

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.383

Ермолкин
Владислав Павлович

Веб-приложение формирования новостей на основе модели пользователя

АВТОРЕФЕРАТ

на соискание академической степени
магистра информатики и вычислительной техники

по специальности 1-40 81 04 – Обработка больших объемов информации

Научный руководитель
Теслюк В. Н.
к.ф.-м.н., доцент

Минск 2020

КРАТКОЕ ВВЕДЕНИЕ

Сфера рекламы и технологий ее распространения развивается ежедневно и покрывает все сферы жизни человека. За последние полвека это одна из самых быстрорастущих и привлекательных отраслей деятельности. Создаются новые рабочие места, привлекаются инвестиции, а значит и новые идеи. Однако в связи с этим появляются и новые проблемы. Для маркетолога фраза “реклама везде” звучит в положительном ключе, однако его целевой аудиторией она воспринимается совершенно в противоположном ключе.

Важнейшей частью общества сегодня является интернет. Он является источником информации, развлечения, досуга и работы, то есть, сложно на данный момент представить себе человека, который не будет так или иначе вовлечен в эту сферу. Очевидно, что рекламные технологии адаптируются под такую относительно молодую и перспективную область рынка. Существует множество видов рекламы в интернете: контекстная, баннерная, SEO, вирусная, реклама в социальных сетях, тизерная, а также при помощи email. Последняя вызывает особенный интерес.

Одним из перспективных направлений области рекламы является email-маркетинг. Данный метод способствует созданию и укреплению качественных отношений с аудиторией, увеличивает показатели повторных покупок, а также дает возможность продвигать новые предложения и оперативно собирать мнения аудитории. Этот вид маркетинга дешев, создает коммуникацию между брендом и потребителем, а также расположен к возможности автоматизации процесса на основе информации о сегментах подписной базы. Однако как собрать эту информацию?

Самый простой и подверженный автоматизации способ сбора - это анкетирование. Его просто проводить в интернете, его можно разнообразить и индивидуализировать, чтобы удержать аудиторию. А также интересна идея автоматизации и оптимизации процесса анкетирования при помощи современных технологий, в частности, при помощи нейронных сетей.

Нейронные сети сделали огромный скачок в развитии в 21 веке. На сегодняшний день модели умеют говорить, понимать человека, а также определять его характерные признаки, такие как пол, возраст, характер, этническую принадлежность по его фотографии.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка веб-приложения, генерирующего индивидуализированную новостную сводку, основываясь на сформированной при помощи анкетирования модели пользователя.

Для достижения поставленной цели необходимо решить следующие задачи:

- 1) Исследовать рынок на наличие аналогов идеи.
- 2) Исследовать возможности сверточных нейронных сетей в задачах классификации в сфере машинного обучения, а также их эффективность.
- 3) Подобрать и предварительно преобразовать (при необходимости) обучающую выборку данных, предназначенную для настройки модели, классифицирующей нейронной сети.
- 4) Разработать архитектуру нейронной сети, а также программной системы, способной обучиться распознавать возраст человека по предоставленной фотографии.
- 5) Составить анкету, на основе которой будет формироваться модель пользователя.
- 6) Создать внешний вид письма, которое будет получать пользователь, грамотно используя пространство как для информации, так и для рекламы.

Объектом исследования являются сверточные нейронные сети, а также новостные рассылки и модель пользователя.

Предметом исследования является веб-приложение формирования новостной сводки, модели пользователя и анкеты, оптимизированной сверточными нейронными сетями.

Основной *гипотезой*, положенной в основу диссертационной работы, является возможность использования нейронных сетей для решения задач классификации фотографий по возрастному признаку, а также формирования оптимизированной модели пользователя и индивидуализированной анкеты, на основе которой модель формируется.

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, 4 глав, заключения, списка использованных источников и приложений. В первой главе представлен обзор предметной области. Во второй главе рассматриваются используемые технологии, в частности, используемый в качестве базы приложения фреймворк Ruby on Rails. В третьей главе рассматриваются алгоритмы обработки и анализа изображения, а именно сверточные нейронные сети; исследуются этапы работы модели сверточной нейронной сети. В четвертой главе находится проектирование и программная реализация: разработка модели определения возраста человека по фотографии, а также разработка приложения и интерфейса.

Общий объем работы составляет 75 страниц, из которых основного текста - 40 страниц, 15 рисунков на 7 страницах, 6 таблицы на 4 страницах, список использованных источников из 14 наименований на 3 страницах и 2 приложений.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В **первой главе** исследованы понятия новости, ее видов, структуры и источников, описывается структура веб-приложения, а также происходит обзор аналогов.

Во **второй главе** рассматриваются используемый в качестве базы приложения фреймворк Ruby on Rails, исследуется область разработки логических уровней архитектуры системы, а также процесс разработки веб-приложения.

В **третьей главе** рассматриваются алгоритмы обработки и анализа изображения, а именно сверточные нейронные сети; исследуются этапы

работы модели сверточной нейронной сети, такие как поиск лиц и оценка возраста.

В четвертой главе описывается разработка модели сверточной нейронной сети для определения возраста человека по фотографии, а также разработка приложения и интерфейса.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

В ходе работы над диссертационной работой была сформулирована и поставлена задача по разработке веб-приложения формирования новостей на основе модели пользователя. В частности, подзадачей является оптимизация процесса формирования модели пользователя при помощи сверточных нейронных сетей.

Данное веб-приложение было разработано с целью повышения эффективности инструменты email-маркетинга при помощи оптимизации нейронными сетями.

Для того чтобы правильно спроектировать систему, необходимо было грамотно проработать предметную область, также разработать основной процесс предметной области. Это позволило подробно разобраться и понять, что должно выполнять приложение.

Важным этапом явилось проектирование такого интерфейса системы, чтобы подготовленному пользователю было привычно пользоваться приложением.

Приложение может быть улучшено за счет дальнейшей разработки нейронных сетей как оптимизаторов формирования модели пользователя. В частности, можно обучить нейронную сеть определять интересы пользователя на основе его социальных сетей.