



Дона : электронный научный журнал. - 2019. - № 1. – 8 с. – Режим доступа: <http://www.ivdon.ru/ru/magazine/archive/n1y2019/5496>.

4. Начертательная геометрия: учеб. для вузов / Н. Н. Крылов, Г.С. Иконникова, В.Л. Николаев, Н.М. Лаврухина ; под. ред. Н.Н. Крылова. – 6 изд., перераб. и доп. – М. : Высш. шк., 1990. – 240 с.

В.В. Ворошилов, Н.В. Башарин

ИНФОРМАЦИОННАЯ СИСТЕМА КЛАСТЕРИЗАЦИИ И КЛАССИФИКАЦИИ ОБЪЕКТОВ В СЛАБОФОРМАЛИЗУЕМЫХ ОБЛАСТЯХ

(Самарский государственный технический университет)

С развитием информационных технологий в современном мире информатизация проникает во все области человечества. И все больше и больше растет гетерогенность информации доступной для обработки и анализа [1]. Появляются новые структуры данных описывающие объекты в слабоформализуемых областях. Эти структуры представляют новые возможности и интерес по их использованию в анализе таких объектов. Одними из таких задач являются задачи кластеризации и классификации.

Традиционные методы кластеризации и классификации работают с объектами, характеристики которых формально определены и заданы в виде определенных значений, формируя кластерные решения, например, на основе оценки расстояний между объектами и центрами кластеров. Такой подход не позволяет эффективно осуществлять анализ объектов в слабоформализуемых областях [2], где, например, по той или иной причине невозможно четко определить расстояние между объектами или нарушается правило треугольника. В связи с этим, актуальной задачей является разработка подходов и методов кластеризации и классификации, способных учитывать слабоформализованую природу объектов.

В работе рассматривается подход кластеризации и классификации объектов на основе анализа тепловых карт их поведения. В качестве апробации работы использованы две области применения задач:

- анализ поведения пользователя в интернете на информационном ресурсе;
- анализ позиционных действий игрока в командных видах спорта.

Классификация и кластеризация пользователя информационного ресурса является не тривиальной задачей. Частично это решается путем регистрации пользователя и требованием дальнейшей авторизации, но это покрывает только часть пользователей и покрывает только формальные характеристики пользователей, например, пол, возраст или платформу выхода в интернет [3]. Анализ тепловых карт поведения пользователя на ресурсе позволяет выявить неформальные характеристики пользователя, например, предметный интерес



или скорость анализа ресурса. Значения таких неформальных характеристик широко используются в маркетинге и в разработке информационных продуктов, рисунок 1 [4].

В работе предлагается подход кластеризации и классификации теплокарт поведения объектов на основе использования нейронной сети.

