

# ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ОБРАБОТКИ ТЕКСТОВОЙ ИНФОРМАЦИИ

Потараев В. В.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: vic229@rambler.ru

На протяжении последних десятилетий объём текстовой информации, хранимой в информационных системах, резко возрос. Это обуславливает актуальность автоматизированной обработки информации, в том числе текстовой. Нейронные и семантические сети являются эффективными инструментами обработки информации. Рассмотрим применение нейронной сети для решения задачи ответа на вопрос.

## ВВЕДЕНИЕ

В современных информационных системах хранятся и обрабатываются довольно большие объёмы данных. Это обуславливает актуальность автоматизированной обработки информации, в том числе текстовой. Для повышения эффективности обработки информации зачастую применяется семантический анализ, то есть учёт смысла данных. Целью данной работы является разработка методов автоматизированной обработки данных, основанных на семантическом анализе.

### I. ЗАДАЧА ПОИСКА ОТВЕТА НА ВОПРОС

Информационные системы зачастую используются в качестве инструмента для нахождения ответа на некоторый запрос [1]. Рассмотрим задачу ответа на вопрос, сформулированный на естественном языке.

Для решения данной задачи может быть использован алгоритм, основанный на представлении текста в виде семантической сети. В данной сети узлами являются слова текста, а связями – семантические отношения между словами [2].

### II. НЕЙРОННАЯ СЕТЬ ХОПФИЛДА

Одной из моделей представления информации, которые могут быть использованы для классификации данных [3], а также для решения других задач, является модель, основанная на искусственной нейронной сети.

Искусственная нейронная сеть (ИНС) – это математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма [4].

Одним из видов ИНС является нейронная сеть Хопфилда. Основной её особенностью является наличие обратных связей.

Сеть Хопфилда можно определить как динамическую систему с обратной связью, у которой выход одной операции служит входом сле-

дующей операции сети. Каждая операция сети называется итерацией [3]. Схема нейронной сети Хопфилда представлена на рисунке 1.

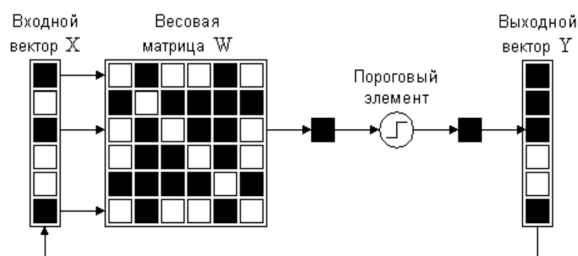


Рис. 1 – Нейронная сеть Хопфилда

К недостаткам сети Хопфилда можно отнести ее сравнительно небольшой объем памяти, то есть в случае записи слишком большого числа образов нейронная сеть перестает их распознавать. Кроме того, ИНС может сойтись к так называемым ложным аттракторам и дать неверный результат [3].

Порядок использования нейронной сети Хопфилда может быть представлен следующим образом:

1. Выбрать параметры сети (размерность входного вектора, множество сохранённых в сети векторов).
2. Рассчитать весовую матрицу на основании параметров.
3. Подать на вход сети распознаваемый входной вектор.
4. Производить перерасчёт выходного вектора сети, пока он не станет постоянным.
5. Полученный вектор является сохранённым сетью вектором, который наиболее «похож» на входной вектор.

### III. ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ АНАЛИЗА ИНФОРМАЦИИ

Модель семантической сети может быть использована в информационной системе для ответа на вопрос [2].

Модели, основанные на нейронных и семантических сетях, весьма сильно отличаются меж-

ду собой. Так, например, по модели нейронной сети далеко не всегда понятно, о каких объектах идёт речь – даже при небольшом количестве нейронов. Семантическая сеть, напротив, даёт прекрасное смысловое представление о предмете исследования, ее непосредственно можно использовать в качестве результата [5].

Нейронная сеть – это не более, чем сложная алгебраическая функция; для любой нейронной сети можно найти эквивалентную матричную систему уравнений [5].

Тем не менее, модели представления информации, основанные на нейронных и семантических сетях, могут быть использованы для решения одних и тех же задач, например, для решения задачи классификации текстовой информации [3].

