

С.Сейфуллиннің 125 жылдығына арналған «Сейфуллин окулары – 15: Жастар, ғылым, технологиялар: жаңа идеялар мен перспективалар» атты халықаралық ғылыми-теориялық конференциясының материалдары = Материалы Международной научно-теоретической конференции «Сейфуллинские чтения – 15: Молодежь, наука, технологии – новые идеи и перспективы», приуроченной к 125-летию С.Сейфуллина. -2019. - Т.II, Ч 1 - С.173-174

МАШИННОЕ ОБУЧЕНИЕ ДЛЯ РАСПОЗНАВАНИЯ РУКОПИСНЫХ ЦИФР

Альдош А.

В 2019 году в мире, наверное, не осталось такой отрасли деятельности человека, где не применялись бы технологии искусственного интеллекта. С появлением первых моделей искусственных нейронных сетей в 1980-х годах, представления о том, на что способна машина без внешнего воздействия человека, были подвергнуты кардинальным изменениям. Взяв за основу модель биологического нейрона, разработчики первых искусственных нейронных сетей попытались воплотить механизм работы человеческого мозга в виде компьютерной программы. Так началась первая эра расцвета искусственного интеллекта.

Несмотря на то, что искусственный интеллект ежедневно встречается нами в повседневной жизни, подавляющее большинство людей не имеют представления о том, как устроены искусственные нейронные сети изнутри и как они работают. Данная работа является попыткой пролить свет на внутреннее устройство нейронных сетей на примере разбора одного конкретного типа нейронных сетей – свёрточной нейронной сети. В качестве тестового задания будет использована известнейшая задача распознавания рукописных цифр[1].

Искусственный интеллект – это обширная область на стыке современных технологий (преимущественно информационных), которая охватывает машинное обучение, глубокое обучение, а также другие способы автоматизации процессов, не связанных с обучением. Более кратко эту область можно описать как автоматизацию интеллектуальных задач, выполняемых людьми.

Искусственная нейронная сеть – это математическое, или, в нашем случае, программное воплощение модели биологических нейронных сетей. Она представляет собой систему, состоящую из соединённых между собой искусственных нейронов. Такой нейрон получает на входе какую-то информацию, обрабатывает её в соответствии со своими предустановками и передаёт её далее к следующему слою, чтобы снова обработать полученные данные и т.д. На выходе мы имеем выходной слой, который состоит из такого количества искусственных нейронов, которое представляет собой количество возможных классов.

Машинное обучение – это область ИИ, которая объединяет подходы автоматизации, связанные с обучением искусственных нейронных сетей. Разработчики вводят входные данные и то, что должно получиться в процессе их обработки (ответы), а задачей такой системы является нахождение правил и закономерностей, которые в дальнейшем можно будет применить на новых данных с получением соответствующих ответов. Таким образом, система обучается автоматически, а не программируется разработчиком с помощью заранее подготовленного набора правил.

Глубокое обучение является более сложной и в то же время более эффективной разновидностью машинного обучения. Под глубиной понимается количество слоёв, содержащихся в ИНС, а не более глубокое восприятие или понимание системой информации.

Задание состоит в том, чтобы создать модель нейронной сети и обучить её, чтобы решать задачи распознавания рукописных цифр. Для этого мы создадим модель свёрточной нейронной сети с помощью TensorFlow и Keras, обучим её на популярном датасете (наборе данных) MNIST, а затем проверим на тестовом наборе данных.

В качестве бэкэнда будет использован по праву считающийся одним из самых популярных фреймворков для машинного обучения – фреймворк компании Google TensorFlow. Для ускорения процесса разработки модели и удобства настройки будет использован высокоуровневый API Keras, который в качестве бэкэнда официально поддерживает TensorFlow, Theano и др. Язык – Python 3[2].

Модель состоит из 8 слоёв, 2 из которых – свёрточные. Свёрточный слой состоит из набора карт признаков, у каждой из которых есть синаптическое ядро. Оно представляет собой фильтр, который проходит по всей области предыдущей карты и обнаруживает определённые признаки объекта, в данном случае – цифр. Ядро разделяет веса внутри слоя, что позволяет сократить общее число связей и позволяет обнаруживать одни и те же признаки на поверхности всего изображения.

Далее следует подвыборочный слой, целью которого является уменьшение размерности карт предыдущего слоя. Так как признаки на свёрточных слоях уже были выявлены, для последующих операций можно уплотнить изображение. На основе свёрточного слоя создаётся новая карта, функцией, отвечающей за этот процесс, обычно является ReLU. Здесь ведётся поиск признаков более высокого уровня.

Одним из последних слоёв является полносвязный обычный слой, целью которого является классификация и улучшение качества распознавания. Последний слой сети – выходной слой, связанный со всеми нейронами из предыдущего слоя. Количество нейронов в этом слое соответствует количеству возможных классов. На двух последних уровнях моделируется сложная нелинейная функция.

В итоге мы получаем модель, в которую остаётся загрузить подготовленные тренировочные данные и провести валидацию на тестовом наборе данных. После процесса обучения данная модель показывает

результат точности в 99,04%. Такой высокий показатель означает что данная модель была хорошо обучена и явления переобучения или overfitting не наблюдается.

Подобные по строению, но с большим количеством скрытых слоёв нейронные сети можно использовать для распознавания более сложных объектов, например, столов, стульев, книг, компьютеров, людей и т.д.

Список литературы

1. Hassan Ghiti Sarand, Bahram Karimi. Observer based robust neuro-adaptive control of non-square MIMO nonlinear systems with unknown dynamics//INTERNATIONAL JOURNAL OF COMPUTATIONAL INTELLIGENCE SYSTEMS. volume 10, Number 1, January - December 2017p. 23-33.

2. Доусон М. Програмируем на Python. - СПб.: Питер, 2014. - 416 с.: ил.

*Научный руководитель: к.б.н. кафедры ВТиПО
Иксымбаева Ж.С.*