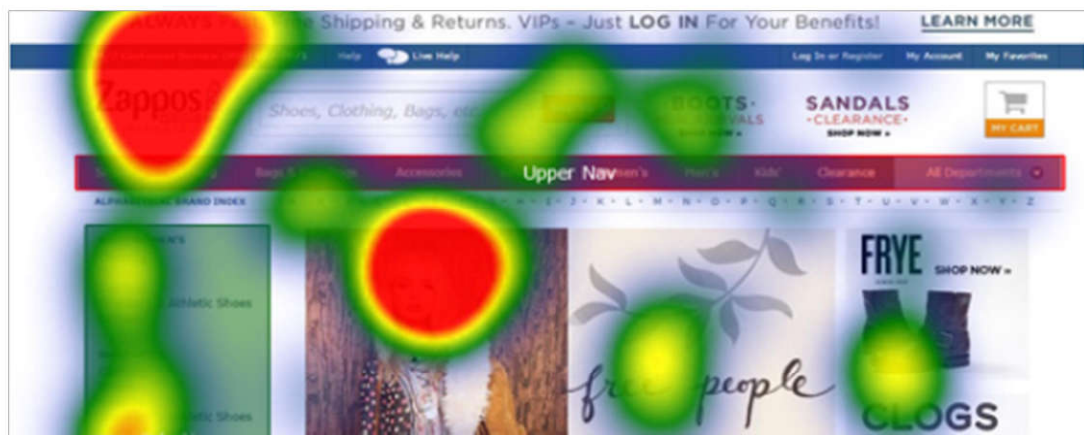


Рис. 1. Тепловая карта зрительного внимания пользователя Интернет магазина

Разработанная нейронная сеть имеет сверточную модель и состоит из четырех сверточных слоев, рисунок 2: двух слоев подвыборки; трех слоев регуляции «dropout»; слоя подготовки и преобразования данных; полносвязного слоя для классификации и выходного слоя. На выходе получается такое же количество нейронов, сколько определено классов для обучения. Для обучения нейронной сети используется размер мини-выборки равный ста двадцати восьми. В результате обучения доля правильных ответов для обучающей выборки составила 98,77%, а для тестовой выборки 88,89%.

```
1 model = Sequential()
2 model.add(Conv2D(32, (4, 4), padding='same',
3                 input_shape=(32, 32, 3), activation='relu'))
4 model.add(Conv2D(32, (4, 4), activation='relu', padding='same'))
5 model.add(MaxPooling2D(pool_size=(2, 2)))
6 model.add(Dropout(0.25))
7 model.add(Conv2D(64, (4, 4), padding='same', activation='relu'))
8 model.add(Conv2D(64, (4, 4), activation='relu'))
9 model.add(MaxPooling2D(pool_size=(2, 2)))
10 model.add(Dropout(0.25))
11 model.add(Flatten())
12 model.add(Dense(512, activation='relu'))
13 model.add(Dropout(0.5))
14 model.add(Dense(nb_classes, activation='softmax'))

1 model.compile(loss='categorical_crossentropy',
2               optimizer='adam',
3               metrics=['accuracy'])

1 model.fit(x_train, y_train,
2           epochs=60,
3           validation_split=0.1,
4           shuffle=True,
5           verbose=2)
```

Рис. 2. Модель нейронной сети в Google Colaboratory



На основе предложенной модели была разработана информационная система классификации и кластеризации поведения пользователя на информационно-новостном ресурсе «Новая газета», рисунок 3. Система позволяет классифицировать пользователя по его поведению с выделением вероятностной оценки для каждого возможного класса. Для этого ресурса можно выделить 4 основные квалификационные группы пользователей:

- читающих преимущественно заголовки;
- читающих преимущественно основной контент страницы;
- читающих преимущественно блок рекомендаций;
- читающих преимущественно комментарии.

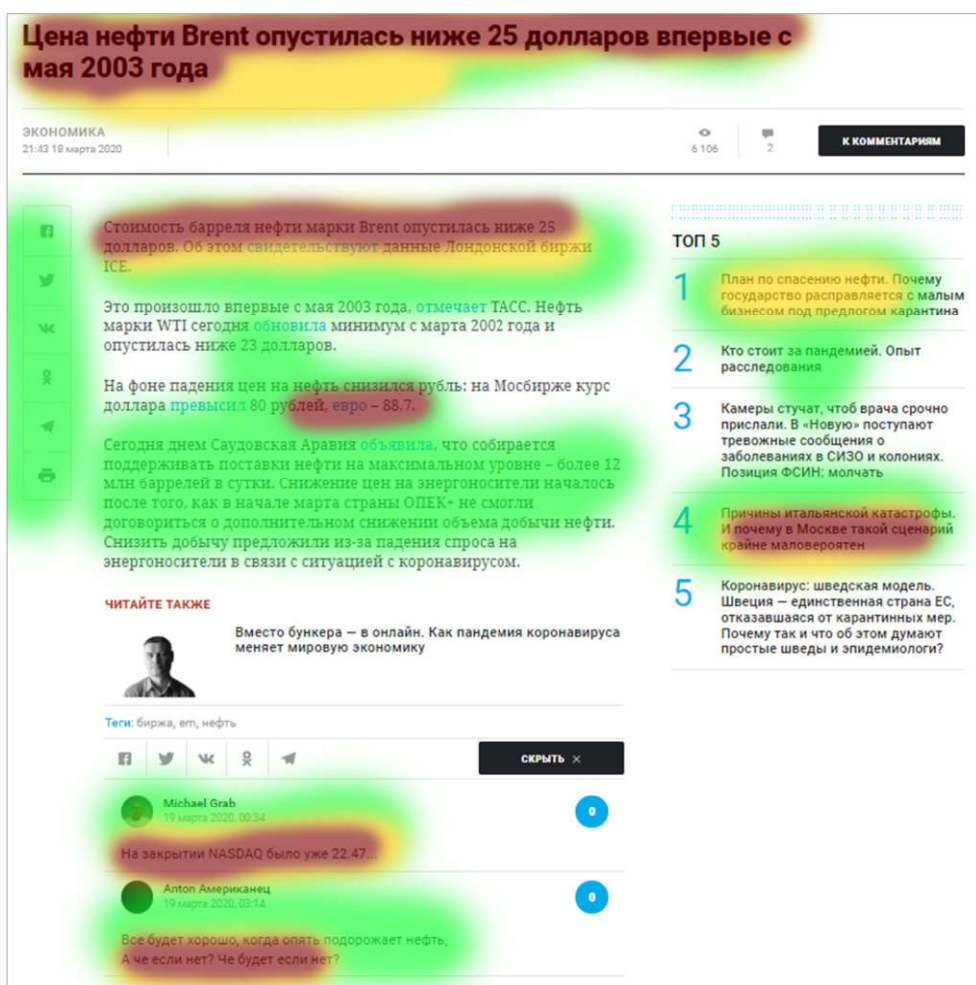


Рис. 3. Пример тепловой карты пользовательского визуального анализа новостной статьи «Новой газеты»

