

УДК 004.8:004.912

СЕМАНТИЧЕСКОЕ АННОТИРОВАНИЕ ТЕКСТОВЫХ ДОКУМЕНТОВ НА ОСНОВЕ ИЕРАРХИЧЕСКОЙ РАДИАЛЬНО- БАЗИСНОЙ НЕЙРОННОЙ СЕТИ

Е. В. Бодянский

Доктор технических наук, профессор*
Контактный тел.: (057) 702-18-90
E-mail: bodya@kture.kharkov.ua

О. В. Шубкина

Аспирант*
Контактный тел.: (057) 702-13-37
E-mail: olga.shubkina@gmail.com

*Кафедра искусственного интеллекта
Харьковский национальный университет
радиоэлектроники
пр. Ленина, 14, г. Харьков, Украина, 61166

В роботі запропоновано ієрархічну радіально-базисну нейронну мережу з багатшаровою архітектурою, яка використовується для видобування знань із текстових джерел із урахуванням максимальної кількості релевантних ознак кожного об'єкта, віднесення його до обраного класу онтології

Ключові слова: семантична анотація, радіально-базисна нейронна мережа, багатшарова архітектура

В работе предложена иерархическая радиально-базисная нейронная сеть с многослойной архитектурой, которая используется для извлечения знаний из текстовых источников с учетом максимального количества релевантных признаков каждого объекта и отнесения его к выбранному классу онтологии

Ключевые слова: семантическая аннотация, радиально-базисная нейронная сеть, многослойная архитектура

The hierarchical radial basis function neural network with a multi-layered architecture is proposed. This neural network is used for extracting knowledge from textual sources with the maximum number of relevant attributes for each object and assigns it to the selected class of ontology

Key words: semantic annotation, radial basis function neural network, multi-layered architecture

1. Введение

Ресурсы знаний различаются в зависимости от отраслей индустрии и приложений, но, как правило, представлены в текстовом виде. Документы представляют собой ресурс знаний любой организации, в котором содержится около 80-85% всей информации [1]. Кроме того, с развитием Web наблюдается огромный рост цифрового контента во внутрикорпоративных сетях, базах данных, архивах, в связи с чем возрастает роль извлечения знаний из текстовых источников, накопленных на разных этапах развития организации. Этот процесс является основной задачей Text Mining (TM) – интеллектуального анализа текстовой информации. В управлении знаниями TM играет важную роль потому, что является механизмом выявления закономерностей, характерных элементов или свойств, которые могут использоваться в качестве метаданных документа, ключевых слов, аннотаций [2].

В настоящее время в системах TM широкое развитие приобретают технологии, основанные на пара-

дигме Semantic Web [3]. Одним из примеров является разработка систем и методов семантического аннотирования текстовых документов. Основная идея заключается в создании описания текстового документа (на основе классов и свойств онтологии предметной области) в машинно-понятной форме для его последующего использования интеллектуальными агентами. Семантическое аннотирование и структурирование знаний – это подход к решению проблемы информационного роста, а также один из многообещающих способов извлечения знаний из хранилищ текстовых документов, позволяющий избежать неоднозначности (неопределенности) информационного поиска, а также повысить возможность взаимодействия и интеграции информации из гетерогенных источников, накопленных организациями в процессе становления и развития.

Разработанные методы семантического аннотирования можно разделить на две группы [4]: внутреннее аннотирование, заключающееся во встраивании семантической разметки внутрь документа, и внешнее

аннотирование, включающее сохранение набора метаданных в RDF-формате [5] в отдельном файле. В области информационного поиска метаданные чаще всего рассматривают как набор элементов, описывающих основные свойства текстового документа (включая тематику), которые могут также содержать элементы из уже существующих схем (например, Dublin Core [6]). Важно заметить, что существуют различные уровни семантического аннотирования (уровни слов, предложений, параграфов, разделов), основное отличие которых состоит в способах предобработки текстовой информации с выделением необходимых информационно-значимых объектов и соответственно разном построении пространства признаков и обработке полученных данных.

Существует набор стандартных решений, которые разработаны для описания метаданных и формирования семантической аннотации, как, например, стандарт Dublin Core, проекты FOAF [7], SKOS [8]. Другие известные методы реализуют создание семантических аннотаций на основе лингвистических шаблонов в качестве описания возможных конструкций предложений (Hearst's patterns) [9], а также путем формирования шаблонов на основе регулярных выражений или разбора предложений по частям речи для последующего сопоставления с полученным образцом.

