериментальных значений.

Когда с ростом размера сети с некоторого момента ошибка обобщения начинает снова возрастать — момент стабилизации и выход индикатора на асимптоту чуть запаздывает по сравнению с моментом достижения минимума ошибки в первой задаче, соответствующей рисунку 1, а в второй задаче на рисунке 2 локальный минимум при размере сети в 6 нейронов точно соответствует минимуму ошибки обобщения. В третьей задаче, график которой представлен на 3 рисунке, не имеется четко выраженного экстремального поведения и существенно нестабилен на всем диапазоне исследованных размеров нейросети  $n$  от 1 до 25 нейронов.

Как видно из графика, при выборе оптимального размера сети опираться только на значение ошибки обучения недостаточно нельзя выявить возникновение переобучения нейросети, поэтому сопоставление поведения нескольких индикаторов (как было сделано на приведенных графиках) позволяет либо более обоснованно подтвердить выбор размера нейросети, либо увидеть возможное существование проблем (например, неадекватности модели из-за возникновения переобучения). Возможность же обойтись без проверки на тестовой вы-



борке позволяет обучать нейросеть на всем доступном наборе примеров, без разделения его на обучающий и тестовый фрагменты, и ожидать, что с ростом числа обучающих примеров снизится и риск переобучения нейросети [2].