Покажем, что модель нейронной сети Хопфилда также может быть использована в информационной системе для ответа на вопрос. Пусть каждому входу сети соответствует некоторое слово (или понятие). Наличие слова в предложении может быть обозначено как наличие сигнала на соответствующем входе. Каждому предложению текста можно поставить в соответствие некоторый вектор.

Порядок использования нейронной сети Хопфилда для ответа на вопрос может быть представлен следующим образом:

1. Выбрать параметры сети (размерность входного вектора – количество уникальных слов текста, множество сохранённых в сети векторов – предложения некоторого текста).
2. Рассчитать весовую матрицу на основании параметров.
3. Подать на вход сети распознаваемое предложение (вопрос).
4. Производить перерасчёт выходного вектора сети, пока он не станет постоянным.
5. Ответ будет содержаться в получившемся выходном векторе.

#### IV. ПРИМЕР

Предположим, в тексте есть два предложения: «На улице идёт дождь» и «Из-за дождя появились лужи на асфальте». Пример векторов для данных предложений представлен на рисунке 2. Весовая матрица для нейронной сети Хопфилда должна быть составлена с учётом запоминания этих двух предложений.

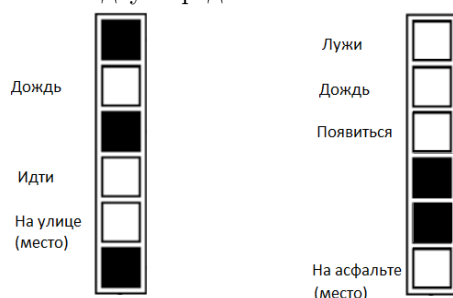


Рис. 2 – Пример векторов для двух предложений

Предположим, нужно ответить на вопрос «Где идёт дождь?». На вход сети подаются 2 сигнала, соответствующие словам «идёт» и «дождь». Очевидно, в результате работы сети на выходе будет вектор, соответствующий словам «дождь», «идти», «на улице» (место). Поскольку «где» – это вопрос определения места, то ответом будет «на улице».

Таким образом, модель нейронной сети может быть использована практически без дополнительных приспособлений для решения задачи ответа на вопрос. Однако, если в вопросе вместо слов текста будут их синонимы, то для получения корректного результата нужно все слова рассматривать в контексте их связей со словами-синонимами. То есть нужен отдельный этап алгоритма, который будет учитывать синонимичность. В случае же семантической сети синонимичность является лишь одним из типов связей, что говорит о большей гибкости семантической сети как модели для обработки текстовых данных.

#### V. ВЫВОД

Нейронная сеть Хопфилда способна хранить информацию из текста. Её построение может быть автоматизировано.

Нейронная сеть является мощным инструментом распознавания образов. Распознавание предложений текста, в свою очередь, может быть использовано для поиска ответа на вопрос. Семантические и нейронные сети могут использоваться при решении одних и тех же задач. Тем не менее, семантическая сеть является более универсальной моделью для решения различных задач обработки текстовой информации.

1. Информационные системы. [Электронный ресурс]. – Электронные данные. – Режим доступа: [http://www.umpro.ru/index.php?page\\_id=17&art\\_id\\_1=261&group\\_id\\_4=72](http://www.umpro.ru/index.php?page_id=17&art_id_1=261&group_id_4=72), Дата доступа : 23.09.2018.
2. Потараев, В. В. Алгоритм применения семантической сети для поиска ответа на вопрос / В. В. Потараев // Компьютерные системы и сети: Материалы 54-ой научной конференции аспирантов, магистрантов и студентов. – Минск : БГУИР, 2018. – С. 103–105.
3. Серебряная, Л. В. Методы классификации текстовой информации на основе искусственной нейронной и семантической сетей / Л. В. Серебряная, В. В. Потараев // Информатика. – 2016. – №4. – С. 95–103.
4. Рот, В. И. Система распознавания символов с помощью искусственной нейронной сети / В. И. Рот // XIV Международная научно-практическая конференция студентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии». – Томск : Изд-во ТПУ, 2016. – С. 54–55.
5. Искусственное сознание (Сети: нейронные или семантические?). [Электронный ресурс]. – Электронные данные. – Режим доступа: <http://alephegg.narod.ru/Survey/NeuralVersusSemantic.htm>, Дата доступа : 23.09.2018.