Кроме того, широкое распространение получили методы семантического аннотирования, в основу которых положены различные алгоритмы машинного обучения, а именно, классификация с помощью машин опорных векторов (SVM) [10 – 12], скрытых марковских моделей (НММ) [13], марковских моделей максимальной энтропии (MEMM) [14], условных случайных полей (CRF) [15, 16]. Некоторые подходы описывают только контексты распределения целевых экземпляров и не выявляют зависимости между ними (SVM и Voted Perceptron), другие же могут моделировать линейную цепочку зависимостей (MEMM и CRF).

Вектор признаков для текстовых объектов зачастую формируется на основе статистической меры для оценки важности слова в контексте документа, являющегося частью коллекции документов или корпуса (TFIDF) [17, 18]. Например, в [19] экземпляром концепта онтологии выступает непосредственно текстовый документ. Кроме статистического распределения и применения метода гистограмм, в системах ТМ может учитываться также позиция и графическая форма (графемная составляющая) слов [20]. При этом возникает проблема того, что необходимо учитывать максимальное число характеристик текстовых объектов, что в данном случае может влиять на качество обработки текстовых данных и соответственно на создание семантических аннотаций.

В настоящей работе в качестве альтернативы существующим методам обработки информации в рамках ТМ предлагается иерархическая радиально-базисная нейронная сеть с многослойной архитектурой, в каждом узле использующая обычную радиально-базисную нейронную сеть, на вход которой подается только некоторая часть признаков, что решает проблему «проклятия размерности». Данный метод позволяет решить проблему извлечения знаний из текстовых источников с учетом максимального числа релевантных признаков для построения семантических аннотаций

текстовых документов выбранной предметной области, предусматривает возможность обработки информации по мере ее поступления в последовательном режиме, характеризуется простотой реализации.

2. Семантическое аннотирование текстовых документов

Одним из способов формирования семантических аннотаций для текстовых документов является идентификация именованных сущностей документа и отнесение их к заданному классу онтологии предметной области. Примером именованных сущностей являются персоны, организации, географические объекты и прочие объекты, обозначаемые в тексте с использованием имен собственных, различных специальных символов (кавычек и прочее), а также графемно-выделенные слова [20]. В [17, 18] описан принцип статистического формирования вектора признаков для текстовых данных.

С другой стороны, если рассматривать семантическое аннотирование на уровне документов, то в таком случае проекция текстового документа на онтологию предметной области будет заключаться в отнесении его к определенному классу онтологии как экземпляра концепта данного класса. При этом следует обратить внимание, что для уменьшения неоднозначности и повышения эффективности дальнейшей интеграции различных информационных источников необходимо найти класс, максимально соответствующий данному текстовому документу.

Таким образом, далее рассмотрим обобщенную модель семантического аннотирования. При этом необходимо учесть, что исходные данные могут быть представлены как в виде именованных сущностей, так и в виде непосредственно текстовых документов, которые необходимо «спроецировать» на онтологию предметной области. Для обобщения будем называть исходные данные текстовыми объектами, а под проекцией будем понимать отнесение полученного на этапе предобработки текстового объекта к какому-либо онтологическому концепту (классу).

Формальное представление семантического аннотирования текстового документа можно получить следующим образом. Для данной онтологии предметной области Ont набор концептов (классов) определяется как $\text{ConceptSet} = (c(1), c(2), \dots, c(i), \dots, c(N_1))$, где $c(i)$ – i -й концепт из Ont. Для рассматриваемого текстового документа TD набор текстовых объектов, полученных на этапе предобработки, можно представить как $\text{ObjectSet} = (x(1), x(2), \dots, x(j), \dots, x(N_2))$, где $x(j)$ – j -й текстовый объект, представленный в виде некоторого набора релевантных признаков большой размерности, N_1 и N_2 – количество концептов (классов) онтологии и мощность исходной выборки текстовых объектов соответственно.

