

ПРИМЕНЕНИЕ КЛАСТЕРНОГО АНАЛИЗА В ОЦЕНКЕ КАЧЕСТВА ГРУЗОВОЙ ТЕХНИКИ

Студент гр. 113511 Войнич К.Э.

Ст. преп. Кондратьева Н.А.

Белорусский национальный технический университет

Главное назначение кластерного анализа – разбиение множества исследуемых объектов и признаков на однородные в соответствующем понимании группы или кластеры. Таким образом, решается задача классификации данных и выявления соответствующей структуры в ней.

В условиях производства МАЗ для проведения сравнения накопленной информации о качестве грузовой техники наиболее эффективно использовать кластерный анализ в пакете STATISTICA. В общем случае он предназначен для объединения некоторых объектов в классы (кластеры) таким образом, чтобы в один класс попадали максимально схожие, а объекты различных классов максимально отличались друг от друга. Количественный показатель сходства получен заданным способом на основании данных, характеризующих объекты. Предварительно исходные данные нормируют в целях устранения различий в единицах измерения показателей. Разбиение на кластеры происходит с помощью некоторой метрики, например, евклидово расстояние.

Возникает задача оценить качество выпускаемой грузовой техники. С помощью кластерного анализа осуществлялось разбиение марок автомобилей и их описания на классы. По результатам наблюдений, попавшие в одну группу марки характеризовались как схожие образцы МАЗ и в процессе эксплуатации их можно заменить подобными.

Кластерный анализ проводился следующими методами: Joining (tree clustering), K – means clustering, Two-way joining. В результате проведенной работы было выявлено, что для анализа о качестве грузовой техники предпочтительно использовать метод Joining (tree clustering) и метод K – means clustering, а метод Two-way joining затруднителен в использовании из-за сложностей с интерпретацией полученных результатов.

Основное достоинство кластерного анализа в том, что он позволяет производить разбиение объектов не по одному параметру, а по целому набору признаков. Кроме того, кластерный анализ в отличие от большинства математико-статистических методов не накладывает никаких ограничений на вид рассматриваемых объектов, и позволяет рассматривать множество исходных данных практически произвольной природы.