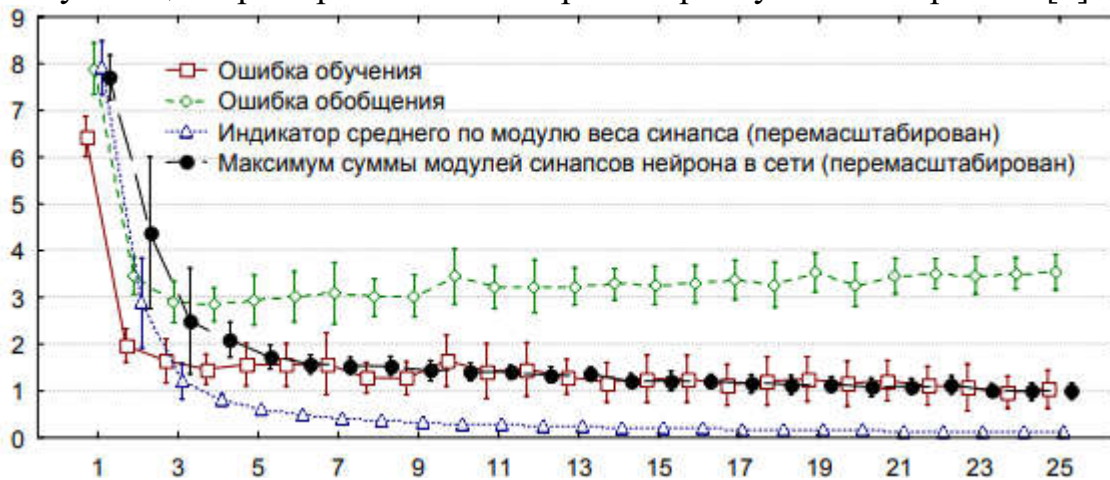


Рис. 1. График для первой тестовой выборки

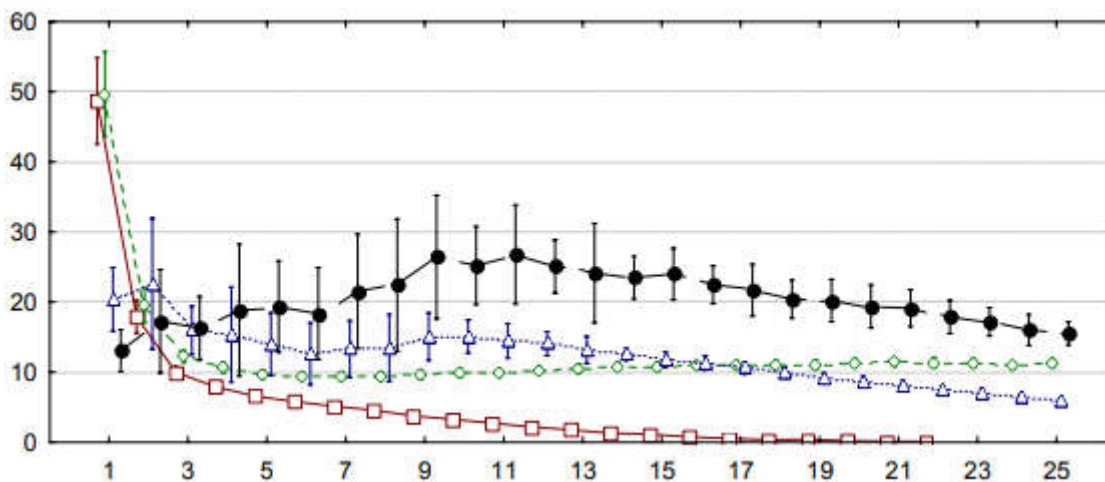


Рис. 2. График для второй тестовой выборки

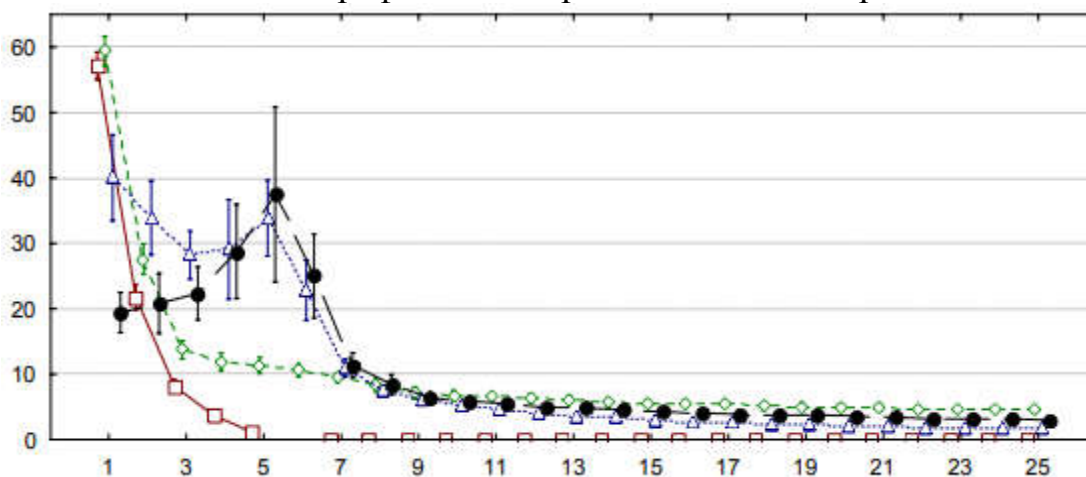


Рис. 3. График для третьей тестовой выборки

Данный вариант кривой обучения — зависимость среднего значения модуля веса синапса в сети от размера нейросети, может помочь достаточно точно надежно определить оптимальный размер сети, обеспечивающий минимум ошибки. Это также позволяет обходиться без расчета ошибки на независимой



тестовой выборке, допускает разные варианты параметров системы (модуль веса, среднее квадратичное значение).

### Литература

1. Cortes C., Jackel L.D., Solla S.A., Vapnik V., Denker J.S. Learning curves: asymptotic values and rate of convergence / Advances in Neural Information Processing Systems 6 (1993). Morgan Kaufmann, 1994. ñ pp.327-334.
2. Bartlett P.L. For valid generalization, the size of the weights is more important than the size of the network / Advances in Neural Information Processing Systems 9 (1996). MIT Press, 1997. ñ pp.134-140.

А.А. Лякишев, И.В. Лёзина

## СРАВНЕНИЕ СЛУЧАЙНОГО АЛГОРИТМА И АЛГОРИТМА РОЯ ЧАСТИЦ ПРИ РЕШЕНИИ ЗАДАЧИ АППРОКСИМАЦИИ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТИ

(Самарский национальный исследовательский университет  
имени академика С.П. Королева)

Основная задача аппроксимации — построение аппроксимирующей (приближенной) функции, которая наиболее близко проходит около заданных точек или около некоторой заданной непрерывной функции. В основном, такая задача возникает, если в исходных данных задачи присутствуют некоторые погрешности, когда алгоритмы интерполяции использовать нецелесообразно.

Нейронная сеть представляет собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов), обрабатывающих входные данные и передающих результат последующим нейронам. На практике наиболее распространёнными функциями активации являются: сигмоидальная, пороговая, гиперболический тангенс, радиально-базисная, логистическая и т. д.

Одной из таких сетей является радиально-базисная сеть (RBF) [1]. Сеть состоит из одного скрытого слоя, выполняющего нелинейное преобразование входной последовательности, и выходного слоя, состоящего из единственного нейрона, суммирующего выходные значения сигналов [2]. Особенностью этой сети является радиальная функция активации скрытых нейронов:

$$\phi(x) = \phi(\|x - c\|),$$

где  $x$  — входной вектор,  $c$  — центр,

$\phi(x)$  — убывающая функция, чаще всего равная 0 вне некоторого отрезка.

Задача аппроксимации состоит в подборе соответствующего количества радиальных функций и их параметров, а также подборе весов. Эту проблему можно свести к минимизации целевой функции:

$$E = \sum_{i=1}^p \left[ \sum_{j=1}^k w_j \phi(\|x_i - c_j\|) - d_i \right]^2,$$