МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И КОГНИТИВНЫЕ ВЫЧИСЛЕНИЯ ПО ДАННЫМ СОЦИАЛЬНЫХ МЕДИА

@ Н.Т. Габдрахманова, М.А. Пильгун

gabdrakhmanova nt@rudn.ru, mpilgun@iling-ran.ru

УДК 517.518

DOI: 10.33184/mnkuomsh2t-2021-10-06.63.

Статья посвящена разработке различных типов моделей с использованием когнитивных вычислений по речевым данным акторов социальных медиа для выявления наличия/отсутствия социальной напряжённости в районах реализации градостроительных проектов на примере строительства транспортно-пересадочного узла «Нижегородская» в городе Москве (Россия). Эмпирической базой исследования послужили данные социальных сетей. Исследование проводилось с использованием трансдисциплинарного подхода, включало семантический анализ с использованием нейросетевых технологий, методов математической статистики и топологического анализа данных.

Ключевые слова: нейронная сеть, восприятие речи, топологический анализ данных

Mathematical modeling and Cognitive Computing from Social Media Data

The paper deals with the development of various types of models using cognitive computing based on speech data of social media actors to reveal the presence/absence of social tension in the areas where urban development projects are being implemented, as illustrated by the construction of the Nizhegorodskaya transport interchange hub in Moscow (Russia). The empirical base of the study was data from social networks. The research was carried out using a transdisciplinary approach, including semantic analysis using neural network technologies and methods of mathematical statistics, topological data analysis

Keywords: neural network, speech perception, topological data analysis

Когнитивные исследования в последнее время привлекают все большее внимание исследователей. Под когнитивными вычислениями обобщают различные технологические подходы, такие как искусственные

Работа выполнена при финансовой поддержке РФФИ (проект № 19-08-00261).

Габдрахманова Наиля Талгатовна, к.т.н., доцент, РУДН (Москва, Россия); Nailia Gabdrakhmanova (RUDN University, Moscow, Russia)

Пильгун Мария Александровна, д.ф.н., профессор, Институт языкознания РАН (Москва, Россия); Maria Pilgun (Institute of Linguistics, RAS, Moscow, Russia)

нейронные сети, нечеткие системы и эволюционные вычисления, для моделирования когнитивных способностей человека (мышление, обучение, рассуждение и т. д.) с помощью компьютерных моделей [1]. Когнитивные вычисления применяются в различных профессиональных сферах. Цель работы: разработка математических моделей с использованием когнитивных вычислений по речевым данным акторов социальных медиа для выявления наличия/отсутствия социальной напряжённости в районах реализации градостроительных проектов на примере строительства транспортно-пересадочного узла (ТПУ) «Нижегородская» в городе Москве (Россия). Эмпирической базой исследования послужили данные социальных сетей, микроблогов, блогов, мессенджеров, видео хосингов, форумов, Интернет-СМИ, тематических порталов и отзывы, посвященные реализации проекта ТПУ «Нижегородская». Date of collection: 1 Январь 2020 - 31 Март 2021. Сбор данных проводился с использованием алгоритмов Brand Analytics. Для интерпретации данных применялся трансдисциплинарный подход. Нейросетевая технология TextAnalyst 2.3. позволила выявить и проанализирована тематическую структуру базы данных, выделить семантическую сеть и проанализировать ядро семантической сети, провести обобщение и анализ словесных ассоциаций. Сентимент-анализ осуществлялся с помощью модуля определения тональности Eureka Engine. Контент-анализ проводился с помощью сервиса AutoMap. Скалярные данные, полученные с использованием TextAnalyst 2.3, можно рассматривать как временные ряды. Достаточно общей математической моделью для временных рядов $x(t)$ служит модель вида: $x(t)=u(t)+v(t)$, где $u(t)$ – детерминированная последовательность или систематическая составляющая, $v(t)$ – случайная составляющая. В работе, для оценки $u(t)$ использованы регрессионные модели [2], для оценки $v(t)$ – авторегрессионные модели. В работе исследованы возможности использования топологического анализа данных (ТДА) [3] для классификации ситуации по данным социальных медиа. По данным временных рядов построены персистентные диаграммы и барркodes с использованием библиотек языка R. Показано, что вычисления оценок характеристик Эйлера позволяют решать актуальные задачи.

Семантический анализ контента с использованием нейросетевых технологий показал нейтральное восприятие жителей реализации проекта, отсутствие социального стресса в районах строительства. Возглавляют рейтинг контента официальные сообщения, которые анонсируют планы московских властей по развитию транспортной системы в районе и ход строительства. Цифровые следы пользователей подтверждают нейтральное восприятие и оценки потребности жителей в изменении транспортной ситуации после окончания реализации проекта. Построенные

математические модели временных рядов по данным интенсивностей сообщений с нейтральным, позитивным и негативным отношением к строительству позволяют: 1) количественно оценить динамику этих рядов; 2) прогнозировать значения временных рядов на несколько шагов. Результаты исследований имеют важное значение для принятия управленческих решений при выявлении и предотвращении градостроительных конфликтов. Алгоритмы, разработанные с использованием методов ТДА, показывают потенциал использования этих методов для решения аналогичных задач.

Литература

1. *Portmann E., D'Onofrio S.* Cognitive Computin. Theorie, Technik und Praxis. — Wien: Springer Vieweg, 2020.
2. *Borovikov V.P., Ivchenko G.I.* Forecasting in the Statistica system in the Windows environment. Fundamentals of theory and intensive practice on a computer. — Moscow: Finance and Statistics, 2006
3. *Carlsson G., Zomorodian A.* Computing persistent homology //Proc. 20th Ann. Sympos.Comput. Geom, 2004, 347–356.