Тогда семантическая аннотация – разметка или набор метаданных рассматриваемого текстового документа – на основе данной онтологии будет определена как $\text{LabelSet} = \{l_i | \exists c_j \in \text{ConceptSet} \wedge l_i = c_j\}$, в котором LabelSet – уникальное множество для каждого текстового документа, состоящее из концептов (классов) онтологии предметной области, полученных путем

проекция множества текстовых объектов, принадлежащих этому документу, на заданную онтологию с использованием методов извлечения знаний из текстовых источников. Полученная разметка может быть впоследствии представлена в виде RDF-описания для дальнейшего использования интеллектуальными агентами в системах управления знаниями.

Далее введем иерархическую радиально-базисную нейронную сеть с многослойной архитектурой для извлечения знаний из текстовых источников с учетом максимального числа релевантных признаков и, таким образом, получения семантических аннотаций текстовых документов.

При этом априорно заданная обучающая выборка $x(1), x(2), \dots, x(N_2)$ состоит из текстовых объектов ObjectSet, полученных после преобработки данных, а значения выхода нейронной сети у формируют множество LabelSet.

3. Многослойная архитектура иерархической радиально-базисной нейронной сети

Радиально-базисная нейронная сеть (Radial Basis Function Network – RBFN) является универсальной аппроксимирующей конструкцией [21, 22], предназначенной для решения широкого класса задач интеллектуального анализа данных и нейроуправления. Входной слой этой сети – это сенсоры, которые связывают искусственную нейронную сеть (ИНС) с окружающей средой. Единственный скрытый слой, образованный, так называемыми, R-нейронами, осуществляет нелинейное преобразование входного пространства R^n в скрытое пространство R^h , как правило, более высокой размерности ($h > n$). Выходной слой, образованный адаптивными линейными ассоциаторами формирует отклик сети $y = (y_1, y_2, \dots, y_m)^T$ на входной сигнал $x = (x_1, x_2, \dots, x_n)^T$.

ИНС такого типа реализуют идею, сформированную Т. Кавером и состоящую в том, что линейно неразделимая задача распознавания образов в пространстве R^n может стать линейно разделимой в пространстве более высокой размерности R^h .

Как уже отмечалось, радиально-базисные сети являются универсальными аппроксиматорами, а в силу того, то в них присутствует только один нелинейный слой, а настраиваются только параметры линейного выходного слоя, то для их обучения могут быть использованы быстродействующие процедуры, обладающие как следящими, так и фильтрующими свойствами. Кроме того, одно из преимуществ использования RBFN в рамках рассматриваемой проблемы заключается в том, что существует возможность добавления нового класса – нового выхода сети без переобучения синаптических весов остальных выходов сети.

Однако основным недостатком радиально-базисных нейронных сетей является, так называемое, «проклятие размерности», ведущее к тому, что при решении сложных задач (распознавание образов, классификация и т.п.) число нейронов скрытого слоя экспоненциально растет с увеличением размерности входного пространства. Это свойство ограничивает использова-

ние таких ИНС в задачах обработки данных с большим количеством признаков.

В связи с этим нами для решения задач ТМ вводится в рассмотрение иерархическая радиально-базисная нейронная сеть (HRBFN) с многослойной архитектурой, которая приведена на рис. 1.

На вход HRBFN подается вектор признаков $x = (x_1, x_2, \dots, x_n)^T$, размерности $(n \times 1)$, при этом значение n таково, что количество R-нейронов обычной RBFN недопустимо велико. Разобьем вектор x достаточно произвольным образом на набор подвекторов $\tilde{x}_l^T, l = 1, 2, \dots, p_1$ так, что

$$x = (x_1, x_2, \dots, x_{n_1}, x_{n_1+1}, x_{n_1+2}, \dots, x_{n_1+n_2}, \dots, x_{n_1+n_2+\dots+n_{p_1-1}+1}, \dots, x_{n_1+n_2+\dots+n_{p_1}})^T = (\tilde{x}_1^T, \tilde{x}_2^T, \dots, \tilde{x}_1^T, \dots, \tilde{x}_{p_1}^T)^T, \quad (1)$$

при этом размерность каждого $\tilde{x}_l^T - (n_l \times 1), \sum_{l=1}^{p_1} n_l = n$ такова, что допускает обработку обычной радиально-базисной нейронной сетью.

