**РЕЦЕНЗИЯ**

**на выпускную квалификационную работу обучающегося СПбГУ**

**Зиннатулиной Беллы Раифовны**

**по теме «Статистические методы кластеризации в больших объемах данных»**

Выпускная работа дипломанта посвящена исследованию проблематики кластерного анализа больших данных. В работе даны описательные характеристики больших данных, в том числе те, которые порождают проблемы при их анализе.

Дан теоретический обзор нескольких методов кластеризации, выстроена структура работы алгоритмов, указаны их особенности, однако отсутствует сравнение между алгоритмами.

Наибольшее внимание уделено методу cтохастической блочной модели, развитие которого является основной ценностью работы. Рассматриваемый метод находится на этапе становления, ему посвящено несколько статей, написанных в последние два года. Учитывая недавнее появление соответствующей теории, описывающие её статьи недостаточно структурированы и содержат путаницу в обозначениях. Дипломант изучил литературу, самостоятельно разобрался в теории, структурировал её, восстановил обозначения и соответствующийалгоритм.

Была написана работающая программа в среде R, производящая кластеризацию и проверку разделимости. Реализована визуализация результатов работы алгоритма. Работа содержит ряд примеров, поясняющих работу алгоритма и объясняющих смысл параметров, а также доказывающих эффективность теории разделимости.

Существенных замечаний к работе нет, а вышеописанные ее достоинства дают возможность оценить работу на «отлично».

Доктор физико-математических наук,

профессор кафедры прикладной математики

СПб Политехнического Университета Петра Великого

Г.Л. Шевляков

29. 05. 2017 