Информационная система представляет собой клиент-серверное приложение: серверная часть написана на языке `c#` с использованием «asp.net core» веб фреймворка; клиентская часть написана на языке `typescript` с использованием «angular» фреймворка. Система развернута на бесплатном хостинге «azure».

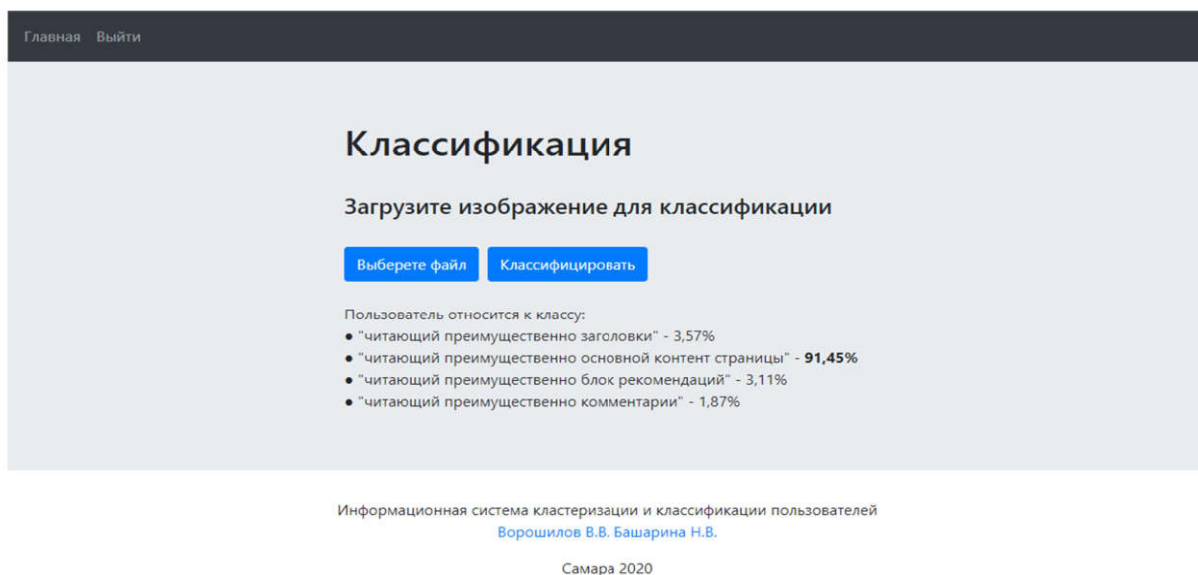


Рис. 4. Экран результата классификации пользователя информационного ресурса

Качество результата классификации и кластеризации поведения пользователей на информационном ресурсе, на основе теплокарт визуального анализа, определяется совокупностью оценок: качества обучения нейронной сети (определяется количеством обучающей выборки и составляет 98,7%); полнотой заполненности классификатора поведения пользователя (определяется экспертной оценкой и составляет 100%).

Литература

1. Рост объема информации - реалии цифровой вселенной // Технологии и средства связи. - 2013. - №1. - С. 24.
2. Определение ИИ. Определение слабоформализуемых задач и их примеры. Определение сложных систем [Электронный ресурс] /. — Электрон. текстовые дан. — Режим доступа: <https://studfile.net/preview/986657/>, свободный (дата обращения 08.04.2020)
3. Три основных модели поведения мобильных пользователей [Электронный ресурс] / Александр Мельник. — Электрон. текстовые дан. — Режим доступа: <https://ain.ua/2012/04/23/tri-osnovnyx-modeli-povedeniya-mobilnyx-polzovatelej/>, свободный (дата обращения 08.04.2020)
4. Школа интернет-маркетинга: Для чего нужны тепловые карты? [Электронный ресурс] /. — Электрон. текстовые дан. — Режим доступа: <https://lpgenerator.ru/blog/2016/01/11/shkola-internet-marketinga-dlya-chego-nuzhny-teplovye-karty/>, свободный (дата обращения 08.04.2020)