Первый скрытый слой HRBFN образован p_1 обычными RBFN, каждая из которых имеет n_l входов и одинаковое число m выходов, которое определяется количеством классов, содержащихся в обрабатываемом массиве данных. Таким образом, число выходов первого скрытого слоя определяется значением $p_1 m < n$. Далее на основании априорно заданной обучающей выборки $x(1), x(2), \dots, x(N)$ с известной классификацией сети первого скрытого слоя обучаются, при этом обучающий сигнал $d(k) = (d_1(k), d_2(k), \dots, d_m(k))^T$ является общим для всех p_1 сетей первого слоя.

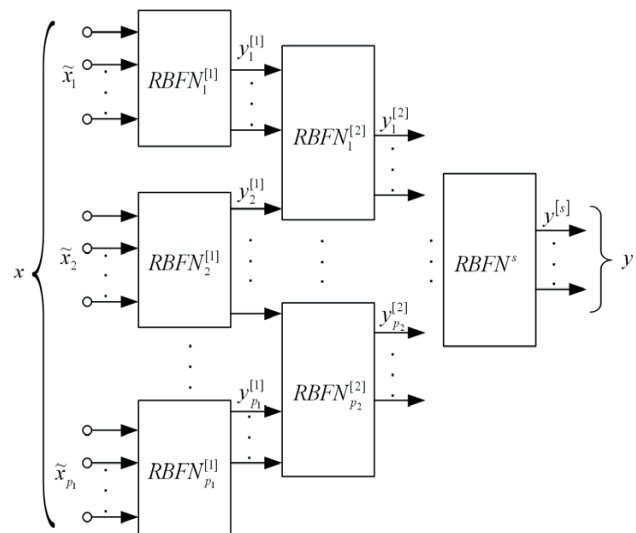


Рис. 1. Архитектура иерархической радиально-базисной нейронной сети

После обучения синаптические веса «замораживаются» и обучающая выборка вновь подается на вход всех сетей, которые формируют выходной сигнал $y^{[i]} = (y_1^{[i]}, y_2^{[i]}, \dots, y_{p_1}^{[i]})^T$, где $y_l^{[i]} = (y_{l_1}^{[i]}, y_{l_2}^{[i]}, \dots, y_{m}^{[i]})^T$. Этот векторный сигнал размерности $(p_1 m \times 1)$ является входным для второго скрытого слоя, содержащего $p_2 < p_1$ радиально-базисных сетей.

Сети второго скрытого слоя обучаются аналогично сетям первого слоя, при этом в качестве обучающего

сигнала используется все та же векторная последовательность $d(k)$, $k=1,2,\dots,N$. Далее «замораживаются» веса второго слоя и формируется его векторный выход $y^{[2]} = (y_1^{[2]}, y_2^{[2]}, \dots, y_{p_2}^{[2]})^T$ размерности $(p_2 \times 1)$.

Процесс наращивания архитектуры HRBFN продолжается до тех пор, пока в выходном s -м слое не останется одна RBFN, чей выход $y^{[s]}$ размерности $(m \times 1)$ является и выходом у иерархической сети в целом.

4. Алгоритм обучения радиально-базисной нейронной сети

Для обучения отдельных RBFN, формирующих слои HRBFN, введем алгоритм обучения, основанный на критерии [23], который предназначен для решения задач распознавания образов (классификации).

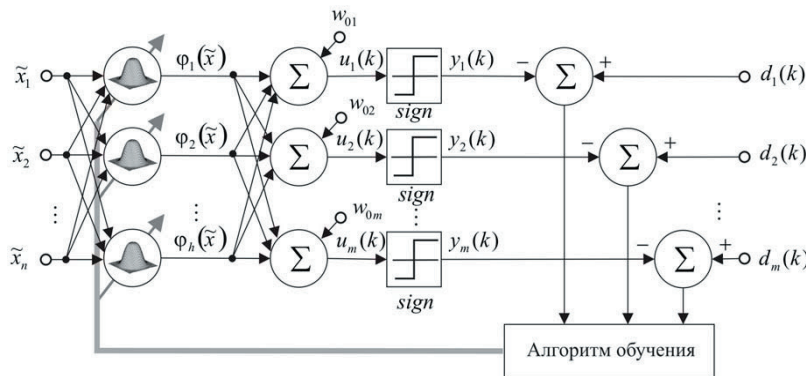


Рис. 2. Архитектура ИНС для решения задачи классификации

Введем в рассмотрение m ошибок обучения [24]

$$e_j(k) = d_j(k) - y_j(k) = d_j(k) - \text{sign } u_j(k) \quad (2)$$

и m локальных критериев [23]

$$E_j(k) = e_j(k)u_j(k) = d_j(k)u_j(k) - |u_j(k)| = (d_j(k) - \text{sign } w_j^T \phi(k))w_j^T \phi(k), \quad (3)$$

