

Рецензия

на выпускную квалификационную работу

Степанченко Анастасии Александровны,

на тему: «Методы анализа данных для стандартизации и оптимизации хранения численных методов в экономике с применением к собранию программных кодов Quantnet»

Работа посвящена кластеризации кусочков программного кода архива Quantnet по их описаниям. Анастасия выполнила свою работу на английском языке.

Настя использует термин «big data», но, судя по графику 2, количество классифицируемых документов примерно равно семи тысячам. Для классификации используются разные алгоритмы и индексы.

Не считая небольших оформительских огрехов работа оформлена хорошо, с использованием системы L^AT_EX. Для реализации алгоритмов классификации используется статистическая среда R.

Основной недостаток работы следующий: непонятно, действительно ли за построенными классификациями стоит понятный человеку смысл? Например, график 14 выглядит симпатично, но осмыслен ли он? Нужен ли он пользователю архива Quantnet? На мой сторонний взгляд, любая иерархическая классификация была бы более интерпретируема. И графическое представление исходных данных очень скудное.

В целом работа Насти заслуживает хорошей оценки (7/10).

Рецензент: Борис Демешев

Дата: 14 ноября 2016