

УДК 004.8

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В ЗАДАЧЕ АНАЛИЗА ПОЛУЧАЕМЫХ ДАННЫХ ОТ IOT-УСТРОЙСТВ

Турсуков Н.О. (Национальный исследовательский университет ИТМО)

Научный руководитель – к.т.н., доцент ФБИТ Попов И.Ю.

(Национальный исследовательский университет ИТМО)

В работе были изучены способы применения машинного обучения для анализа получаемых данных от устройств интернета вещей. Был представлен метод обработки и оценки получаемых данных. Анализ данных реализован с применением рекуррентных нейронных сетей, а также работа метода была протестирована на реальных наборах данных.

Введение. В условиях стремительного развития Интернета вещей, все больше предприятий стараются анализировать большое количество данных, собираемых устройствами. Анализ и оценка данных, получаемых от датчиков, могут как предупредить о возможной неисправности устройства, так и во предполагаемом нарушении работы аппарата, например злоумышленником. Целью классификации данных, при обработке информации, полученной с IoT-устройств, как правило, является оценка корректности работы данного устройства. Так, если собираемая акселерометром информация с автономного беспилотного робота, в какой-то момент начала значительно искажаться, то с большой вероятностью, возникла неисправность аппарата.

Основная часть. Рекуррентные нейронные сети являются частным случаем глубоких нейронных сетей, и более применимы к задаче анализа временных промежутков, поскольку представляют собой цепочку, обрабатывающую данные. В реализованном методе применяется нейронная сеть LSTM, с целью выявления необходимых последовательностей на большом объеме данных.

Распространенная проблема, с которой сталкиваются при анализе данных – это проблема обработки сигналов, а именно, как отфильтровать шум от базового сигнала. В работе проводилась оценка различных фильтров для обработки входных данных, с целью удаления шума и более точного прогнозирования. Фильтр Калмана показал наилучший результат, оставив большинство одиночных пиковых значений, в то же время убрав шум, отображающийся на графиках.

Реализация нейронной сети была проведена при помощи языка программирования Python 3.9 библиотеки keras. В результате построения модели рекуррентной нейронной сети LSTM, строится график, отображающий исторические показатели исследуемых данных, а также графически выделяет интервалы, которым модель присвоила соответствующий класс.

Данные для тестирования методов представляют собой реальные размеченные показатели, собираемые акселерометром и гироскопом, в количестве 10 млн. измерений. В ходе проведения тестов на рассмотренных данных, нейронная сеть показала хорошие результаты в задаче классификации, правильно обнаружив более 80% специфичных отрезков данных.

Выводы. В результате исследования были изучены способы применения машинного обучения для анализа получаемых данных. Был представлен метод, основанный на рекуррентной нейронной сети LSTM, позволяющий обрабатывать, а также проводить классификацию получаемых данных, с целью выявления подозрительной активности.

Турсуков Н.О. (автор)

Подпись

Попов И.Ю. (научный руководитель)

Подпись