

Выходом из данной ситуации служит терминальный доступ с рабочего ПК к компьютеру со всеми необходимыми программами. Для обеспечения такого доступа необходимо лишь наличие клиентской части ПО Citrix и данные пользователя для подключения к серверу Citrix.

Плюсами данного способа является возможность использования одного и того же приложения несколькими пользователями одновременно, независимо друг от друга.

**Список использованных источников:**

1. Харун Даниил, Создание бездисковых терминалов Windows RDP (Citrix ICA) на базе FreeBSD.
2. Информационные технологии и управление : материалы 49 науч. конф. аспирантов, магистрантов и студентов, Минск, 6–10 мая 2013 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. – Минск : БГУИР, 2013. – 103 с.
3. William Manning, Citrix Certified Administrator for Citrix XenApp 5 for Windows Server 2008 Certification Exam Preparation Course, London, UK, UK ©2010

## **ВНУТРЕННЯЯ И ВНЕШНЯЯ ОЦЕНКА КОНВЕРСИИ СЛОВА В ВЕКТОР**

*Игнатъев Е.Д.*

*Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь*

*Иванов Н.Н. – к.т.н., доцент*

В последнее время количество исследований в сфере семантики неуклонно растёт. Для проведения таких исследований языка, зачастую прибегают к использованию конвертированных в вектор слов. Но при проведении конвертации могут встречаться различные ошибки которые в конечном итоге повлияют на точность и качество исследования. Поэтому необходимо проводить оценку такой конвертации.

В 2013 году компания состоялась релиз инструмента word2vec предлагая новый подход в представлении слова [1]. В 2015 году была проведена систематизация всех методов оценки и их разделение на *внутренние* и *внешние* [2].

Методы внутренней оценки — это эксперименты, в которых представления слов сравниваются с человеческими суждениями о словесных отношениях. Созданные вручную наборы слов часто используются для получения оценок человека, а затем эти оценки сравниваются с вложениями слов (этот метод сбора суждений называется абсолютной внутренней оценкой). Методы внутренней оценки-это эксперименты, в которых вложения слов сравниваются с человеческими суждениями о словесных отношениях. Созданные вручную наборы слов часто используются для получения оценок от человека, а затем эти оценки сравниваются с представлениями слов (этот метод сбора суждений называется абсолютной внутренней оценкой).

Иногда людей просят оценить качество вложений слов непосредственно, например, когда разные модели производят разные суждения о словесных отношениях, и задача человека — сказать, какая модель работает лучше (такой метод называется сравнительной внутренней оценкой). Сравнительная внутренняя оценка позволяет не оценивать абсолютное качество вложений, а находить наиболее адекватные вложения в заданном множестве.

Большинство методов абсолютной внутренней оценки предназначены для сбора данных, являющихся результатом сознательных процессов в человеческом мозге (другими словами, у оценщиков есть время подумать над ответами). Следовательно, существует вероятность того, что такие ответы предвзяты определенными субъективными факторами (например, из-за отсутствия четкого определения значения каждый человек интерпретирует отношения слов по-своему, внося вариативность в оценки). И неясно, действительно ли сознательные оценки способны сообщить структуру семантики на естественном языке.

Методы внешней оценки основаны на возможности представления слов в качестве векторов признаков алгоритмов контролируемого машинного обучения (например, модели максимальной энтропии), используемых в одной из различных задач обработки естественного языка. Производительность контролируемой модели (измеряемая в наборе данных для задачи обработки естественного языка) функционирует как мера качества представления слов. Некоторые исследования предполагают, что представление слов, показывающее хороший результат по одной задаче, покажет хороший результат по другим, а результаты встраивания слов по разным задачам коррелируют, определяя какой-то глобальный оценочный балл для дистрибутивной семантики [3].