где $d_j(k)$ -обучающий сигнал, принимающий значение 1, если входной образ принадлежит заданному онтологическому классу, и -1 – в противном случае и кодирующий множество классов онтологии предметной области; $w_j = (w_{j1}, w_{j2}, \dots, w_{jh})^T$ – вектор синаптических весов, подлежащий определению; $\phi(k) = (\phi_1(k), \phi_2(k), \dots, \phi_h(k))^T$ – вектор выходных сигналов радиально-базисного слоя. Естественно, что в обучающей выборке обязательно должны присутствовать все возможные состояния из анализируемого массива данных $\{\tilde{x}(k)\}$, ибо в противном случае некоторые из «не предъявленных» структур сеть просто не распознает.

Для настройки синаптических весов будем использовать градиентную процедуру

$$w_{ji}(k+1) = w_{ji}(k) - \eta(k) \frac{\partial E_j(k)}{\partial w_{ji}(k)}, \quad (4)$$

$$j=0,1,\dots,m; i=1,2,\dots,h,$$

где $\eta(k)$ – параметр шага поиска, который с учетом (3) в векторной форме может быть записан в виде

$$w_j(k+1) = w_j(k) + \eta(k)e_j(k)\phi(k) = w_j(k) + \eta(k)(d_j(k) - \text{sign } w_j^T(k)\phi(k))\phi(k), \quad (5)$$

$$j=0,1,\dots,m$$

Вводя общий критерий классификации

$$E(k) = \sum_{j=1}^m E_j(k) = - \sum_{j=1}^m e_j(k)u_j(k), \quad (6)$$

можно записать алгоритм одновременного обучения с учетом (6) всех синаптических весов каждой RBFN в форме

$$W(k+1) = W(k) + \eta(k)(d(k) - \text{sign } W(k)\phi(k))\phi^T(k), \quad (7)$$

где

$$\text{sign}(u_1(k), u_2(k), \dots, u_m(k))^T = (\text{sign } u_1(k), \text{sign } u_2(k), \dots, \text{sign } u_m(k))^T,$$

$$d(k) = (d_1(k), d_2(k), \dots, d_m(k))^T,$$

$$W(k+1) = \begin{pmatrix} w_1^T(k) \\ w_2^T(k) \\ \vdots \\ w_m^T(k) \end{pmatrix} - (m \times h) - \text{матрица настраиваемых синаптических весов.}$$

весов.

Работа в условиях «зашумленных» данных требует, чтобы параметры, задающие шаг обучения, удовлетворяли условиям Дворецкого [25], т.е. по сути, превращали алгоритмы настройки в процедуры стохастической аппроксимации.

С этой целью алгоритм обучения (9) можно модифицировать к виду

$$\left\{ \begin{aligned} w_j(k+1) &= w_j(k) + \frac{d_j(k) - \text{sign } w_j^T(k)\phi(k)}{\eta(k)} \phi(k), j=1,2,\dots,m, \\ \eta(k) &= \eta(k-1) + \|\phi(k)\|^2, \eta(0) = 1, \end{aligned} \right. \quad (8)$$

являющемуся расширением процедуры идентификации Гудвина-Рэмеджа-Кэйнеса [26] на задачу обучения ИНС.

5. Результаты экспериментальных исследований

Предложенный метод семантического аннотирования текстовых документов, основой которого является HRBFN, оценивался на выборке текстов, принадлежа-

щих к разным онтологическим классам узкоспециализированной предметной области (30 признаков, 100 объектов). При этом для обычной RBFN число нейронов составило бы величину порядка $e^{30} \approx 1,0687 \times 10^{13}$, что не позволило бы решить задачу на заданной обучающей выборке.

