Отзыв научного руководителя на выпускную квалификационную работу аспиранта Ложкинса Алексейса по теме: «Исследование робастности кластеризации методами статистического и имитационного моделирования»

В выпускной квалификационной работе рассматривается задача сравнение кластеризаций на возмущенных данных и оценка устойчивости кластеризации по отношению к возмущению исходных данных. Предложены два критерия устойчивости кластеризации. Предложены новые алгоритмы проведения статистических экспериментов, разработано программное обеспечение. Проведены численные эксперименты на искусственных и реальных наборах данных. Полученные результаты показали эффективность предложенного подхода и разработанных алгоритмов.

Кроме того, в работе исследована задача робастного размещения хабов, в которой применяется алгоритм определения устойчивой кластеризации. Проведен обширный эксперимент на реальных данных, в котором найдено устойчивое количество хабов.

Основными теоретическими результатами проведенного исследования являются:

1. Разработка нового метода определения устойчивой кластеризации, для формулировки которого вводятся новые статистические показатели для кластеризации (частота сменяемости кластеризации, уровень ожидаемых отклонений кластеризации).
2. Эмпирически продемонстрированы работоспособность метода, а также показаны типы данных на которых алгоритм имеет преимущества перед другими подходами валидации кластеризации.
3. Разработан новый алгоритм определения устойчивой сети хабов, который основан на процедуре определения устойчивой кластеризации.

Результаты исследования могут быть использованы в статистическом анализе экономических показателей (автором продемонстрировано применение к анализу уровней безработицы и инфляции стран Европы), демографических показателей (автором продемонстрировано применение к анализу индексов смертности, рождаемости, миграции и др.), географическим показателям, в социальных исследованиях и других областях, в машинном обучении: кластеризация текстов, кластерная регрессия, сжатие данных и др.

Результаты работы опубликованы в трех статьях, индексированных в базах web of science, две из них индексированы в базе scopus, автор выступал на 6 конференциях, труды одной из них были индексированы в базах web of science и scopus. Имеется свидетельство о государственной регистрации программы для ЭВМ. Четыре статьи опубликованы в журналах из списка ВАК.

Работа, несомненно, заслуживает оценки «отлично».

Научный руководитель:

Профессор, д.т.н. /В. М. Буре/