

Отзыв
научного руководителя

на магистерскую диссертацию *СХИРТЛАДЗЕ Владимира Георгиевича*
«Методы анализа данных формата BigData»

Магистерская диссертация Схиртладзе В.Г. связана с изучением и применением в практической части работы методов анализа BigData. Цель диссертационной работы заключалась в систематизации и изучении методов анализа данных большой размерности и применении для их обработки одного из наиболее популярных современных алгоритмов машинного обучения – алгоритма случайного леса, предназначенного для бинарной классификации. Необходимость анализа и динамической обработки Big Data является характерной неотъемлемой чертой нашего времени, поэтому любое развитие методологии обработки Big Data является, бесспорно, актуальным.

В.Г. Схиртладзе проделана большая работа по систематизации методов машинного обучения и реализующих их программных средств. Он разобрался в особенностях алгоритма Random Forest (случайного леса), применил его в своей работе, причем использовал алгоритм RF в комбинации с методом главных компонент для снижения размерности Big Data без существенной потери информации, что позволило ему обработать собранные им данные за приемлемое время компьютерных вычислений. Для проведения вычислений была использована система Python, выбор которой аргументированно обоснован в работе.

Из недостатков работы можно отметить присутствие небольшого числа стилистических погрешностей и местами недостаточно детальное описание проведенных вычислений.

Общее количество страниц – 103. Работа написана достаточно ясно, логично, хорошим языком. По теме работы имеется публикация.