Преобразование корпуса текстов в векторное пространство осуществлялось на основе статистической оценки TFIDF с использованием WordNet 2.0 (для формирования признаков рассматривались только существительные, стоп-слова были удалены). После получения выборки исходных текстовых объектов применялся принцип построения разности гистограмм признаков для исключения «голода» нейронной сети из-за большой разреженности данных (большого числа нулей в данных).

В ходе эксперимента рассматривалось, прежде всего, качество работы HRBFN. На рис. 3 приведен график зависимости между количеством эпох и точностью предложенного метода. Также в табл. 1 представлены результаты сравнения работы методов на основе стандартной RBFN и HRBFN для одинакового количества признаков.

При этом HRBFN содержит 2 слоя, первый слой образован тремя RBFN с 7 входами. Второй слой представляет собой RBFN, на вход которой подаются выходы предыдущего слоя (значения у первого слоя) – 9 входов и 3 выхода. Для обучения использовалось 75% исходных данных.

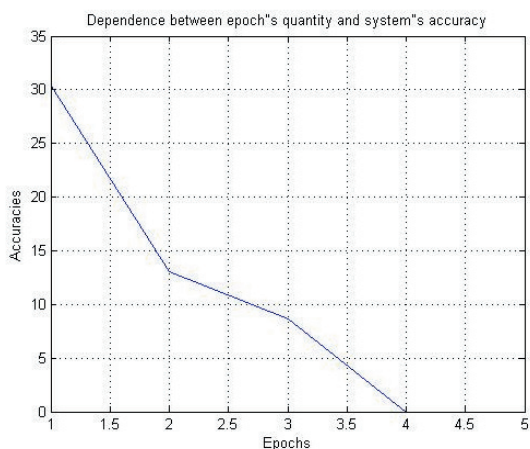


Рис. 3. График, иллюстрирующий точность обучения системы на каждой эпохе

Таблица 1

Сравнение результатов на исходной выборке

Модель	Ошибка распознавания на тестирующей выборке, %
RBFN	10,2
HRBFN	3,33

Стоит отметить, что в данном случае скорость настройки синаптических весов для второго слоя оказалась выше. Кроме того, в ходе экспериментальных исследований было установлено, что предложенный метод позволяет понизить пространство признаков,

поступающих на вход отдельной RBFN каждого слоя иерархии, что исключает проблему «проклятия размерности» и дает возможность повысить качество извлечения знаний из текстовых источников при ограниченной выборке.

6. Выводы

Предложена иерархическая радиально-базисная нейронная сеть с многослойной архитектурой, которая в каждом узле использует однотипную радиально-базисную нейронную сеть, что позволяет понизить количество признаков, поступающих на вход каждого слоя, и, таким образом, исключает проблему «проклятия размерности» при ограниченной обучающей выборке. В работе предложен алгоритм обучения нейронной сети с многослойной архитектурой.

Также рассмотрен подход семантического аннотирования, в основе которого лежит иерархическая радиально-базисная нейронная сеть. Данный метод позволяет решить проблему извлечения знаний из текстовых источников с учетом максимального числа релевантных признаков для построения семантических аннотаций текстовых документов узкоспециализированной предметной области, предусматривает возможность обработки информации по мере ее поступления в последовательном режиме, характеризуется простотой реализации.

Литература

1. Uren V. Semantic annotation for knowledge management: Requirements and a survey of the state of the art / V. Uren, Ph. Cimiano, J. Iria, S. Handschuh, M. Vargas-Vera, E. Motta, F. Ciravegna // Web Semantics: Science, Services and Agents on the World Wide Web. – 2006. – V. 4 – N. 1. – P. 14–28.
2. Chen H. Knowledge Management Systems: A Text Mining Perspective / H. Chen // Arisona: Knowledge Computing Corporation. – 2004. – 50 p.
3. Хорошевский В.Ф. Пространства знаний в сети Интернет и Semantic Web / В.Ф. Хорошевский // Искусственный интеллект и принятие решений. – 2008. – 1. – С. 80–97.
4. Al-Khalifa H. The evolution of metadata from standards to semantics in E-learning applications / H. Al-Khalifa, H. Davis // Proceedings of the 17th conference on Hypertext and hypermedia table of contents. – Odense, Denmark. – 2006. – P. 69–72.
5. Resource Description Framework: Overview. [Электронный ресурс] // Режим доступа: <http://www.w3.org/RDF/>.
6. The Dublin Core Metadata Initiative. [Электронный ресурс] // Режим доступа: <http://dublincore.org/>.
7. FOAF project. [Электронный ресурс] // Режим доступа: <http://www.foaf-project.org/>.
8. Simple Knowledge Organization System: Intoduction. [Электронный ресурс] // Режим доступа: <http://www.w3.org/2004/02/skos/intro>.