Представление слов, вероятно, может быть использовано практически в любой задаче обработки естественного языка, и исследователи (описывая возможные варианты использования представлений слов в других последующих задачах) не упоминают задачу оценки представлений

слов. Тем не менее, согласно определению внешней оценки, приведенному выше, любая последующая задача может рассматриваться в качестве метода оценки.

Конечно, внешняя оценка имеет определенные преимущества, и в тех случаях, когда встраивание слов предполагается использовать только для решения конкретной задачи, оценка производительности контролируемой модели по этой задаче даст наиболее адекватную оценку производительности встраивания слов. Но внешняя оценка терпит неудачу, если вложения, которые вы хотите оценить, обучены служить в широком диапазоне различных задач, так как оценки производительности вложений word в различных задачах не коррелируют между собой.

**Список использованных источников:**

1. Mikolov et al., 2013a. Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space.
2. Schnabel et al., 2015. Schnabel, T., Labutov, I., Mimno, D. M., and Joachims, T. (2015). Evaluation methods for unsupervised word embeddings. // EMNLP, 2015, 298–307.
3. Bakarov A. A Survey of Word Embeddings Evaluation Methods // CoRR, vol. abs/1801.09536, 2018. Режим доступа: <http://arxiv.org/abs/1801.09536>.

## ПРИМЕНЕНИЕ АЛГОРИТМА СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ГРАФОВ ДЛЯ АНАЛИЗА ФОТОСНИМКОВ ТОПОЛОГИЙ ИНТЕГРАЛЬНЫХ МИКРОСХЕМ

Канаш В.Н.

Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь

Воронов А.А. – к.т.н., доцент

Описан алгоритм сегментации изображений на основе графов, а также рассмотрена возможность применения алгоритма для решения задач предварительной обработки изображений интегральных микросхем для последующего анализа с применением нейронных сетей.

Изображения топологий интегральных микросхем, полученные при увеличении отличаются большим разрешением, а также наличием множества мелких элементов, затрудняющих процесс анализа и контроля. Для обработки таких изображений в настоящее время используются различные методы интеллектуального анализа данных, в том числе нейронные сети. Однако для того, чтобы эффективно использовать нейронные сети для анализа изображений большого размера, необходимо выделять на них некоторые “точки интереса” - отдельные фрагменты, которые содержат в себе больше полезной информации, чем другие части изображения. Для поиска таких фрагментов в настоящее время обычно применяется метод скользящего окна. Он заключается в том, что в контексте решаемой задачи задается конфигурация «окна обработки» – двумерной области, охватывающей конечное множество пикселей входного изображения. В процессе обработки это окно смещается по изображению, последовательно занимая все возможные положения. Фрагменты изображения, покрываемые окном на текущем шаге используются в качестве входных данных для более сложных методов анализа. Недостатком данного подхода является то, что для прохождения окна по всему изображению и параллельного применения нейронной сети для каждого положения окна требуется дополнительное время. Также требуется решить задачу выбора оптимального масштаба окна.

В данном докладе рассматривается возможность применения алгоритма сегментации изображения на основе графов (*Graph-Based Segmentation*) для выделения отдельных фрагментов на изображениях топологий интегральных микросхем в целях подготовки изображения для последующего анализа с применением сверточных нейронных сетей.

Алгоритм сегментации изображения на основе графов (*Efficient Graph-Based Image Segmentation*) был впервые описан в 2004 году [1]. Он широко используется в задачах анализа изображений из-за простоты реализации и высокой производительности. Алгоритм использует представление изображения в виде взвешенного связного неориентированного графа, где каждый пиксель изображения является отдельной вершиной. Ребра графа связывают соседние пиксели между собой. Возможно 2 варианта построения такого графа: каждый пиксель связывается с 4 ближайшими соседями, либо с 8. В случае полутоновых изображений, вес (длина) ребра, соединяющего соседние вершины, выражается формулой (1)

$$w(v_i, v_j) = |I(p_i) - I(p_j)| \quad (1)$$