9. Hearst M. Automatic acquisition of hyponyms from large text corpora / M. Hearst // Proceedings of the 14th conference on Computational linguistics. – 1992. – V. 2. – P. 539–545.
10. Giles L. Automatic document metadata extraction using Support Vector Machines / L. Giles, E. Manavoglu, H. Zha // Proceedings of the 2003 Joint Conference on Digital Libraries. – 2003. – P. 37–48.
11. Wu Yu-Ch. Extracting named entities using Support Vector Machines / Yu-Ch. Wu, T.-K. Fan, Y.-Sh. Lee, Sh.-J. Yen // Lecture Notes in Computer Science. – 2006. – 3886. – P. 91–103.
12. Xu K. Self-teaching semantic annotation method for knowledge discovery from text / K. Xu, S. Liao, R. Lau, L. Liao, H. Tang // HICSS '09. – 2009. – P. 1–7.
13. Seymore K. Learning for information extraction learning Hidden Markov Model structure for information extraction / K. Seymore, A. McCallum, R. Rosenfeld // AAAI'99 Workshop on Machine Learning. – 1999. – P. 37–42.
14. McCallum A. Maximum entropy Markov Models for information extraction and segmentation / A. McCallum, D. Freitag, F. Pereira // 17th International Conference on Machine Learning. – 2000. – P. 591–598.
15. Grilheres B. A platform for semantic annotations and ontology population using conditional random fields / B. Grilheres, S. Canu, C. Beauce, S. Brunessaux // The 2005 IEEE/WIC/ACM International Conference. – 2005. – P. 790–793.
16. Tang J. Tree-structured conditional random fields for semantic annotation / J. Tang, M. Hong, J. Li, B. Liang // Lecture Notes in Computer Science. – 2006. – 4273. – P. 640–653.
17. Salton G. Term-weighting approaches in automatic text retrieval / G. Salton, C. Buckley // Information Processing & Management. – 1988. – 24(5). – P. 513–523.
18. Chifu E. St. A neural model for unsupervised named entity classification / E. St. Chifu, V. R. Chifu. // The IEEE International Conference on Intelligent Agents, Web Technologies and Internet Commerce. – 2008. – P.1073–1078.
19. Todorov K. Mining concept similarities for heterogeneous ontologies / K. Todorov, P. Geibel, K.-U. K hnberger // Lecture Notes in Computer Science. – 2010. – 6171. – P. 86-100.
20. Habib M. Addressing scalability issues of named entity recognition using Multi-class Support Vector Machines/ M. Habib // World Academy of Science, Engineering and Technology. – 2008. – 37. – P. 69-78.
21. Бодянский Е.В. Искусственные нейронные сети: архитектуры, обучение, применения / Бодянский Е.В., Руденко О.Г. – Харьков: ТЕЛЕТЕХ, 2004. – 372 с.
22. Haykin S. Neural Networks. A Comprehensive Foundation / S. Haykin – Upper Saddle River, N. J.: Prentice Hall, Inc, 1999. – 842 p.
23. Shynk J.J. Performance surfaces of a single-layer perceptron / J.J. Shynk. – IEEE Trans. on Neural Networks. – 1990. – 1. – P. 268–274.
24. Бодянский Е.В. Нейро-фаззи сети Петри в задачах моделирования сложных систем / Бодянский Е.В., Кучеренко Е.И., Михалев А.И. – Днепропетровск: Системные технологии, 2005. – 311 с.
25. Dvoretzky A. On stochastic approximation / A. Dvoretzky // Proc. 3-rd Berkley Symp. Math. Statistics and Probability. – 1956. – 1. – P. 39–55.
26. Goodwin G.C. A globally convergent adaptive predictor / G.C. Goodwin, P.J. Ramadge, P.E. Caines // Automatica. – 1981. – 1. – P. 135–140.